Minireview

Mechanisms of Macromolecular Interactions Mediated by Protein Intrinsic Disorder

Sunghyun Hong[1,2], Sangmin Choi[1,2], Ryeonghyeon Kim[1,2], and Junseock Koh[1,*]

[1]School of Biological Sciences, Seoul National University, Seoul 08826, Korea, [2]These authors contributed equally to this work.
*Correspondence: junseockkoh@snu.ac.kr
https://doi.org/10.14348/molcells.2020.0186
www.molcells.org

Intrinsically disordered proteins or regions (IDPs or IDRs) are widespread in the eukaryotic proteome. Although lacking stable three-dimensional structures in the free forms, IDRs perform critical functions in various cellular processes. Accordingly, mutations and altered expression of IDRs are associated with many pathological conditions. Hence, it is of great importance to understand at the molecular level how IDRs interact with their binding partners. In particular, discovering the unique interaction features of IDRs originating from their dynamic nature may reveal uncharted regulatory mechanisms of specific biological processes. Here we discuss the mechanisms of the macromolecular interactions mediated by IDRs and present the relevant cellular processes including transcription, cell cycle progression, signaling, and nucleocytoplasmic transport. Of special interest is the multivalent binding nature of IDRs driving assembly of multicomponent macromolecular complexes. Integrating the previous theoretical and experimental investigations, we suggest that such IDR-driven multiprotein complexes can function as versatile allosteric switches to process diverse cellular signals. Finally, we discuss the future challenges and potential medical applications of the IDR research.

Keywords: allostery, coupled folding and binding, dynamic binding, intrinsically disordered proteins or regions, macromolecular complex, multivalent binding

INTRODUCTION

In the classical structure-function paradigm, a protein must fold into a well-defined three-dimensional structure in order to carry out its function. The landmark achievements underpinning this paradigm were the determination of the first three-dimensional protein structures (myoglobin and hemoglobin) by Kendrew (Kendrew et al., 1958) and Perutz (Perutz et al., 1960). In parallel, the elegant biochemical experiments by Anfinsen and colleagues further demonstrated that the native three dimensional structure of a protein is determined by its amino acid sequence (Anfinsen et al., 1961). For more than half a century, the structure-function paradigm has been one of the most fundamental frameworks in understanding complex biological processes at the molecular level. However, in the mid-1990s, with advances in bioinformatics, it was realized that a significant portion of the proteomes from various species contains proteins or regions in a protein with the amino acid contents distinct from those of ordered globular proteins (Romero et al., 1998; Wootton, 1994). Concurrently, nuclear magnetic resonance (NMR) experiments observed regulatory proteins that are disordered but fully functional under physiological conditions (Daughdrill et al., 1997; Kriwacki et al., 1996). Following these counter-intuitive discoveries, numerous biophysical and bioinformatic investigations have accumulated a substantial amount of evidence demonstrating the prevalence and the biological significance of disordered proteins, collectively termed intrinsically disordered proteins (IDPs) or regions in a protein (IDRs)

(Dunker et al., 2002; Uversky, 2002; van der Lee et al., 2014; Wright and Dyson, 1999). About a third of the eukaryotic proteome contains IDRs of 30 or more residues in length (Oates et al., 2013; Ward et al., 2004) that are particularly abundant in the proteomes associated with dynamic cellular processes including gene expression and signaling as well as with cancer and neurodegenerative diseases (Babu et al., 2011; Iakoucheva et al., 2002; Liu et al., 2006; Uversky et al., 2008; Xie et al., 2007).

In order to fully appreciate the functional role of IDRs in such diverse biological processes, it is required to understand how IDRs interact with other macromolecules. Of particular interest is the unique interaction features of IDRs that arise from conformational disorder and potentially provide functional advantages in specific cellular processes. In this review, we briefly overview the sequence and conformational features of IDRs and then discuss the mechanisms of the IDR-mediated macromolecular interactions distinct from those of ordered proteins. To expedite the discussion, the mechanisms are classified into three broad categories: i) coupled folding and binding, ii) dynamic binding, and iii) multivalent binding (Fig. 1). We also present the exemplary cellular processes that are exquisitely regulated by these interaction principles. In particular, we focus on the third principle that enables IDRs to drive assembly of multicomponent macromolecular complexes that potentially function as manifold allosteric switches to process and integrate diverse cellular signals. The more comprehensive physical and biological attributes of IDRs including multisite posttranslational modifications and liquid-liquid phase separations have been discussed in many excellent reviews (Csizmok et al., 2016; Uversky, 2018; van der Lee et al., 2014; Wright and Dyson, 2015; Wu and Fuxreiter, 2016).

SEQUENCE AND CONFORMATIONAL FEATURES

The amino acid sequences of IDRs exhibit compositional biases distinct from those of ordered proteins (Radivojac et al., 2007; Uversky et al., 2000; Weathers et al., 2004). In general, while the IDR sequences are enriched in polar and charged amino acids, they contain bulky hydrophobic amino acids in low abundance. These sequence characteristics directly dictates the conformational features of IDRs. One of the major driving forces in protein folding is the hydrophobic effect leading to compaction of a polypeptide chain into a stable tertiary structure that buries hydrophobic amino acids from the aqueous solvent (Chothia, 1974; Spolar and Record, 1994). The hydrophobic effect and other noncovalent interactions cooperatively stabilize the folded form, minimizing the free energy of the polypeptide best represented as funnel-shaped energy landscapes (Dill and Chan, 1997; Onuchic et al., 1995). In contrast, folding of an IDR into a unique and stable tertiary structure is unfavorable due to the low abundance of hydrophobic amino acids. Instead, they interact with the solvent through hydrophilic residues and dynamically interconvert among multiple conformational states (Csizmok et al., 2016; Dyson and Wright, 2005; Mukhopadhyay et al., 2007). Therefore, the typical energy landscape of an IDR is flat with multiple local minima (heterogeneity) that
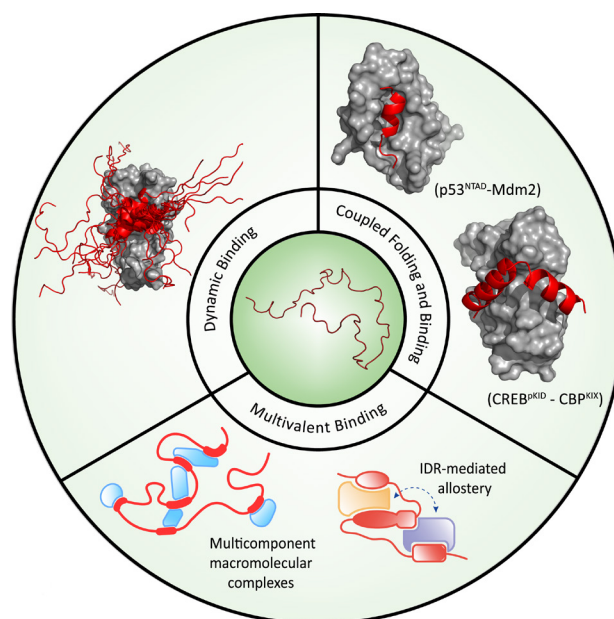

Fig. 1. Mechanisms of the interactions of IDRs with macromolecules. A schematic illustration is shown for coupled folding and binding, dynamic binding, and multivalent binding. Coupled folding and binding: interaction between pKID of CREB (red helix) and KIX of CBP (grey surface) (pdb code: 1kdx), interaction between NTAD of p53 (red helix) and Mdm2 (grey surface) (pdb code: 1ycr); Dynamic binding (fuzzy complex): interaction between cAD of GCN4 (red helix and loops) and ABD1 of Med15 (grey surface) (pdb code: 2lpb); Multivalent binding: schematic representation of multicomponent macromolecular complexes and of IDR-mediated allostery.

are separated by low activation energy barriers (fast interconversion) (Fisher and Stultz, 2011; Papoian, 2008; Wei et al., 2016).

COUPLED FOLDING AND BINDING

The majority of IDRs utilize linear peptide motifs to interact with their target macromolecules. These linear motifs are categorized into two broad groups, namely short linear motifs (SLiMs) and molecular recognition features (MoRFs) (Davey et al., 2012; Mohan et al., 2006; Oldfield et al., 2005; Tompa et al., 2014). While SLiMs contain 3-10 amino acids among which 3-4 residues are directly involved in the binding specificity, MoRFs are longer peptide motifs with 10-70 amphipathic amino acids. All MoRFs and selected SLiMs undergo disorder-to-order transitions upon binding to their targets, a process termed coupled folding and binding (Dyson and Wright, 2002; Spolar and Record, 1994). The type of the ordered structures adopted by IDRs includes α-helices, β-strands, and rigid loops. The coupled folding and binding mechanism has been extensively investigated for the interactions of the transcriptional co-activator CBP (CREB-binding protein) with its various binding partners (Dyson and Wright, 2016). A representative example is the binding of the phos-

phorylated KID (kinase-inducible domain) of CREB (cAMP-response element-binding protein) to KIX (KID interacting domain) of CBP. While pKID is disordered in the free form, it folds into two helices upon binding to KIX (Fig. 1) (Radhakrishnan et al., 1997; Sugase et al., 2007).

A suggested functional role of coupled folding and binding is to achieve the structural complementarity or specificity of the binding interface formed between an IDR and its target protein or DNA (Spolar and Record, 1994; Wright and Dyson, 1999; 2009). From the thermodynamic point of view, folding of a disordered motif into a well-defined structure is entropically unfavorable and must reduce the overall binding affinity. Such an entropic cost has been predicted to fine-tune the affinity between strong and weak binding in order to achieve both the binding specificity and reversibility required for robust and rapid signaling processes in the cell (Dyson and Wright, 2005; Fuxreiter et al., 2004; Zhou, 2012).

Some IDRs transiently adopt secondary structures in the free states that closely resemble the conformations of the target-bound states (Tompa, 2002). These observations invoked a conformational selection mechanism in which the preformed structures are the binding competent states and therefore selected out of other disordered states by the binding partners (Fuxreiter et al., 2004). A physiological significance of the preformed structures of IDRs was suggested from the investigation of the interaction between the E3 ubiquitin ligase Mdm2 and the N-terminal transactivation domain (NTAD) of the p53 tumor suppressor protein (Borcherds et al., 2014). In the free form, NTAD is mostly disordered with the low helical content, but folds into an amphipathic helix upon binding Mdm2. Mutating one of the conserved flanking prolines (Pro27) into alanine increased the population of the helical form and the affinity of NTAD for Mdm2 by an order of magnitude, likely due to the reduction in the entropic cost for folding. However, the increased affinity altered the stability of p53, impaired target gene expression, and ultimately caused failure to induce cell cycle arrest upon DNA damage. Such a dramatic effect strongly suggests that the abundance of the preformed structure is optimized in the conformational ensemble of NTAD for signaling fidelity. However, it may be ambiguous in some instances to interpret mutational or protein engineering data in part because these molecular interventions can affect not only the equilibrium for the formation of the preformed structure but also the binding equilibrium itself. Thus, in order to properly assess the functionality of the preformed structures, it is critical to combine spectroscopic and thermodynamic approaches to dissect the observed effects into these two (or more) contributions.

DYNAMIC BINDING

Because of the weak and reversible nature of the target binding, IDRs often retain the dynamic property even in the target-bound states, forming an ensemble of conformationally heterogeneous complexes (fuzzy complex) (Baker et al., 2007; Fuxreiter, 2018; Mittag et al., 2010; Tompa and Fuxreiter, 2008). A direct visual demonstration of the fuzzy complex was provided in the NMR investigation of the inter-

action between the transcription activator GCN4 and one of the multiple subunits of the Mediator complex (transcription co-activator) (Brzovic et al., 2011; Tuttle et al., 2018). The central activation domain (cAD) of GCN4 comprising 35 amino acids is disordered in the free form, but folds into an amphipathic α-helix in the residues 117 to 124 upon binding to the first activator binding domain (ABD1) of Med15. Remarkably, the folded helical region adopts multiple orientations on the ABD1 surface and forms a fuzzy complex (Fig. 1). The fuzziness arises from the surprisingly simple binding interface in which a few key hydrophobic residues of cAD are inserted into the shallow hydrophobic clefts on the ABD1 surface. The hydrophobic residues of cAD dynamically sample the multiple clefts on ABD1, leading to the multiple orientations of the helical segment. A subsequent study has shown that the peptides mimicking cAD of GCN4 but with the increased fuzziness bind ABD1 tighter than the wild type cAD (Warfield et al., 2014). Remarkably, these peptides enhanced the transcriptional activities of the Med15-dependent genes, underscoring the physiological significance of the fuzzy complexes.

Conformational fluctuations of IDRs in the target-bound states allow their limited regions to be transiently exposed to the solvent and make alternative interactions with other proteins as demonstrated in the extensive structural and biophysical investigations of the interaction between p27 and Cdk2/cyclin A (Fig. 2) (Galea et al., 2008; Grimmler et al., 2007; Lacy et al., 2004; Tsytlonok et al., 2019). The disordered KID (kinase inhibitory domain, residues 29-90) of p27 uses three subdomains, designated D1, D2, and 3_{10}, to bind the Cdk2/cyclin A complex and inhibit the catalytic activity of Cdk2. In particular, the D2 and 3_{10} subdomains adopt a beta hairpin (and an intermolecular beta-sheet) and a 3_{10} helix, respectively, in the bound state in which a tyrosine residue (Y88) in the 3_{10} helix is inserted into the active site of Cdk2. However, persistent flexibility of the D2 and 3_{10} subdomains allows these regions to dynamically sample the bound and solvent-exposed states. In particular, Y88 is transiently exposed during this "breathing" motion for phosphorylation by the non-receptor tyrosine kinase BCR-ABL. The phosphorylation unfolds and completely ejects the 3_{10} subdomain from the active site of Cdk2, which partially restores the catalytic activity of Cdk2. Subsequently, a threonine residue (T187) located near the C-terminus of p27 is phosphorylated by Cdk2 via a pseudounimolecular mechanism. In this phosphorylation step, the extremely flexible nature of the entire C-terminal IDR of p27 permits close proximity between the residue T187 and the Cdk2 active site. In turn, p27 with the phosphorylated T187 can be polyubiquitinated by the E3 ligase SCF[Skp2] and selectively degraded by the 26S proteasome. Finally, the fully active Cdk2/cyclin A complex drives progression from G_1 to S phase of the cell division cycle. In summary, p27 exploits its intrinsic disorder and large-scale dynamic motions in order to exquisitely regulate the catalytic activity of Cdk2/cyclin A through the multistep post-translational modification cascade.
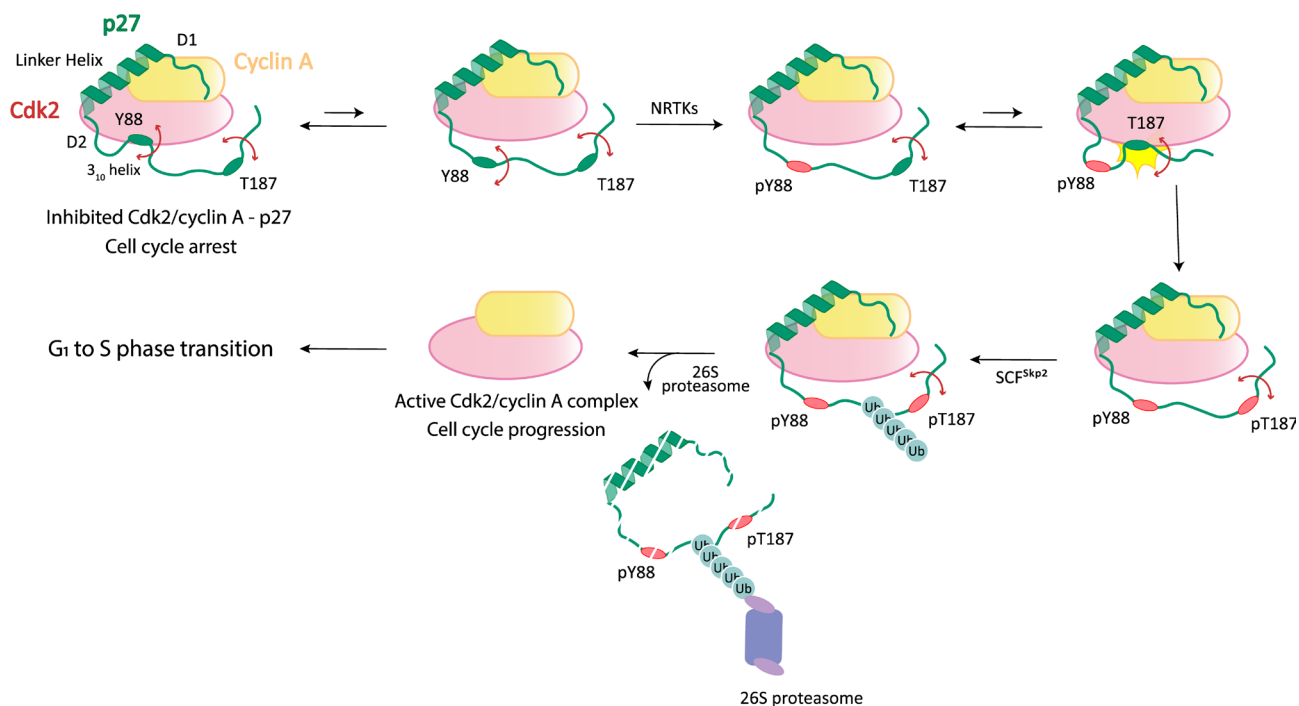
Fig. 2. Dynamic binding of p27 to Cdk2/cyclin A in the cell cycle regulation. A schematic illustration is shown for the sequential steps of the activation of Cdk2/cyclin A driving progression from G_1 to S phase of the cell division cycle. Conformational fluctuations of the subdomains of p27 (D2 and 3_{10}) are indicated by the double-sided arrows. See the text for the detailed discussion.

MULTIVALENT BINDING

Assembly of multicomponent macromolecular complexes

IDRs driving the formation of higher-order assemblies have caught much attention for the last decade (Fung et al., 2018; Fuxreiter et al., 2014; Wu and Fuxreiter, 2016). These IDRs present multiple MoRFs or SLiMs to make multivalent interactions with many binding partners (Fig. 1) (Cumberworth et al., 2013; Wright and Dyson, 2015). Multiple binding sites can be identical, promoting cooperative and high-affinity binding of an IDR to multiple copies of a target protein (Praefcke et al., 2004). Otherwise, multiple distinct MoRFs or SLiMs embedded in an IDR recruit diverse binding partners, enabling the IDR to function as an interaction hub (Cortese et al., 2008; Dunker et al., 2005; Haynes et al., 2006; Hegyi et al., 2007; van der Lee et al., 2014; Wright and Dyson, 2015). A molecular basis for the scaffolding property of IDRs was deduced to be their disordered and extended conformations that expose greater surface area available for potential interactions as compared to the ordered proteins of the same size (Dosztanyi et al., 2006; Gunasekaran et al., 2003). In addition, IDRs utilize linear motifs comprising relatively small numbers of amino acids for target binding (Tompa et al., 2014) while ordered proteins arrange numerous amino acids, often distant from each other on a linear sequence, into complex three-dimensional binding sites. Therefore, it should be more efficient to incorporate multiple target binding sites into an IDR than into an ordered protein (Cumberworth et al., 2013).

Axin is a representative hub protein interacting with various proteins in the Wnt, JNK, TGF-β, and p53 signaling pathways (Cortese et al., 2008). In the Wnt signaling pathway, β-catenin, casein kinase Iα, and glycogen synthetase kinase 3β (GSK-3β) bind the distinct motifs present in the central disordered region (residues ~200 to ~800) of axin to assemble the β-catenin destruction complex (Cortese et al., 2008; Xue et al., 2013). The assembly has been predicted to increase the local concentrations of these axin binding proteins and consequently facilitate the interactions among them (i.e., reduction of dimensionality) for the efficient phosphorylation and degradation of β-catenin (Noutsou et al., 2011; Xue et al., 2013). Of note, the binding of GSK-3β was observed to inhibit the interaction of the JNK signaling pathway protein MEKK1 with axin (Zhang et al., 2001), which suggests a negative coupling mechanism to prevent crosstalks among different signaling pathways. Since the binding sites of these two proteins on axin are non-overlapping, the observed inhibitory effect is not competitive. Furthermore, it was suggested that the facilitated interactions among the proteins in the β-catenin destruction complex may be driven not only by the aforementioned reduction of dimensionality but also by conformational changes of axin (Cortese et al., 2008). Collectively, it is plausible to hypothesize that allosteric coupling exists among target binding sites of an IDR hub for sophisticated regulation of the binding affinities and catalytic activities of hub-bound proteins.

Allosteric coupling mediated by IDRs in macromolecular complexes

In the structure-function paradigm, allostery has been conceived as a communication between two distinct ligand binding sites on a macromolecule connected by a well-defined structured pathway (Changeux, 2013; Monod et al., 1965). However, a theoretical framework has been proposed to show the feasibility of allosteric coupling between two sites within an IDR (Hilser and Thompson, 2007; Motlagh et al., 2014). This framework, termed ensemble allosteric model (EAM), elegantly demonstrates that coupling between two sites can be achieved by an intricate balance among the intrinsic stabilities of the two sites, their affinities for the respective ligands, and the interaction energy between the two sites. Furthermore, the EAM predicted allosteric coupling to be optimal when one or both of the ligand binding sites are disordered in the free forms.

The experimental demonstration for IDR-mediated allosteric coupling is beginning to emerge from a handful of systems: the interaction between glucocorticoid receptor and its target DNA (Li et al., 2017); the transcriptional regulation of the *phd/doc* toxin-antitoxin operon of bacteriophage P1 (Garcia-Pino et al., 2010); the interactions between the adenoviral oncoprotein E1A and host regulatory proteins (Fig.

Fig. 3. Allostery mediated by IDRs. (A) A schematic illustration is shown for positive allosteric coupling between the TAZ2 domain of CBP and the pocket domain of pRB that are bound to the distinct regions (N-terminal region and conserved region 1) of the adenoviral oncoprotein E1A. Truncation of the N-terminal region of E1A drives a transition in allostery from positive to negative coupling. (B) A schematic illustration is shown for the core complex of the NPC in which Nup53 utilizes the N and C-terminal IDRs to recruit Nic96 and Nup157, respectively. Binding of a karyopherin (Kap121) to the C-terminal IDR allosterically destabilizes the core complex by dissociating Nup157 from the nearby binding site and reducing the affinity of Nic96 for Nup53 by about an order of magnitude.

3A) (Ferreon et al., 2013); the interactions of nuclear pore proteins with transport factors (Fig. 3B) (Blus et al., 2019; Koh and Blobel, 2015). E1A is an intrinsically disordered hub protein that uses multiple promiscuous binding motifs to interact with numerous host proteins (Pelka et al., 2008). For instance, E1A uses its N-terminal region and two conserved regions (termed CR1 and CR2) to bind the TAZ2 (transcriptional adaptor zinc-finger 2) domain of CBP and the pocket domain of pRb (retinoblastoma protein) in order to reprogram the cell cycle and transcriptional regulation (Ferrari et al., 2008; Horwitz et al., 2008). For the E1A variant containing the N-terminal region and CR1, the binding sites of the two host proteins are distinct and positively coupled through disorder-to-order transitions to promote the ternary complex formation (Ferreon et al., 2013). However, truncation of the N-terminal region of E1A drives a striking transition in allostery from positive to negative coupling, which favors the formation of the binary complexes (Fig. 3A). The positive allosteric coupling was suggested critical to cooperatively recruit the two host proteins and facilitate the acetylation of pRb by the HAT (histone acetyl transferase) domain of CBP. In turn, the acetylation triggers degradation of pRb, causing uncontrolled cell cycle progression to S phase and proliferation of infected cells. Conversely, the negative coupling implies that the concentration and the overall activity of the ternary complex can be fine-tuned by the availability of the N-terminal region of E1A. The *in vivo* availability of the N-terminal region can be regulated by binding of other cellular proteins to this region. Therefore, in order to corroborate the physiological significance of the truncation-driven allosteric switch, it is important to test whether negative allostery can be induced by occupation of the N-terminal region of intact E1A by other proteins.

Recent thermodynamic and structural investigations of the interactions of nuclear pore proteins (nucleoporins) with transport factors (karyopherins) have demonstrated that interactions of karyopherins with IDRs of nucleoporins allosterically modulate the interaction networks among nucleoporins (Blus et al., 2019; Koh and Blobel, 2015). In particular, Nup53 utilizes its N and C-terminal IDRs to interact with other nucleoporins (Nic96 and Nup157) and assemble into the core of the nuclear pore complex (NPC). The C-terminal IDR of Nup53 interacts with a karyopherin (Kap121) as well, and the karyopherin binding allosterically destabilizes the nucleoporin interactions at both the N and C-terminal IDRs (Blus et al., 2019). The flexibility of the NPC core induced by the allosteric destabilization may be required to accommodate potential conformational changes in the central channel of the NPC during various transport events (Blus et al., 2019; Koh and Blobel, 2015).

These pioneering discoveries suggest a novel feature of allostery in multiprotein complexes assembled from IDRs (Fig. 4). Because of the reversible target binding and conformational fluctuation of IDRs, an IDR-based macromolecular complex exists as an ensemble of various states with similar free energy levels and low energy barriers among them (Fig. 4A). Each state has a unique conformation/function determined by the affinities of the bound protein subunits and cooperativities among them. At the same time, IDRs present
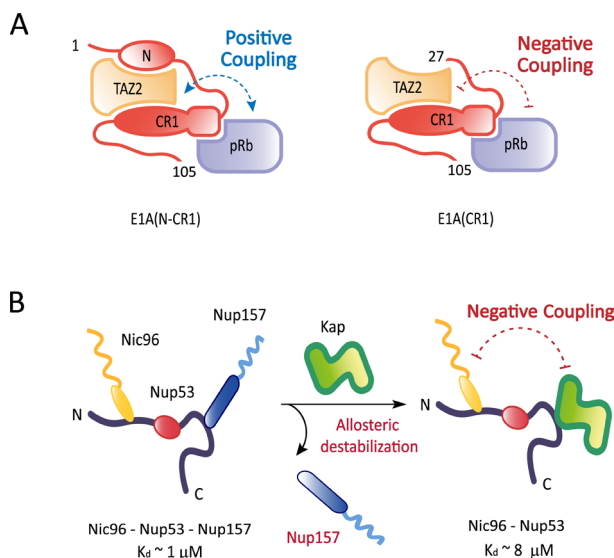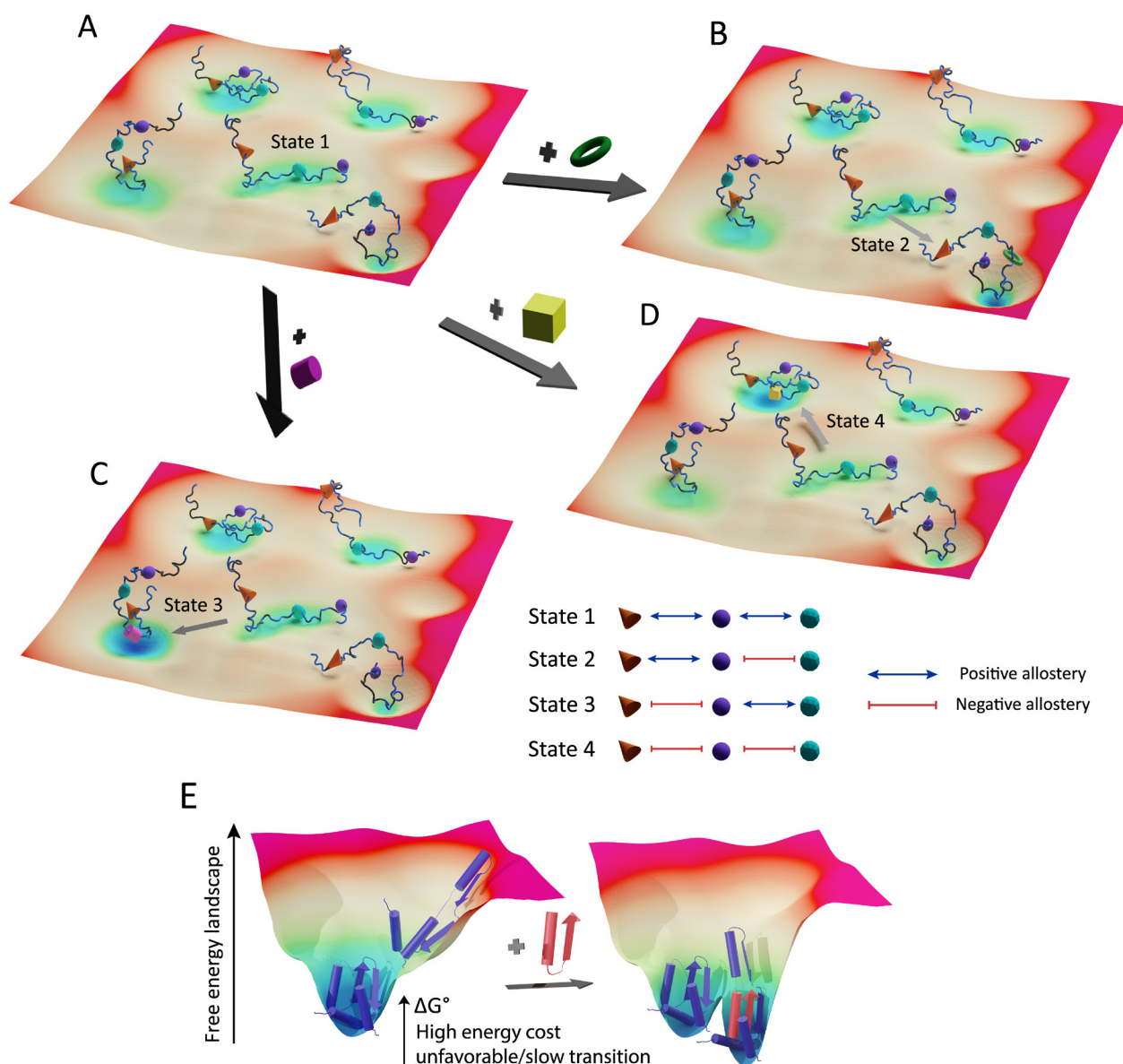
Fig. 4. Multifaceted allosteric coupling mediated by IDRs in macromolecular complexes. (A) A schematic illustration is shown for a hypothetical energy landscape of the ensemble of a multiprotein complex in which three protein subunits (orange cone, purple sphere, and green polyhedron) are bound to distinct linear motifs embedded in an IDR of the hub protein. Four different states of the complex are shown with similar free energy levels and low activation energy barriers among them (i.e., flat energy landscape). Each state exhibits a unique conformation determined by the binding affinities of the protein subunits and the cooperativities among them. Hypothetical (positive or negative) cooperativities were assigned to each state as shown in the appended schematic table. In the absence of external signals (protein binding, PTMs), state 1 is marginally more stable than others. (B-D) Upon binding of different external proteins (green torus, yellow cube, and purple cylinder) to the distinct regulatory sites on the IDR, the ensemble population shifts toward the specific conformational/functional states (state 2-4). Such a multifaceted allosteric shift is energy-efficient and fast due to the preconditioned heterogeneity and low energy barriers in the ensemble. (E) Hypothetical allosteric transition in an ordered protein in which two end states greatly differ in free energy level.

multiple regulatory sites to recognize diverse external signals (protein binding, PTMs) that allosterically shift the ensemble toward distinct conformational and functional states (Figs. 4B-4D). Because of the preconditioned heterogeneity and low energy barriers within the ensemble, the multifaceted allosteric modulation of the macromolecular complex is thermodynamically favorable and kinetically fast, particularly as compared to allosteric coupling in ordered proteins (Fig. 4E). In short, allostery mediated by IDRs confers versatility and agility on the macromolecular complexes that must promptly

process and adapt to the various external stimuli.

CONCLUSION

For the last two decades, our perception of IDRs has dramatically evolved from nonfunctional terminal tails or passive domain tethers to critical regulatory components in dynamic cellular processes. Despite the paradigm shift, it is far from complete to understand the staggering complexity of how IDRs interact with various targets and thereby regulate biological processes especially in the context of large macromolecular complexes such as the NPC and the Mediator complex. As discussed earlier (Fig. 4), IDRs may function not only as architectural scaffolds but also as manifold allosteric switches in these large complexes. In order to test this model, it is essential to purify intact proteins that contain both IDRs and ordered domains, which is a nontrivial technical problem. A theoretical challenge is to develop analytical tools quantifying the stoichiometries, binding constants, and cooperativities of non 1:1, multivalent interactions among IDRs and target proteins. In parallel, an integrative structural approach utilizing x-ray crystallography (target-bound motif structure), NMR (conformational ensemble), cryo-EM (global architecture), and low-resolution methods (e.g., light scattering, fluorescence resonance energy transfer [FRET], cross-linking) is required to explore changes in the conformational and functional states of the macromolecular complexes driven by IDR-mediated allostery. The combined effort will eventually make significant contributions in addressing the long-standing biological questions, for example, how the Mediator complex, enriched in IDRs, interacts with multiple transcription factors of enhanceosomes and accordingly modulates its conformation to recognize and activate specific promoters (Malik and Roeder, 2010).

Given the various regulatory roles of IDRs, it is not surprising that mutations affecting the property of IDRs are associated with many pathological conditions including cancer and neurodegenerative diseases (Uversky et al., 2008). Notably, a significant portion of dosage-sensitive genes, harmful to cells when overexpressed, was found to encode IDRs (Vavouri et al., 2009). This dosage-sensitivity is suggested to originate from the promiscuous binding property of IDRs. Overexpression of a promiscuous IDR involved in a specific signaling pathway would result in off-target interactions, disrupting a balance among signaling pathways. Hence, IDRs are attractive targets in treating devastating human diseases (Cheng et al., 2006). Indeed, recent efforts with optimization for enhancing the binding affinity and specificity have successfully designed small molecules targeting IDRs to inhibit their interactions with binding partners or substrates (Cheng et al., 2006; Joshi and Vendruscolo, 2015; Metallo, 2010). The representative examples include the oncogenic transcription factor c-Myc (Follis et al., 2008; Hammoudeh et al., 2009; Harvey et al., 2012), the nuclear protein 1 (Neira et al., 2017), and the protein tyrosine phosphatase 1B (PT1B) (Krishnan et al., 2014). In particular, an allosteric site in the C-terminal IDR of PT1B was targeted by a natural product to lock the protein in an inactive form (Krishnan et al., 2014). In conclusion, IDRs present the unparalleled variety and complexity in the inter-

action with macromolecules. Although extremely challenging, deconvoluting the complicated interaction features into quantitative molecular terms will culminate in discovering novel regulatory mechanisms of fundamental biological processes. In turn, such discoveries will serve as a firm foundation for future medical applications.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST

The authors have no potential conflicts of interest to disclose.

ORCID

Sunghyun Hong   https://orcid.org/0000-0003-4676-3313
Sangmin Choi   https://orcid.org/0000-0002-8061-6336
Ryeonghyeon Kim   https://orcid.org/0000-0003-4542-5339
Junseock Koh   https://orcid.org/0000-0002-2920-2656

REFERENCES

Anfinsen, C.B., Haber, E., Sela, M., and White, F.H., Jr. (1961). The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. Proc. Natl. Acad. Sci. U. S. A. *47*, 1309-1314.

Babu, M.M., van der Lee, R., de Groot, N.S., and Gsponer, J. (2011). Intrinsically disordered proteins: regulation and disease. Curr. Opin. Struct. Biol. *21*, 432-440.

Baker, J.M., Hudson, R.P., Kanelis, V., Choy, W.Y., Thibodeau, P.H., Thomas, P.J., and Forman-Kay, J.D. (2007). CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices. Nat. Struct. Mol. Biol. *14*, 738-745.

Blus, B.J., Koh, J., Krolak, A., Seo, H.S., Coutavas, E., and Blobel, G. (2019). Allosteric modulation of nucleoporin assemblies by intrinsically disordered regions. Sci. Adv. *5*, eaax1836.

Borcherds, W., Theillet, F.X., Katzer, A., Finzel, A., Mishall, K.M., Powell, A.T., Wu, H., Manieri, W., Dieterich, C., Selenko, P., et al. (2014). Disorder and residual helicity alter p53-Mdm2 binding affinity and signaling in cells. Nat. Chem. Biol. *10*, 1000-1002.

Brzovic, P.S., Heikaus, C.C., Kisselev, L., Vernon, R., Herbig, E., Pacheco, D., Warfield, L., Littlefield, P., Baker, D., Klevit, R.E., et al. (2011). The acidic transcription activator Gcn4 binds the mediator subunit Gal11/Med15 using a simple protein interface forming a fuzzy complex. Mol. Cell *44*, 942-953.

Changeux, J.P. (2013). 50 years of allosteric interactions: the twists and turns of the models. Nat. Rev. Mol. Cell Biol. *14*, 819-829.

Cheng, Y., LeGall, T., Oldfield, C.J., Mueller, J.P., Van, Y.Y., Romero, P., Cortese, M.S., Uversky, V.N., and Dunker, A.K. (2006). Rational drug design via intrinsically disordered protein. Trends Biotechnol. *24*, 435-442.

Chothia, C. (1974). Hydrophobic bonding and accessible surface area in proteins. Nature *248*, 338-339.

Cortese, M.S., Uversky, V.N., and Dunker, A.K. (2008). Intrinsic disorder in scaffold proteins: getting more from less. Prog. Biophys. Mol. Biol. *98*, 85-106.

Csizmok, V., Follis, A.V., Kriwacki, R.W., and Forman-Kay, J.D. (2016). Dynamic protein interaction networks and new structural paradigms in

signaling. Chem. Rev. *116*, 6424-6462.

Cumberworth, A., Lamour, G., Babu, M.M., and Gsponer, J. (2013). Promiscuity as a functional trait: intrinsically disordered regions as central players of interactomes. Biochem. J. *454*, 361-369.

Daughdrill, G.W., Chadsey, M.S., Karlinsey, J.E., Hughes, K.T., and Dahlquist, F.W. (1997). The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, sigma 28. Nat. Struct. Biol. *4*, 285-291.

Davey, N.E., Van Roey, K., Weatheritt, R.J., Toedt, G., Uyar, B., Altenberg, B., Budd, A., Diella, F., Dinkel, H., and Gibson, T.J. (2012). Attributes of short linear motifs. Mol. Biosyst. *8*, 268-281.

Dill, K.A. and Chan, H.S. (1997). From Levinthal to pathways to funnels. Nat. Struct. Biol. *4*, 10-19.

Dosztanyi, Z., Chen, J., Dunker, A.K., Simon, I., and Tompa, P. (2006). Disorder and sequence repeats in hub proteins and their implications for network evolution. J. Proteome Res. *5*, 2985-2995.

Dunker, A.K., Brown, C.J., Lawson, J.D., Iakoucheva, L.M., and Obradovic, Z. (2002). Intrinsic disorder and protein function. Biochemistry *41*, 6573-6582.

Dunker, A.K., Cortese, M.S., Romero, P., Iakoucheva, L.M., and Uversky, V.N. (2005). Flexible nets. The roles of intrinsic disorder in protein interaction networks. FEBS J. *272*, 5129-5148.

Dyson, H.J. and Wright, P.E. (2002). Coupling of folding and binding for unstructured proteins. Curr. Opin. Struct. Biol. *12*, 54-60.

Dyson, H.J. and Wright, P.E. (2005). Intrinsically unstructured proteins and their functions. Nat. Rev. Mol. Cell Biol. *6*, 197-208.

Dyson, H.J. and Wright, P.E. (2016). Role of intrinsic protein disorder in the function and interactions of the transcriptional coactivators CREB-binding protein (CBP) and p300. J. Biol. Chem. *291*, 6714-6722.

Ferrari, R., Pellegrini, M., Horwitz, G.A., Xie, W., Berk, A.J., and Kurdistani, S.K. (2008). Epigenetic reprogramming by adenovirus e1a. Science *321*, 1086-1088.

Ferreon, A.C., Ferreon, J.C., Wright, P.E., and Deniz, A.A. (2013). Modulation of allostery by protein intrinsic disorder. Nature *498*, 390-394.

Fisher, C.K. and Stultz, C.M. (2011). Constructing ensembles for intrinsically disordered proteins. Curr. Opin. Struct. Biol. *21*, 426-431.

Follis, A.V., Hammoudeh, D.I., Wang, H., Prochownik, E.V., and Metallo, S.J. (2008). Structural rationale for the coupled binding and unfolding of the c-Myc oncoprotein by small molecules. Chem. Biol. *15*, 1149-1155.

Fung, H.Y.J., 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.

Fuxreiter, M. (2018). Fuzziness in protein interactions-a historical perspective. J. Mol. Biol. *430*, 2278-2287.

Fuxreiter, M., Simon, I., Friedrich, P., and Tompa, P. (2004). Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. J. Mol. Biol. *338*, 1015-1026.

Fuxreiter, M., Toth-Petroczy, A., Kraut, D.A., Matouschek, A., Lim, R.Y., Xue, B., Kurgan, L., and Uversky, V.N. (2014). Disordered proteinaceous machines. Chem. Rev. *114*, 6806-6843.

Galea, C.A., Wang, Y., Sivakolundu, S.G., and Kriwacki, R.W. (2008). Regulation of cell division by intrinsically unstructured proteins: intrinsic flexibility, modularity, and signaling conduits. Biochemistry *47*, 7598-7609.

Garcia-Pino, A., Balasubramanian, S., Wyns, L., Gazit, E., De Greve, H., Magnuson, R.D., Charlier, D., van Nuland, N.A., and Loris, R. (2010). Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity. Cell *142*, 101-111.

Grimmler, M., Wang, Y., Mund, T., Cilensek, Z., Keidel, E.M., Waddell, M.B., Jakel, H., Kullmann, M., Kriwacki, R.W., and Hengst, L. (2007). Cdk-inhibitory activity and stability of p27Kip1 are directly regulated by oncogenic

tyrosine kinases. Cell *128*, 269-280.

Gunasekaran, K., Tsai, C.J., Kumar, S., Zanuy, D., and Nussinov, R. (2003). Extended disordered proteins: targeting function with less scaffold. Trends Biochem. Sci. *28*, 81-85.

Hammoudeh, D.I., Follis, A.V., Prochownik, E.V., and Metallo, S.J. (2009). Multiple independent binding sites for small-molecule inhibitors on the oncoprotein c-Myc. J. Am. Chem. Soc. *131*, 7390-7401.

Harvey, S.R., Porrini, M., Stachl, C., MacMillan, D., Zinzalla, G., and Barran, P.E. (2012). Small-molecule inhibition of c-MYC:MAX leucine zipper formation is revealed by ion mobility mass spectrometry. J. Am. Chem. Soc. *134*, 19384-19392.

Haynes, C., Oldfield, C.J., Ji, F., Klitgord, N., Cusick, M.E., Radivojac, P., Uversky, V.N., Vidal, M., and Iakoucheva, L.M. (2006). Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. PLoS Comput. Biol. *2*, e100.

Hegyi, H., Schad, E., and Tompa, P. (2007). Structural disorder promotes assembly of protein complexes. BMC Struct. Biol. *7*, 65.

Hilser, V.J. and Thompson, E.B. (2007). Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. Proc. Natl. Acad. Sci. U. S. A. *104*, 8311-8315.

Horwitz, G.A., Zhang, K., McBrian, M.A., Grunstein, M., Kurdistani, S.K., and Berk, A.J. (2008). Adenovirus small e1a alters global patterns of histone modification. Science *321*, 1084-1085.

Iakoucheva, L.M., Brown, C.J., Lawson, J.D., Obradovic, Z., and Dunker, A.K. (2002). Intrinsic disorder in cell-signaling and cancer-associated proteins. J. Mol. Biol. *323*, 573-584.

Joshi, P. and Vendruscolo, M. (2015). Druggability of intrinsically disordered proteins. Adv. Exp. Med. Biol. *870*, 383-400.

Kendrew, J.C., Bodo, G., Dintzis, H.M., Parrish, R.G., Wyckoff, H., and Phillips, D.C. (1958). A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. Nature *181*, 662-666.

Koh, J. and Blobel, G. (2015). Allosteric regulation in gating the central channel of the nuclear pore complex. Cell *161*, 1361-1373.

Krishnan, N., Koveal, D., Miller, D.H., Xue, B., Akshinthala, S.D., Kragelj, J., Jensen, M.R., Gauss, C.M., Page, R., Blackledge, M., et al. (2014). Targeting the disordered C terminus of PTP1B with an allosteric inhibitor. Nat. Chem. Biol. *10*, 558-566.

Kriwacki, R.W., Hengst, L., Tennant, L., Reed, S.I., and Wright, P.E. (1996). Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. Proc. Natl. Acad. Sci. U. S. A. *93*, 11504-11509.

Lacy, E.R., Filippov, I., Lewis, W.S., Otieno, S., Xiao, L., Weiss, S., Hengst, L., and Kriwacki, R.W. (2004). p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. Nat. Struct. Mol. Biol. *11*, 358-364.

Li, J., White, J.T., Saavedra, H., Wrabl, J.O., Motlagh, H.N., Liu, K., Sowers, J., Schroer, T.A., Thompson, E.B., and Hilser, V.J. (2017). Genetically tunable frustration controls allostery in an intrinsically disordered transcription factor. Elife *6*, e30688.

Liu, J., Perumal, N.B., Oldfield, C.J., Su, E.W., Uversky, V.N., and Dunker, A.K. (2006). Intrinsic disorder in transcription factors. Biochemistry *45*, 6873-6888.

Malik, S. and Roeder, R.G. (2010). The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. Nat. Rev. Genet. *11*, 761-772.

Metallo, S.J. (2010). Intrinsically disordered proteins are potential drug targets. Curr. Opin. Chem. Biol. *14*, 481-488.

Mittag, T., Marsh, J., Grishaev, A., Orlicky, S., Lin, H., Sicheri, F., Tyers, M., and Forman-Kay, J.D. (2010). Structure/function implications in a dynamic complex of the intrinsically disordered Sic1 with the Cdc4 subunit of an SCF ubiquitin ligase. Structure *18*, 494-506.

Mohan, A., Oldfield, C.J., Radivojac, P., Vacic, V., Cortese, M.S., Dunker, A.K., and Uversky, V.N. (2006). Analysis of molecular recognition features (MoRFs). J. Mol. Biol. *362*, 1043-1059.

Monod, J., Wyman, J., and Changeux, J.P. (1965). On the nature of allosteric transitions: a plausible model. J. Mol. Biol. *12*, 88-118.

Motlagh, H.N., Wrabl, J.O., Li, J., and Hilser, V.J. (2014). The ensemble nature of allostery. Nature *508*, 331-339.

Mukhopadhyay, S., Krishnan, R., Lemke, E.A., Lindquist, S., and Deniz, A.A. (2007). A natively unfolded yeast prion monomer adopts an ensemble of collapsed and rapidly fluctuating structures. Proc. Natl. Acad. Sci. U. S. A. *104*, 2649-2654.

Neira, J.L., Bintz, J., Arruebo, M., Rizzuti, B., Bonacci, T., Vega, S., Lanas, A., Velazquez-Campoy, A., Iovanna, J.L., and Abian, O. (2017). Identification of a drug targeting an intrinsically disordered protein involved in pancreatic adenocarcinoma. Sci. Rep. *7*, 39732.

Noutsou, M., Duarte, A.M., Anvarian, Z., Didenko, T., Minde, D.P., Kuper, I., de Ridder, I., Oikonomou, C., Friedler, A., Boelens, R., et al. (2011). Critical scaffolding regions of the tumor suppressor Axin1 are natively unfolded. J. Mol. Biol. *405*, 773-786.

Oates, M.E., Romero, P., Ishida, T., Ghalwash, M., Mizianty, M.J., Xue, B., Dosztanyi, Z., Uversky, V.N., Obradovic, Z., Kurgan, L., et al. (2013). D(2) P(2): database of disordered protein predictions. Nucleic Acids Res. *41*, D508-D516.

Oldfield, C.J., Cheng, Y., Cortese, M.S., Romero, P., Uversky, V.N., and Dunker, A.K. (2005). Coupled folding and binding with alpha-helix-forming molecular recognition elements. Biochemistry *44*, 12454-12470.

Onuchic, J.N., Wolynes, P.G., Luthey-Schulten, Z., and Socci, N.D. (1995). Toward an outline of the topography of a realistic protein-folding funnel. Proc. Natl. Acad. Sci. U. S. A. *92*, 3626-3630.

Papoian, G.A. (2008). Proteins with weakly funneled energy landscapes challenge the classical structure-function paradigm. Proc. Natl. Acad. Sci. U. S. A. *105*, 14237-14238.

Pelka, P., Ablack, J.N., Fonseca, G.J., Yousef, A.F., and Mymryk, J.S. (2008). Intrinsic structural disorder in adenovirus E1A: a viral molecular hub linking multiple diverse processes. J. Virol. *82*, 7252-7263.

Perutz, M.F., Rossmann, M.G., Cullis, A.F., Muirhead, H., Will, G., and North, A.C. (1960). Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-A. resolution, obtained by X-ray analysis. Nature *185*, 416-422.

Praefcke, G.J., Ford, M.G., Schmid, E.M., Olesen, L.E., Gallop, J.L., Peak-Chew, S.Y., Vallis, Y., Babu, M.M., Mills, I.G., and McMahon, H.T. (2004). Evolving nature of the AP2 alpha-appendage hub during clathrin-coated vesicle endocytosis. EMBO J. *23*, 4371-4383.

Radhakrishnan, I., Perez-Alvarado, G.C., Parker, D., Dyson, H.J., Montminy, M.R., and Wright, P.E. (1997). Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. Cell *91*, 741-752.

Radivojac, P., Iakoucheva, L.M., Oldfield, C.J., Obradovic, Z., Uversky, V.N., and Dunker, A.K. (2007). Intrinsic disorder and functional proteomics. Biophys. J. *92*, 1439-1456.

Romero, P., Obradovic, Z., Kissinger, C.R., Villafranca, J.E., Garner, E., Guilliot, S., and Dunker, A.K. (1998). Thousands of proteins likely to have long disordered regions. Pac. Symp. Biocomput. 437-448.

Spolar, R.S. and Record, M.T., Jr. (1994). Coupling of local folding to site-specific binding of proteins to DNA. Science *263*, 777-784.

Sugase, K., Dyson, H.J., and Wright, P.E. (2007). Mechanism of coupled folding and binding of an intrinsically disordered protein. Nature *447*, 1021-1025.

Tompa, P. (2002). Intrinsically unstructured proteins. Trends Biochem. Sci. *27*, 527-533.

Tompa, P., Davey, N.E., Gibson, T.J., and Babu, M.M. (2014). A million peptide motifs for the molecular biologist. Mol. Cell *55*, 161-169.

Tompa, P. and Fuxreiter, M. (2008). Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. Trends Biochem. Sci. *33*, 2-8.

Tsytlonok, M., Sanabria, H., Wang, Y., Felekyan, S., Hemmen, K., Phillips, A.H., Yun, M.K., Waddell, M.B., Park, C.G., Vaithiyalingam, S., et al. (2019). Dynamic anticipation by Cdk2/Cyclin A-bound p27 mediates signal integration in cell cycle regulation. Nat. Commun. *10*, 1676.

Tuttle, L.M., Pacheco, D., Warfield, L., Luo, J., Ranish, J., Hahn, S., and Klevit, R.E. (2018). Gcn4-mediator specificity is mediated by a large and dynamic fuzzy protein-protein complex. Cell Rep. *22*, 3251-3264.

Uversky, V.N. (2002). Natively unfolded proteins: a point where biology waits for physics. Protein Sci. *11*, 739-756.

Uversky, V.N. (2018). Intrinsic disorder, protein-protein interactions, and disease. Adv. Protein Chem. Struct. Biol. *110*, 85-121.

Uversky, V.N., Gillespie, J.R., and Fink, A.L. (2000). Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins *41*, 415-427.

Uversky, V.N., Oldfield, C.J., and Dunker, A.K. (2008). Intrinsically disordered proteins in human diseases: introducing the D^2 concept. Annu. Rev. Biophys. *37*, 215-246.

van der Lee, R., Buljan, M., Lang, B., Weatheritt, R.J., Daughdrill, G.W., Dunker, A.K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D.T., et al. (2014). Classification of intrinsically disordered regions and proteins. Chem. Rev. *114*, 6589-6631.

Vavouri, T., Semple, J.I., Garcia-Verdugo, R., and Lehner, B. (2009). Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity. Cell *138*, 198-208.

Ward, J.J., Sodhi, J.S., McGuffin, L.J., Buxton, B.F., and Jones, D.T. (2004). Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. J. Mol. Biol. *337*, 635-645.

Warfield, L., Tuttle, L.M., Pacheco, D., Klevit, R.E., and Hahn, S. (2014). A sequence-specific transcription activator motif and powerful synthetic variants that bind Mediator using a fuzzy protein interface. Proc. Natl. Acad. Sci. U. S. A. *111*, E3506-E3513.

Weathers, E.A., Paulaitis, M.E., Woolf, T.B., and Hoh, J.H. (2004). Reduced amino acid alphabet is sufficient to accurately recognize intrinsically disordered protein. FEBS Lett. *576*, 348-352.

Wei, G., Xi, W., Nussinov, R., and Ma, B. (2016). Protein ensembles: how does nature harness thermodynamic fluctuations for life? The diverse functional roles of conformational ensembles in the cell. Chem. Rev. *116*, 6516-6551.

Wootton, J.C. (1994). Non-globular domains in protein sequences: automated segmentation using complexity measures. Comput. Chem. *18*, 269-285.

Wright, P.E. and Dyson, H.J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J. Mol. Biol. *293*, 321-331.

Wright, P.E. and Dyson, H.J. (2009). Linking folding and binding. Curr. Opin. Struct. Biol. *19*, 31-38.

Wright, P.E. and Dyson, H.J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. Nat. Rev. Mol. Cell Biol. *16*, 18-29.

Wu, H. and Fuxreiter, M. (2016). The structure and dynamics of higher-order assemblies: amyloids, signalosomes, and granules. Cell *165*, 1055-1066.

Xie, H., Vucetic, S., Iakoucheva, L.M., Oldfield, C.J., Dunker, A.K., Uversky, V.N., and Obradovic, Z. (2007). Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. J. Proteome Res. *6*, 1882-1898.

Xue, B., Romero, P.R., Noutsou, M., Maurice, M.M., Rudiger, S.G., William,

A.M., Jr., Mizianty, M.J., Kurgan, L., Uversky, V.N., and Dunker, A.K. (2013). Stochastic machines as a colocalization mechanism for scaffold protein function. FEBS Lett. *587*, 1587-1591.

Zhang, Y., Qiu, W.J., Liu, D.X., Neo, S.Y., He, X., and Lin, S.C. (2001). Differential molecular assemblies underlie the dual function of Axin in modulating the WNT and JNK pathways. J. Biol. Chem. *276*, 32152-32159.

Zhou, H.X. (2012). Intrinsic disorder: signaling via highly specific but short-lived association. Trends Biochem. Sci. *37*, 43-48.