PORTLAND PRESS

The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease

M. Madan Babu*

MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, U.K.

Correspondence: M. Madan Babu (madanm@mrc-lmb.cam.ac.uk)

OPEN ACCESS

In the 1960s, Christian Anfinsen postulated that the unique three-dimensional structure of a protein is determined by its amino acid sequence. This work laid the foundation for the sequence–structure–function paradigm, which states that the sequence of a protein determines its structure, and structure determines function. However, a class of polypeptide segments called intrinsically disordered regions does not conform to this postulate. In this review, I will first describe established and emerging ideas about how disordered regions contribute to protein function. I will then discuss molecular principles by which regulatory mechanisms, such as alternative splicing and asymmetric localization of transcripts that encode disordered regions, can increase the functional versatility of proteins. Finally, I will discuss how disordered regions contribute to human disease and the emergence of cellular complexity during organismal evolution.

Introduction

Understanding how proteins, which are polymers of amino acids, carry out different functions in a cell has been a problem of considerable interest. In a series of elegant publications, Christian Anfinsen and colleagues proposed that the sequence of a protein contains the information required to adopt a defined structure and, hence, function. This led to what is now called as Anfinsen's postulate or the thermodynamic hypothesis, which states that 'the three-dimensional structure of the native protein in its normal physiological milieu is the one in which the Gibbs-free energy of the whole system is the lowest; that is, that the native conformation is determined by the totality of the interatomic interactions and hence by the amino acid sequence, in a given environment' [1]. The biochemical studies of Anfinsen and colleagues, along with the unprecedented molecular insights obtained from crystallographic studies of proteins, such as hemoglobin and numerous enzymes, established the structure–function paradigm (Figure 1A). This concept postulates that function is achieved by the unique three-dimensional structure adopted by a protein, which in turn is determined by its amino acid sequence (see ref. [2] for a general historical overview).

While most proteins and polypeptide segments fold co-operatively into defined three-dimensional structures, numerous studies, primarily over the last couple of decades, have discovered that a large number of polypeptide segments do not fold into defined tertiary structure. Instead, they adopt an ensemble of different conformations and can still carry out their function in an unstructured/disordered state [3–6]. These studies are now establishing the disorder–function paradigm (Figure 1B), which states that certain polypeptide segments can be functional without achieving a defined tertiary structure [7–15]. Recent studies that have investigated genome sequences of many organisms have established that over 40% of any eukaryotic proteome contains such disordered regions [16–18]. More importantly, altered abundance and mutations in many proteins with disordered segments have been implicated in human diseases, such as neurodegeneration and cancer [19–28].
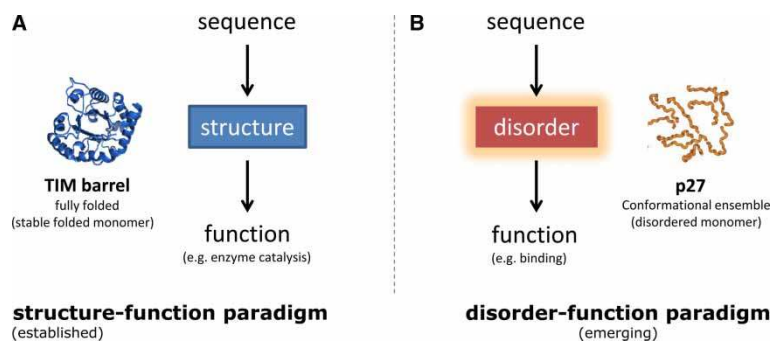
Figure 1. Sequence to function relationship.
(**A**) Structure–function paradigm and (**B**) disorder–function paradigm. Reprinted with permission from ref. [7].

Conformational states of intrinsically disordered regions

A major determinant of polypeptide segments folding co-operatively into a defined tertiary structure is the long-range hydrophobic interaction between amino acids in the linear sequence [29,30]. Intrinsically disordered regions (IDRs) are polypeptide segments that do not contain sufficient hydrophobic amino acids to mediate co-operative folding. Instead, they typically contain a higher proportion of polar or charged amino acids [31]. Thus, IDRs lack a unique three-dimensional structure either entirely or in parts in their native state. They generally sample a variety of conformations that are in dynamic equilibrium under physiological conditions [14,32–34].

This, however, does not mean that they are completely flexible and adopt all possible conformations. Computational analysis of sequences, single-molecule studies, and molecular dynamics simulations has revealed that the amino acid composition affects the IDR conformational states and can determine whether they adopt a totally extended conformation (segments with high net charge and low hydrophobicity) or a compact conformation (depending on the balance between hydrophobicity and net charge) [35–38]. This can further influence the functional elements (e.g. motifs or posttranslational modification sites) that are embedded within IDRs and can affect critical processes such as the cell cycle [39]. For the same number of charged residues, the charge patterning has also been shown to determine whether the polypeptide segment will be fully extended (e.g. alternating positively and negatively charged residues) or a collapsed globule (e.g. clearly separated stretches of positively and negatively charged residues), or somewhere in between (Figure 2) [38,40].

Advantages and functions mediated by IDRs

IDRs can provide many advantages to proteins (Figure 3). These include the following: (a) exposing short linear motifs that can mediate domain peptide interactions [41–44]. This permits interaction of the same protein with a large number of interaction partners in a functionally promiscuous manner or assembly of multiple proteins by serving as a scaffold (e.g. as seen in the AP2 adaptor protein during endocytosis [45,46]). (b) Facilitating the regulation of protein function via diverse posttranslational modification (PTM) of residues within the IDR [42,47,48]. Owing to their conformational flexibility, IDRs serve as excellent substrates to encode and decode information via posttranslational modifications (e.g. as seen in the tails of histone proteins or in the cytoplasmic tails of receptor tyrosine kinases and GPCRs [49–52]). (c) Regulating protein half-life by efficiently engaging proteins that have been targeted for degradation by the proteasome [53–59]. (d) Adopting different conformations when binding to different interaction partners [12,60–66]. These properties of IDRs make them well suited to perform signaling and regulatory functions. Indeed, genome-scale analyses of the functions of proteins with IDRs have revealed that they are enriched in signaling proteins and nucleic acid-binding proteins such as kinases, transcription factors and splicing factors [67–70].

Folding upon binding of IDRs

An important aspect by which IDRs contribute to protein function is by adopting a defined conformation when binding a specific interaction partner [6,9,32,34,53,71–74]. Although a large fraction of the polypeptide
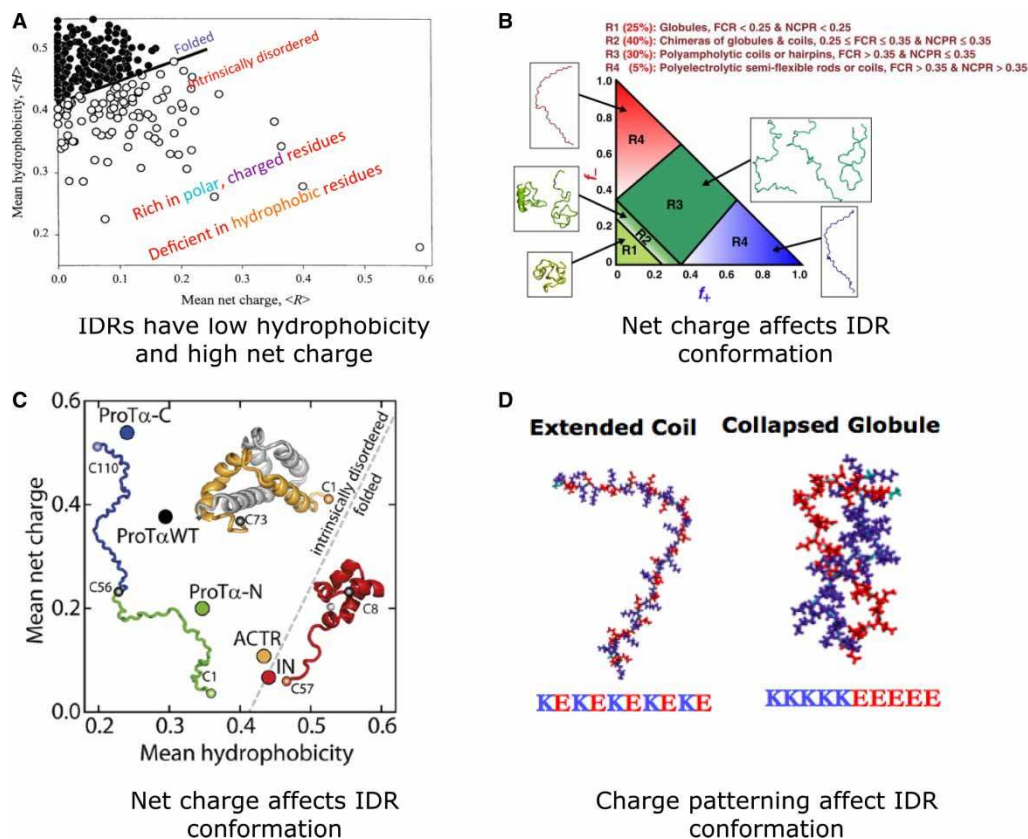
Figure 2. The relationship between sequence composition and conformations adopted by IDRs.

(**A**) Plot of mean net charge versus mean hydrophobicity reveals the clear separation between structured proteins and IDPs. Reprinted with permission from ref. [31]. (**B**) Phase diagram showing the conformations of IDRs for different fractions of positive ($f+$) and negative charges ($f–$). Reprinted with permission from ref. [38]. FCR, fraction of charged residues; NCPR, Net charge per residue. (**C**) IDRs with sufficient hydrophobicity tend to fold upon binding (yellow, ACTR). Reprinted with permission from ref. [157]. ACTR, activator for thyroid hormones and retinoid receptors; ProTα-C, prothymosin α C-terminal segment; ProTα-WT, prothymosin α wild type; ProTα-N: prothymosin α N-terminal segment; IN, HIV integrase. (**D**) For the same net charge, the patterning can determine if the IDR adopts an extended coil or a collapsed globule conformation. Reprinted with permission from ref. [40].

adopts a defined structure upon complex formation, distinct segments can still remain disordered. This phenomenon has been referred to as fuzzy complex formation [75–78]. The folding and binding of IDRs facilitates interaction with their targets with relatively high specificity and low affinity [79,80]. This can permit highly specific associations to trigger signaling events while facilitating rapid disassociation when signaling is completed (e.g. p27 interaction with cyclin–CDK during the cell cycle; Figure 4A) [80]. The low-free energy of binding is due to the fine balance associated with the high entropic cost of folding and a comparable enthalpic gain of binding [60,72]. Thus, small perturbations either to entropy or to enthalpy of binding, such as via posttranslational modifications, can trigger association or disassociation from their interaction partners (e.g. CBP–CREB interaction; Figure 4B) [72]. While 'weak but specific' binding is often observed for IDRs, they also display very tight binding in several cases, which is often overlooked [81]. In terms of the kinetics of interactions, such proteins can have a wide spectrum of association and disassociation rates depending on the mode of interaction (e.g. conformational selection versus induced folding) [65,80,82–84]. For a given K_d value, the kinetic constants can vary widely [81]. Several different intrinsically disordered proteins (IDPs; proteins with IDRs) have exploited this property in order to facilitate robust cellular decision-making (e.g. as seen in the PUMA–MCL1 interaction involved in apoptosis; Figure 4C) [85].
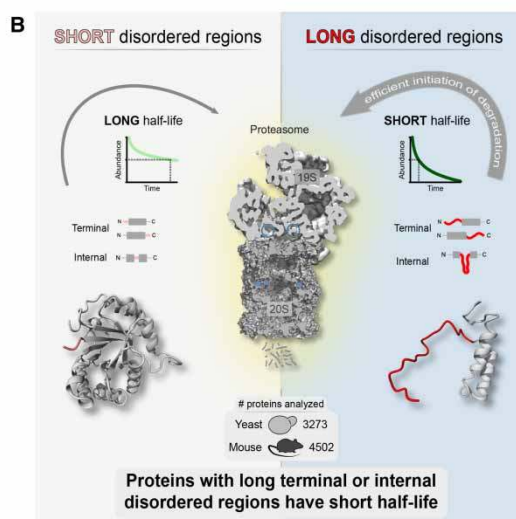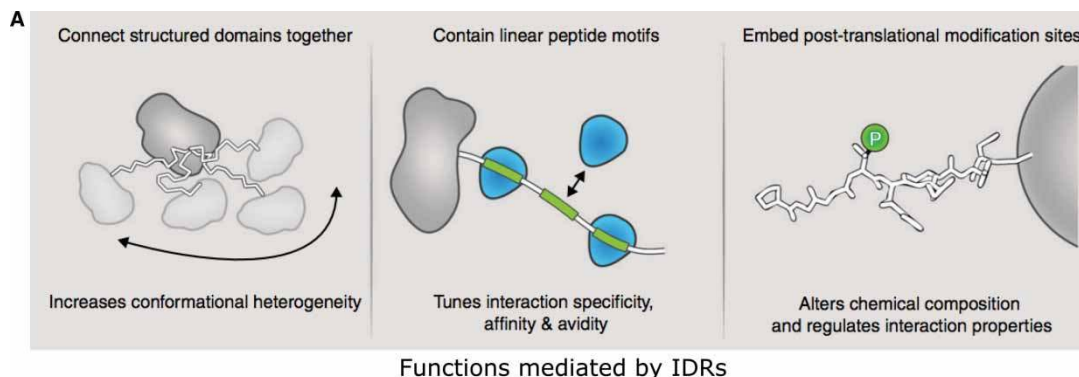
Figure 3. Advantages and functions mediated by IDRs.

(**A**) IDRs can link structured domains, where their flexibility permits the protein to adopt multiple conformations; linear motifs within IDRs mediate protein interactions; posttranslational modification of residues within IDRs permits encoding and decoding of information [106]. (**B**) IDRs in protein sequences can increase the efficiency of degradation by the proteasome, thereby regulating protein half-life [53].
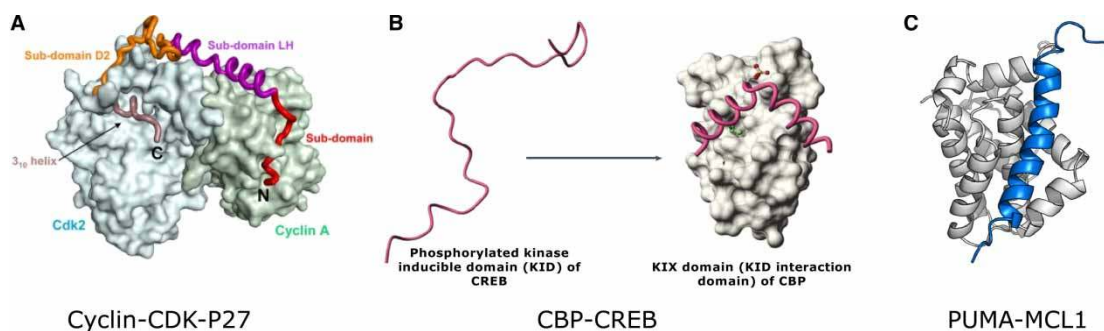

Figure 4. Coupled folding and binding of IDRs.

(**A**) p27–cyclin–CDK complex. Reprinted with permission from ref. [158]. (**B**) CBP–CREB interaction regulated by phosphorylation. Reprinted with permission from ref. [32]. (**C**) PUMA–MCL1 interaction in apoptosis. Reprinted with permission from ref. [159].

Formation of higher-order assemblies by IDRs

Recently, it has been demonstrated that many low-complexity regions and IDRs with repeating peptide motifs can form nonmembrane-bound organelles and higher-order assemblies, often in a highly reversible manner [86–98]. For instance, Q/N-rich regions are important for forming cellular assemblies, such as P-bodies, FG-rich regions are critical in forming the hydrogel-like structure of the nuclear pore, and repeats of multiple linear motifs can mediate phase separation and organize matter in cells, as seen in certain actin regulatory proteins (Figure 5) [92,99–102]. Thus, IDRs can mediate functions comparable to structured domains, such as (i) the formation of protein complexes and higher-order assemblies of variable stoichiometry of subunits [86], (ii) conformational transition (disorder-to-order and order-to-disorder) in response to specific environmental changes, context, or ligands [94], and (iii) allosteric communication [15,60,103–105]. Since most proteins contain structured and disordered regions in varying proportions, together with structured domains in the same polypeptide chain, IDRs can synergistically increase the functional versatility of proteins [12,15].

In this award lecture review, I will first describe emerging ideas on how alternative splicing of disordered regions can rewire protein interaction networks in a tissue-specific (TS) manner, thereby leading to increased complexity and diversity of interactomes of different tissues [106,107]. I will then describe our studies on how asymmetric mRNA localization and local translation of transcripts encoding IDRs can facilitate organization of higher-order assemblies in distinct parts of the cell and contribute to increased fidelity of signaling networks [108]. Finally, I will discuss how altered regulation and mutations within IDRs can cause many diseases [24,27,109].

Splicing of disordered regions and functional versatility

Alternative splicing is a molecular mechanism that results in the formation of multiple transcripts from the same gene. In this manner, alternative splicing increases the potential number of distinct protein products that can be encoded by a single gene [110–119]. Many next-generation sequencing studies have established that over 90% of human genes are expected to undergo alternative splicing [111–113,120,121]. Interestingly, these studies also estimate that nearly 50% of the isoforms are likely to be expressed in a TS manner [113,120]. While high-throughput studies have established the extent of splicing at the transcript level, the roles played by the different variants at the protein level are not fully understood. In one of our studies, we investigated the characteristics of tissue-specific spliced exons and how they could have an impact on the function of the encoded protein. To this end, we systematically collected the complete transcriptome sequence of 10 human tissues and 5 human cell lines [120] and classified the exons into three groups: those that are (a) constitutively expressed, (b) alternatively included or excluded but expressed in multiple tissues, and (c) alternatively included/excluded but in a tissue-specific manner. We then investigated the structural properties of the encoded protein segment of these exons, analyzed their functional features (such as linear motifs and PTM
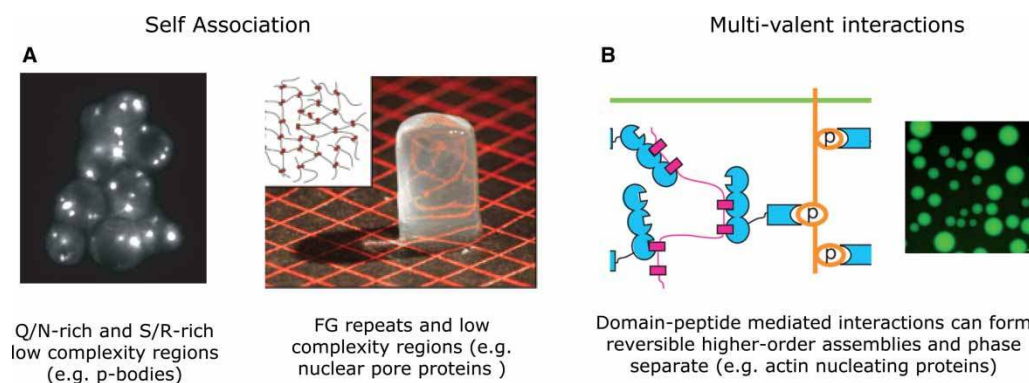

Figure 5. Formation of nonmembrane-bound organelles and higher-order assemblies by IDRs.
(**A**) Self-association. Q/N-rich regions are important for P-bodies. Reprinted with permission from ref. [160]. FG-rich regions form hydrogel-like structures at the nuclear pore. Reprinted with permission from ref. [100]. (**B**) Multivalent interactions. Contacts between proteins containing repeating domains and peptide motifs can mediate phase transition that can be regulated via posttranslational modification. Reprinted with permission from ref. [92].

sites) and evolutionary conservation, and studied various properties in terms of the protein interaction networks that they participate in within different tissues [107].

This analysis allowed us to make many observations. First, <5% of the TS exons map to complete protein domains, suggesting that protein segments encoded by TS exons are unlikely to adopt a defined tertiary structure. Further investigation revealed that such segments are enriched in disordered regions, which contain linear motifs and PTM sites that are evolutionarily conserved between human and mouse orthologs [107]. A detailed analysis of the proteins containing such segments revealed that they tend to have more interaction partners and mediate a higher number of TS interactions in the respective tissues where they are known to be alternatively spliced. Collectively, these observations suggested that disordered TS segments are unlikely to be passive linkers that connect structured domains, but have the potential to mediate new interactions via peptide motifs and PTM sites [107].

To understand the molecular details of how TS exons can affect protein interactions, we systematically mapped the TS exons onto the known three-dimensional structures of proteins and protein complexes that were available in the Protein Data Bank [107]. This allowed us to describe the following general principles. TS splicing of disordered regions that contain (a) binding motifs can rewire protein interactions and thus contribute to the specificity of an interaction (Figure 6A) and (b) posttranslational modification sites can rewire signaling networks and make the protein a substrate for specific signaling proteins in a TS manner (Figure 6B). In this way, TS splicing of disordered segments can contribute to the rewiring of protein interactions and signaling networks in a TS manner and increase the diversity of protein networks in different tissues (Figure 7A) [107].

In addition to affecting protein interactions, such splicing events can also affect protein complexes. More specifically, expression of two different isoforms that can still interact (e.g. through the protein segment encoded by a constitutive exon) can lead to hybrid complexes. Depending on the region that is spliced, different isoforms can sequester and compete for the same interaction partner, which can lead to dominant-negative response, ultra-sensitive response, or transient gain- or loss-of-function effects — depending on which proteins are sequestered into nonfunctional complexes (Figure 7B) [107]. This is highlighted by the expression of an isoform of p53, which contains the DNA-binding domain but not the transactivation domain, a disordered region that is required to recruit the transcriptional machinery [107,122]. Expression of this isoform competes for the same regulatory elements in the genome in the promoter region of the p53 target genes, but ends up repressing gene expression of the targets since the transactivation domain is missing in this isoform. This mechanism has been exploited during development, by influencing pluripotency and differentiation of embryonic stem cells [122].

In this manner, TS splicing of disordered segments leads to the recruitment of the same biochemical activity (often carried out by structured domains encoded by constitutive exons) to different molecular contexts by mediating new protein interactions through the differentially spliced unstructured segment [107]. For example, even though the substrate protein may be expressed in a certain tissue, the TS inclusion or exclusion of a disordered substrate-docking motif in kinases can determine whether the kinase domain can phosphorylate a substrate protein or not. Similarly, even though the kinase may be expressed in the cell type, the TS inclusion/exclusion of a disordered modifiable residue within a substrate can determine whether it can be regulated by that particular kinase or not. In the case of transcription factors, splicing of the disordered transactivation domain in a TS manner or during specific times in development can convert a transcription factor from an activator to a repressor and hence, the same set of target genes can show very different transcriptional responses in different tissue types [107].

A number of related studies have all independently described similar observations [106,111,119,123–130]. Taken together, these studies reveal that alternative splicing of disordered segments can have important consequences [without affecting structured domain(s)] by rewiring signaling and regulatory networks in different cell types or during development. In this way, they increase the functional versatility of proteins by providing new contexts and expand the diversity of interaction networks in the different tissue types or at different time points during development. The plasticity associated with the divergence of alternative splicing between different organisms may have led to the emergence of novel phenotypes and increased complexity during organismal evolution [121,131,132].

Localized translation of IDPs and cellular complexity

Asymmetric localization of proteins is a key to a wide variety of functions ranging from signal transduction in neurons and asymmetric cell division during development to maintaining cellular morphology [133–135]. The
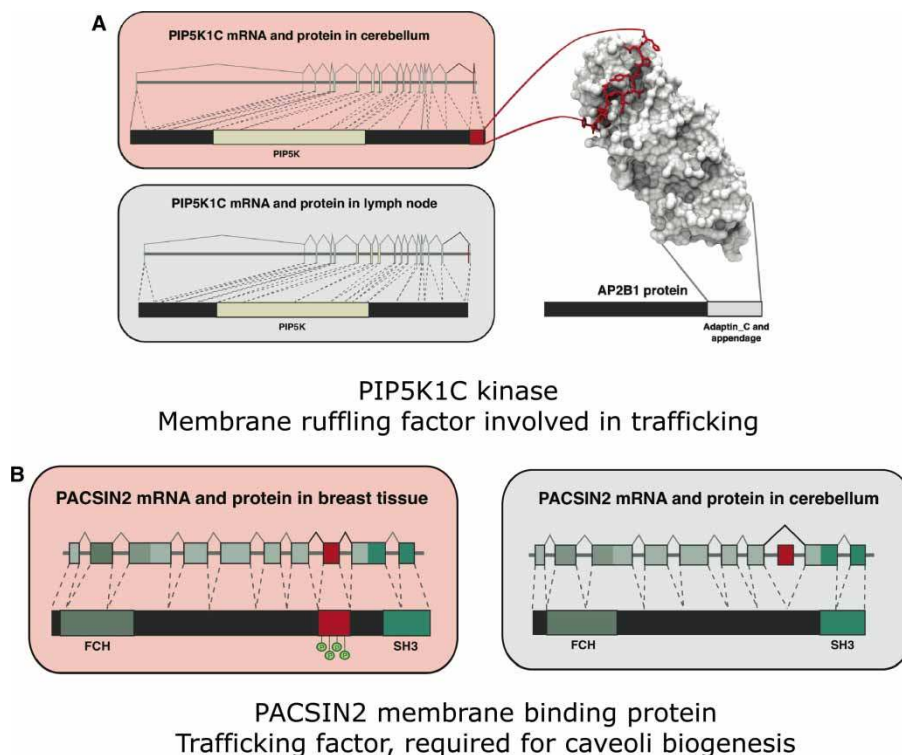
Figure 6. TS splicing can rewire protein interaction and signaling networks.

TS splicing of IDRs can (**A**) affect interactions with other proteins by differential inclusion of linear peptide motifs and (**B**) influence whether a signaling enzyme can regulate a protein by differential inclusion of IDRs that contain posttranslational modification sites. Reprinted with permission from ref. [107].
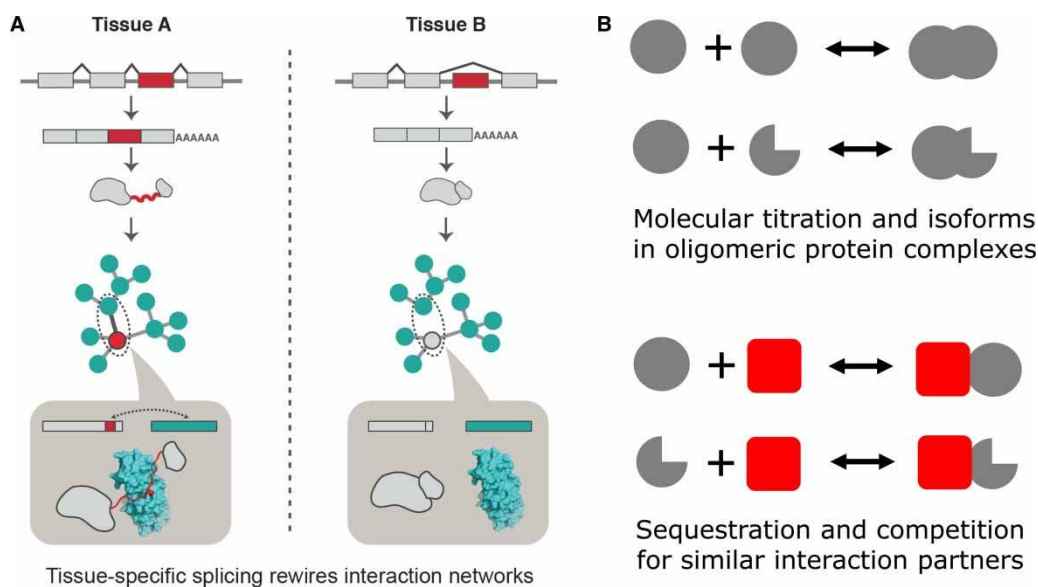

Figure 7. Impact of TS splicing of IDRs on protein networks and complexes.

(**A**) Rewiring of protein interaction networks and signaling pathways by TS splicing. Reprinted with permission from ref. [107].
(**B**) Expression of multiple isoforms can affect response kinetics and influence cellular decision-making (ultra-sensitive behavior, dominant-negative response, and sequestration of interacting partners, leading to gain or loss of function).

importance of asymmetric localization is apparent when localization goes awry, thereby leading to developmental defects and disease. For instance, in *Drosophila* embryos, mislocalization of *Oskar* mRNA to the anterior side produces embryos that have two abdomens with mirror-image symmetry [136]. Many studies over the last years have identified that there are two major mechanisms by which asymmetric localization of proteins can be achieved: The first mechanism involves protein transport after synthesis (translation), whereas the second one involves transport of the mRNA to specific locations in the cell followed by localized translation (Figure 8A) [134,137–142]. Both mechanisms can generate asymmetric localization of proteins and often operate in the same cell type [134].

Are there differences between proteins that are transported after synthesis (TAS) compared with those that undergo on-site synthesis (OSS) after mRNA transport? To investigate this question, we first compiled multiple large-scale datasets that experimentally identified asymmetrically localized transcripts and proteins in fibroblasts, mouse neuronal cells, and rat dorsal root ganglion cells from embryo and adult [140,141,143–146]. We then systematically integrated multiple large-scale datasets that described the structural properties of the proteins, mRNA, and protein expression levels, half-life, and functional features of proteins and transcripts to uncover the underlying trends [108].

This integrated large-scale analysis allowed us to make many observations, based on which we described general principles of how localized translation of transcripts that encode disordered proteins can (a) enhance signaling fidelity and sensitivity and (b) increase cellular complexity through precise spatial localization of proteins and formation of nonmembrane-bound assemblies (Figure 8B). A systematic comparison of structural properties revealed that TAS proteins are enriched in structured domains, whereas OSS proteins are enriched in disordered regions. Furthermore, proteins that are synthesized on-site are enriched in repeating linear motifs that have the potential to form higher-order assemblies. Such proteins also tend to be posttranslationally modified either within the motif or just around the motif, suggesting that OSS proteins might direct the flow of information and regulate the formation of reversible assemblies by using posttranslational modifications to switch protein interactions on/off. In addition, there was enrichment for OSS proteins encoding low-complexity regions; specifically, Q/N-rich regions and FG repeat-rich regions, both of which can undergo phase separation and form reversible, nonmembrane-bound assemblies. Taken together, these observations suggested that in contrast with transport after synthesis proteins, OSS proteins encode disordered regions, which contain multivalent, assembly-promoting segments that are surrounded by posttranslational modification sites (interaction/PTM switches; [41,47,147]). The trends were consistent across different cell types, organisms, and developmental stages, suggesting that these observations are likely to be applicable to different organisms [108].

Given the potentially promiscuous nature of such proteins, we then investigated how their availability is regulated. A systematic analysis of the protein abundance, protein half-life, transcript abundance, and transcript half-life revealed that OSS proteins and their transcripts are tightly regulated at almost every stage along the process of gene expression compared with the TAS group of proteins. An investigation of how the abundance and the PTM status of the two groups of asymmetrically localized proteins change over time after stimulating cells revealed that OSS proteins tend to increase their abundance more rapidly after stimulation compared with the TAS proteins. In summary, these findings suggested that proteins that are synthesized on-site are generally present in low abundance and are tightly temporally regulated. However, upon receiving a signal (e.g. stimulation with growth factors), they display a rapid increase in abundance and distinct phosphorylation dynamics [108].

There are many implications of the observations described here. Since many of the proteins that are synthesized on-site are likely to mediate promiscuous interactions and form higher-order assemblies, spatial localization of their transcripts to where they are required and their synthesis on demand by local translation can significantly restrict the likelihood of off-target interactions. Furthermore, since asymmetric mRNA localization decentralizes gene expression by decoupling transcription and translation, such a mechanism ensures that cells can rapidly respond to signals at the site where the signal is received and can process information within specific sub-cellular locations. In this way, localized translation after mRNA transport can sharpen the sensitivity of signaling networks and lead to nonlinear input–output responses for efficient information processing. It also ensures that while the overall copy number of the regulatory and signaling proteins may be low in a cell, at specific locations their local concentrations can be sufficiently high to help mediate their function (Figure 8B) [108].

On-site synthesis of proteins could further act as a general mechanism to ensure that nucleating proteins are available at the right place, in appropriate amounts, and only when required. In this manner, spatial control by localized translation may play a central role in signaling, by enhancing interaction fidelity and sensitivity, and by minimizing noisy, off-target interactions. Thus, together with other modes of regulation and temporal cues,
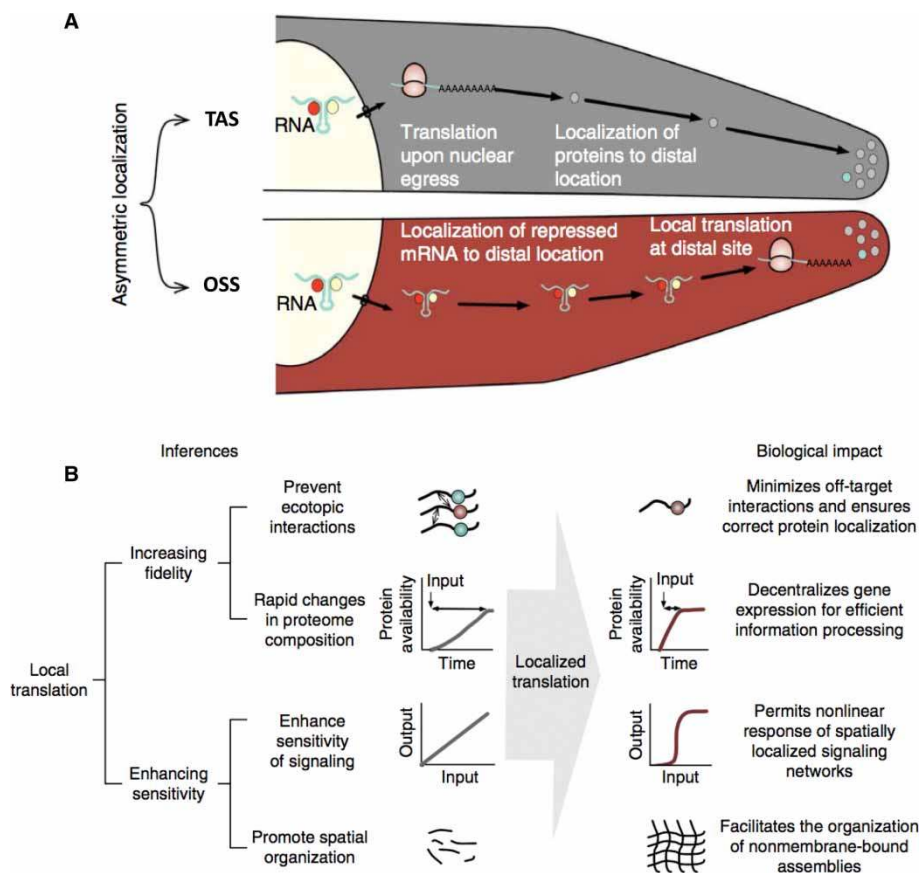
PORTLAND
PRESS

Figure 8. Asymmetric localization of proteins with IDRs.
(**A**) Mechanisms to achieve asymmetric protein localization — transported after synthesis (TAS) or on-site synthesis (OSS) after asymmetric mRNA localization. (**B**) Advantages and implications of localized translation upon asymmetric mRNA localization. Reprinted with permission from ref. [108].

such as signal integration via posttranslational modification, spatial control of proteins by localized translation can have a significant impact on cell signaling [108].

IDPs and disease

While many studies have shown how proteins with IDRs can contribute to increased functional versatility and cellular complexity, research over the last several years has also revealed the importance of IDRs in many human diseases [19,24]. Mutations that lead to the alteration in the levels of proteins with IDRs can result in protein aggregation, leading to diseases such as neurodegeneration. Not surprisingly, it has been reported that aggregates of IDPs are found in very high concentrations in plaques and brain deposits of patients with neurodegenerative diseases (Figure 9A). Similarly, mutations within IDRs that increase the aggregation propensity, such as those seen in the amyloid β-peptide, α-synuclein, and huntingtin, have been directly linked to diseases such as Alzheimer's, Parkinson's, and Huntington's diseases, respectively [7,19,24,86,87,148–153].

It has been shown that IDRs are enriched in genes that participate in cell signaling and cancer-associated proteins, such as oncogenes or tumor suppressor genes [21]. Since IDRs typically contain motifs that can mediate low-affinity promiscuous interactions, altered abundance can form undesirable ectopic interactions and sequester other proteins into nonproductive complexes. In this manner, they can disturb the fine balance in many signaling and regulatory networks, leading to diseases such as cancer. Not surprisingly, gene fusions and missplicing of proteins with IDRs have also been associated with cancer [28,126,154]. Given our observations on the role of TS splicing and asymmetric localization of proteins with IDRs [107,108], it is likely that altered abundance of splicing factors and RNA-binding proteins that regulate the composition and the localization of mRNA may alter the
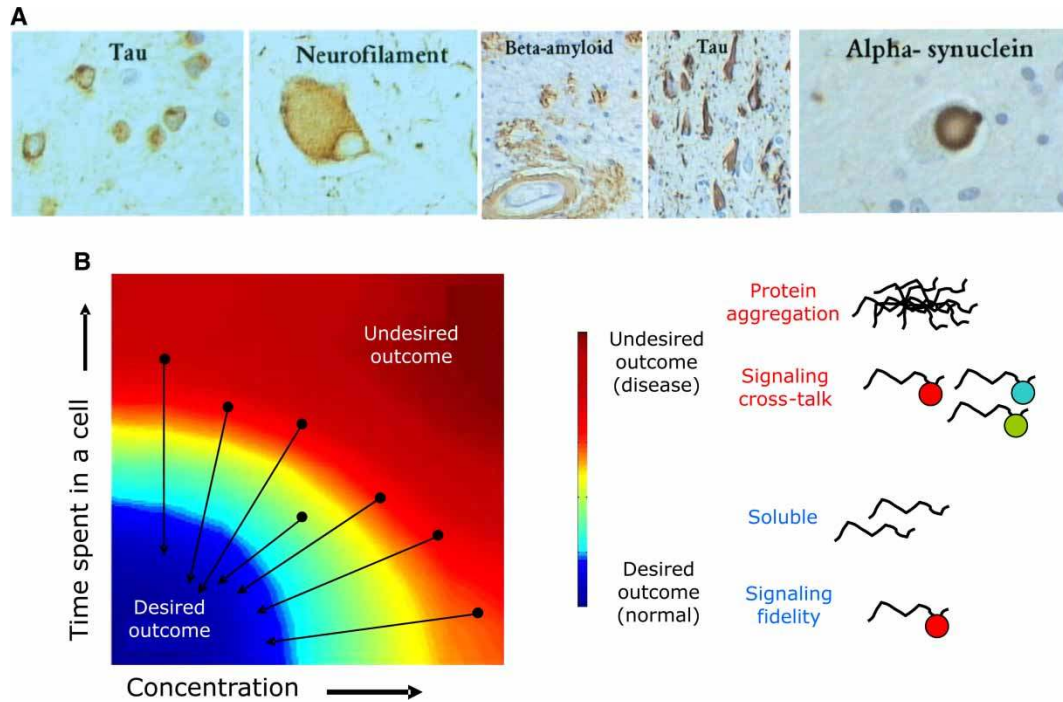
Figure 9. IDRs and disease.

(**A**) IDRs are found in plaques and cellular deposits of patients with neurodegenerative disease. Reprinted with permission from ref. [161]. (**B**) Protein availability–outcome landscape. Tight regulation of proteins with IDRs (black arrow) ensures that they are present in the right amount and not longer than required. Adapted and reprinted with permission from refs [24,27,108].

protein sequence, localization, and availability of IDPs. This may result in off-target and potentially ectopic signaling events and might explain the molecular basis for cell type-specific disease phenotypes.

To address how the beneficial and potentially detrimental roles of proteins with IDRs are balanced in the cell, we investigated the availability of such proteins in a cell, both in terms of the time spent in the cell as well
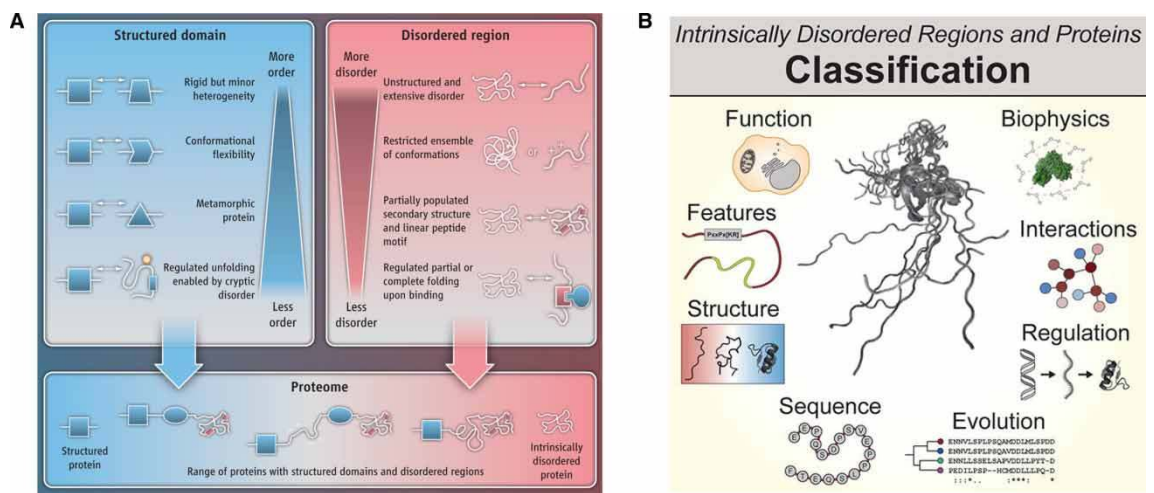

Figure 10. IDRs are fundamental units of protein function, regulation and evolution.

(**A**) Synergy between structured domains and IDRs increases the functional versatility of proteins. Reprinted with permission from ref. [12]. (**B**) Classification of IDRs and IDPs. Reprinted with permission from ref. [7].

as in the steady-state amounts of IDPs and their transcripts in many organisms, ranging from yeast to human (Figure 9B). We and others have observed that proteins with IDRs are more tightly regulated than those with structured domains at multiple stages of gene expression, ranging from transcript synthesis to protein degradation [27,109,155]. In this manner, IDPs are tightly regulated to be present in the right abundance and for the appropriate amount of time in a cell. As long as this happens, the desirable outcome, such as interaction fidelity and solubility, is achieved. However, if their half-life or protein abundance is significantly altered, that may lead to undesirable outcomes such as protein aggregation or signaling cross-talk due to nonfunctional promiscuous interactions [24,27,109]. We suggested that, within a cell, a co-ordinated tight regulation of IDPs at several stages of transcription and translation ensures that they are present for short amounts of time and in low quantities [27,109]. This strategy minimizes the harmful effects of IDPs and at the same time permits their vital contribution to the functioning of the cell. An important implication of this observation is that, in addition to mutations that affect the IDPs and cause disease, mutations affecting genes that regulate IDPs availability can be an important class of disease genes that should be closely investigated in genome-wide association studies of human diseases [24,27].

Conclusion

We have come a long way in our understanding of how proteins carry out their function in cells. In addition to the structured domains, which ensure precise positioning of side chains of specific amino acids in spatial proximity to carry out their function, IDRs, which adopt multiple conformational states, are emerging to be fundamental units of protein function and regulation. IDRs are not just passive linkers that connect different structured domains, but actively provide new contexts to structured domains and, hence, enhance the functional space associated with proteins (Figure 10A). IDRs are not all the same, but they can be classified into different groups based on various properties (Figure 10B) [7]. Since the number of functional residues in IDRs is small and clustered in the linear sequence (e.g. short linear motifs that mediate protein interactions), they can be gained and lost rapidly during evolution [41,42,156]. Thus, IDRs, in otherwise less evolvable proteins (e.g. developmentally important proteins), facilitate the exploration of new functional landscapes by changing the context in which the biochemical function can be applied. Therefore, disordered regions need to be studied in the right biological context to understand how complex functions emerge in cellular systems.

In conclusion, it is an exciting time for researchers who are investigating proteins with IDRs. Given the emerging importance of IDRs and a newfound understanding of their biomedical relevance, many discoveries regarding their myriad roles are yet to be unraveled. IDRs are now to researchers what the first few protein structures were to biologists half a century ago. We have witnessed the knowledge and impact on human health of the structure–function paradigm in the last 50 years. If structured proteins are only half the story, it brings to our attention the enormous possibilities and the potential of disordered proteins that remains to be tapped for bettering human health and revolutionizing medicine.

Abbreviations

AP2, adaptor protein complex 2; CBP, CREB-binding protein; CDK, cyclin dependent kinase; CREB, cAMP response element binding protein; FG repeat, phenylalanine-glycine repeat; GPCRs, G-protein coupled receptors; IDPs, intrinsically disordered proteins; IDRs, intrinsically disordered regions; MCL1, myeloid cell leukaemia 1; OSS, on-site synthesis; P-bodies, processing bodies; PTM, post-translational modification; PUMA, p53 up-regulated modulator of apoptosis; TAS, transported after synthesis; TS, tissue-specific.

Competing Interests

The Author declares that there are no competing interests associated with this manuscript.

References

1. Anfinsen, C.B. (1973) Principles that govern the folding of protein chains. *Science* **181**, 223–230 doi:10.1126/science.181.4096.223
2. Fersht, A.R. (2008) From the first protein structures to our current knowledge of protein folding: delights and scepticisms. *Nat. Rev. Mol. Cell Biol.* **9**, 650–654 doi:10.1038/nrm2446
3. Wright, P.E. and Dyson, H.J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293**, 321–331 doi:10.1006/jmbi.1999.3110
4. 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. USA* **93**, 11504–11509 doi:10.1073/pnas.93.21.11504
5. Romero, P., Obradovic, Z., Kissinger, C.R., Villafranca, J.E., Garner, E., Guilliot, S. et al. (1998) Thousands of proteins likely to have long disordered regions. *Pac. Symp. Biocomput.* 437–448 PMID: 9697202
6. Mittag, T., Kay, L.E. and Forman-Kay, J.D. (2010) Protein dynamics and conformational disorder in molecular recognition. *J. Mol. Recognit.* **23**, 105–116 doi:10.1002/jmr.961
7. van der Lee, R., Buljan, M., Lang, B., Weatheritt, R.J., Daughdrill, G.W., Dunker, A.K. et al. (2014) Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **114**, 6589–6631 doi:10.1021/cr400525m
8. Tompa, P. (2012) Intrinsically disordered proteins: a 10-year recap. *Trends Biochem. Sci.* **37**, 509–516 doi:10.1016/j.tibs.2012.08.004
9. Wright, P.E. and Dyson, H.J. (2015) Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **16**, 18–29 doi:10.1038/nrm3920
10. Uversky, V.N. and Dunker, A.K. (2010) Understanding protein non-folding. *Biochim. Biophys. Acta, Proteins Proteomics* **1804**, 1231–1264 doi:10.1016/j.bbapap.2010.01.017
11. Oldfield, C.J. and Dunker, A.K. (2014) Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.* **83**, 553–584 doi:10.1146/annurev-biochem-072711-164947
12. Babu, M.M., Kriwacki, R.W. and Pappu, R.V. (2012) Structural biology. Versatility from protein disorder. *Science* **337**, 1460–1461 PMID: 22997313
13. Dunker, A.K. and Kriwacki, R.W. (2011) The orderly chaos of proteins. *Sci. Am.* **304**, 68–73 doi:10.1038/scientificamerican0411-68
14. Forman-Kay, J.D. and Mittag, T. (2013) From sequence and forces to structure, function, and evolution of intrinsically disordered proteins. *Structure* **21**, 1492–1499 doi:10.1016/j.str.2013.08.001
15. Latysheva, N.S., Flock, T., Weatheritt, R.J., Chavali, S. and Babu, M.M. (2015) How do disordered regions achieve comparable functions to structured domains? *Protein Sci.* **24**, 909–922 doi:10.1002/pro.2674
16. Oates, M.E., Romero, P., Ishida, T., Ghalwash, M., Mizianty, M.J., Xue, B. et al. (2013) D^{2}P^2: database of disordered protein predictions. *Nucleic Acids Res.* **41**, D508–D516 doi:10.1093/nar/gks1226
17. 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 doi:10.1016/j.jmb.2004.02.002
18. Potenza, E., Domenico, T.D., Walsh, I. and Tosatto, S.C.E. (2015) MobiDB 2.0: an improved database of intrinsically disordered and mobile proteins. *Nucleic Acids Res.* **43**, D315–D320 doi:10.1093/nar/gku982
19. 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 doi:10.1146/annurev.biophys.37.032807.125924
20. Uversky, V.N. (2014) The triple power of D^3: protein intrinsic disorder in degenerative diseases. *Front. Biosci.* **19**, 181–258 doi:10.2741/4204
21. Iakoucheva, L.M., Brown, C.J., Lawson, J.D., Obradović, Z. and Dunker, A.K. (2002) Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* **323**, 573–584 doi:10.1016/S0022-2836(02)00969-5
22. Vacic, V., Markwick, P.R.L., Oldfield, C.J., Zhao, X., Haynes, C., Uversky, V.N. et al. (2012) Disease-associated mutations disrupt functionally important regions of intrinsic protein disorder. *PLoS Comput. Biol.* **8**, e1002709 doi:10.1371/journal.pcbi.1002709
23. Vacic, V. and Iakoucheva, L.M. (2012) Disease mutations in disordered regions — exception to the rule? *Mol. Biosyst.* **8**, 27–32 doi:10.1039/C1MB05251A
24. 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 doi:10.1016/j.sbi.2011.03.011
25. Pajkos, M., Mészáros, B., Simon, I. and Dosztányi, Z. (2012) Is there a biological cost of protein disorder? Analysis of cancer-associated mutations. *Mol. Biosyst.* **8**, 296–307 doi:10.1039/C1MB05246B
26. Uyar, B., Weatheritt, R.J., Dinkel, H., Davey, N.E. and Gibson, T.J. (2014) Proteome-wide analysis of human disease mutations in short linear motifs: neglected players in cancer? *Mol. Biosyst.* **10**, 2626–2642 doi:10.1039/C4MB00290C
27. Gsponer, J. and Babu, M.M. (2012) Cellular strategies for regulating functional and nonfunctional protein aggregation. *Cell Rep.* **2**, 1425–1437 doi:10.1016/j.celrep.2012.09.036
28. Latysheva, N.S. and Babu, M.M. (2016) Discovering and understanding oncogenic gene fusions through data intensive computational approaches. *Nucleic Acids Res.* **44**, 4487–4503 doi:10.1093/nar/gkw282
29. Chothia, C. (1974) Hydrophobic bonding and accessible surface area in proteins. *Nature* **248**, 338–339 doi:10.1038/248338a0
30. Chothia, C. (1975) Structural invariants in protein folding. *Nature* **254**, 304–308 doi:10.1038/254304a0
31. Uversky, V.N., Gillespie, J.R. and Fink, A.L. (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins: Struct., Funct., Bioinf.* **41**, 415–427 doi:10.1002/1097-0134(20001115)41:3<415::AID-PROT130>3.0.CO;2-7
32. Dyson, H.J. and Wright, P.E. (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**, 197–208 doi:10.1038/nrm1589

33   Burger, V.M., Nolasco, D.O. and Stultz, C.M. (2016) Expanding the range of protein function at the far end of the order-structure continuum. *J. Biol. Chem.* **291**, 6706–6713 doi:10.1074/jbc.R115.692590

34   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 doi:10.1021/acs.chemrev.5b00562

35   Mao, A.H., Crick, S.L., Vitalis, A., Chicoine, C.L. and Pappu, R.V. (2010) Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc. Natl Acad. Sci. USA* **107**, 8183–8188 doi:10.1073/pnas.0911107107

36   Muller-Spath, S., Soranno, A., Hirschfeld, V., Hofmann, H., Ruegger, S., Reymond, L. et al. (2010) Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc. Natl Acad. Sci. USA* **107**, 14609–14614 doi:10.1073/pnas.1001743107

37   Marsh, J.A. and Forman-Kay, J.D. (2010) Sequence determinants of compaction in intrinsically disordered proteins. *Biophys. J.* **98**, 2383–2390 doi:10.1016/j.bpj.2010.02.006

38   Das, R.K., Ruff, K.M. and Pappu, R.V. (2015) Relating sequence encoded information to form and function of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **32**, 102–112 doi:10.1016/j.sbi.2015.03.008

39   Das, R.K., Huang, Y., Phillips, A.H., Kriwacki, R.W. and Pappu, R.V. (2016) Cryptic sequence features within the disordered protein p27 Kip1 regulate cell cycle signaling. *Proc. Natl Acad. Sci. USA* **113**, 5616–5621 doi:10.1073/pnas.1516277113

40   Das, R.K. and Pappu, R.V. (2013) Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl Acad. Sci. USA* **110**, 13392–13397 doi:10.1073/pnas.1304749110

41   Van Roey, K., Uyar, B., Weatheritt, R.J., Dinkel, H., Seiler, M., Budd, A. et al. (2014) Short linear motifs: ubiquitous and functionally diverse protein interaction modules directing cell regulation. *Chem. Rev.* **114**, 6733–6778 doi:10.1021/cr400585q

42   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 doi:10.1016/j.molcel.2014.05.032

43   Davey, N.E., Van Roey, K., Weatheritt, R.J., Toedt, G., Uyar, B., Altenberg, B. et al. (2012) Attributes of short linear motifs. *Mol. Biosyst.* **8**, 268–281 doi:10.1039/C1MB05231D

44   Diella, F., Haslam, N., Chica, C., Budd, A., Michael, S., Brown, N.P. et al. (2008) Understanding eukaryotic linear motifs and their role in cell signaling and regulation. *Front. Biosci.* **13**, 6580–6603 doi:10.2741/3175

45   Schmid, E.M. and McMahon, H.T. (2007) Integrating molecular and network biology to decode endocytosis. *Nature* **448**, 883–888 doi:10.1038/nature06031

46   Praefcke, G.J.K., Ford, M.G.J., Schmid, E.M., Olesen, L.E., Gallop, J.L., Peak-Chew, S.-Y. et al. (2004) Evolving nature of the AP2 α-appendage hub during clathrin-coated vesicle endocytosis. *EMBO J.* **23**, 4371–4383 doi:10.1038/sj.emboj.7600445

47   Van Roey, K., Gibson, T.J. and Davey, N.E. (2012) Motif switches: decision-making in cell regulation. *Curr. Opin. Struct. Biol.* **22**, 378–385 doi:10.1016/j.sbi.2012.03.004

48   Bah, A., Vernon, R.M., Siddiqui, Z., Krzeminski, M., Muhandiram, R., Zhao, C. et al. (2015) Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* **519**, 106–109 doi:10.1038/nature13999

49   Venkatakrishnan, A.J., Flock, T., Prado, D.E., Oates, M.E., Gough, J. and Madan, B.M. (2014) Structured and disordered facets of the GPCR fold. *Curr. Opin. Struct. Biol.* **27**, 129–137 doi:10.1016/j.sbi.2014.08.002

50   Dinkel, H., Van Roey, K., Michael, S., Kumar, M., Uyar, B., Altenberg, B. et al. (2016) ELM 2016—data update and new functionality of the eukaryotic linear motif resource. *Nucleic Acids Res.* **44**, D294–D300 doi:10.1093/nar/gkv1291

51   Seet, B.T., Dikic, I., Zhou, M.-M. and Pawson, T. (2006) Reading protein modifications with interaction domains. *Nat. Rev. Mol. Cell Biol.* **7**, 473–483 doi:10.1038/nrm1960

52   Gsponer, J. and Babu, M.M. (2009) The rules of disorder or why disorder rules. *Prog. Biophys. Mol. Biol.* **99**, 94–103 doi:10.1016/j.pbiomolbio.2009.03.001

53   van der Lee, R., Lang, B., Kruse, K., Gsponer, J., Sánchez de Groot, N., Huynen, M.A. et al. (2014) Intrinsically disordered segments affect protein half-life in the cell and during evolution. *Cell Rep.* **8**, 1832–1844 doi:10.1016/j.celrep.2014.07.055

54   Fishbain, S., Inobe, T., Israeli, E., Chavali, S., Yu, H., Kago, G. et al. (2015) Sequence composition of disordered regions fine-tunes protein half-life. *Nat. Struct. Mol. Biol.* **22**, 214–221 doi:10.1038/nsmb.2958

55   Prakash, S., Inobe, T., Hatch, A.J. and Matouschek, A. (2009) Substrate selection by the proteasome during degradation of protein complexes. *Nat. Chem. Biol.* **5**, 29–36 doi:10.1038/nchembio.130

56   Inobe, T., Fishbain, S., Prakash, S. and Matouschek, A. (2011) Defining the geometry of the two-component proteasome degron. *Nat. Chem. Biol.* **7**, 161–167 doi:10.1038/nchembio.521

57   Sasaki, K., Dockerill, S., Adamiak, D.A., Tickle, I.J. and Blundell, T. (1975) X-ray analysis of glucagon and its relationship to receptor binding. *Nature* **257**, 751–757 doi:10.1038/257751a0

58   Guharoy, M., Bhowmick, P., Sallam, M. and Tompa, P. (2016) Tripartite degrons confer diversity and specificity on regulated protein degradation in the ubiquitin-proteasome system. *Nat. Commun.* **7**, 10239 doi:10.1038/ncomms10239

59   Guharoy, M., Bhowmick, P. and Tompa, P. (2016) Design principles involving protein disorder facilitate specific substrate selection and degradation by the ubiquitin-proteasome system. *J. Biol. Chem.* **291**, 6723–6731 doi:10.1074/jbc.R115.692665

60   Flock, T, Weatheritt, R.J., Latysheva, N.S. and Babu, M.M. (2014) Controlling entropy to tune the functions of intrinsically disordered regions. *Curr. Opin. Struct. Biol.* **26**, 62–72 doi:10.1016/j.sbi.2014.05.007

61   Mittag, T., Orlicky, S., Choy, W.-Y., Tang, X., Lin, H., Sicheri, F. et al. (2008) Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor. *Proc. Natl Acad. Sci. USA* **105**, 17772–17777 doi:10.1073/pnas.0809222105

62   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 doi:10.1021/bi8006803

63   Jakob, U., Kriwacki, R. and Uversky, V.N. (2014) Conditionally and transiently disordered proteins: awakening cryptic disorder to regulate protein function. *Chem. Rev.* **114**, 6779–6805 doi:10.1021/cr400459c

64   Rogers, J.M., Oleinikovas, V., Shammas, S.L., Wong, C.T., De Sancho, D., Baker, C.M. et al. (2014) Interplay between partner and ligand facilitates the folding and binding of an intrinsically disordered protein. *Proc. Natl Acad. Sci. USA* **111**, 15420–15425 doi:10.1073/pnas.1409122111

65 Shammas, S.L., Travis, A.J. and Clarke, J. (2013) Remarkably fast coupled folding and binding of the intrinsically disordered transactivation domain of cMyb to CBP KIX. *J. Phys. Chem. B* **117**, 13346–13356 doi:10.1021/jp404267e

66 Kragelund, B.B., Schenstrøm, S.M., Rebula, C.A., Panse, V.G. and Hartmann-Petersen, R. (2016) DSS1/Sem1, a multifunctional and intrinsically disordered protein. *Trends Biochem. Sci.* **41**, 446–459 doi:10.1016/j.tibs.2016.02.004

67 Lobley, A., Swindells, M.B., Orengo, C.A. and Jones, D.T. (2007) Inferring function using patterns of native disorder in proteins. *PLoS Comput. Biol.* **3**, e162 doi:10.1371/journal.pcbi.0030162

68 Dunker, A.K., Oldfield, C.J., Meng, J., Romero, P., Yang, J.Y., Chen, J. et al. (2008) The unfoldomics decade: an update on intrinsically disordered proteins. *BMC Genomics* **9**(Suppl. 2), S1 doi:10.1186/1471-2164-9-S2-S1

69 Xie, H., Vucetic S., Iakoucheva, L.M., Oldfield, C.J., Dunker, A.K., Obradovic, Z. et al. (2007) Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications and diseases associated with intrinsically disordered proteins. *J. Proteome Res.* **6**, 1917–1932 doi:10.1021/pr060394e

70 Dunker, A.K., Bondos, S.E., Huang, F. and Oldfield, C.J. (2015) Intrinsically disordered proteins and multicellular organisms. *Semin. Cell Dev. Biol.* **37**, 44–55 doi:10.1016/j.semcdb.2014.09.025

71 Wright, P.E. and Dyson, H.J. (2009) Linking folding and binding. *Curr. Opin. Struct. Biol.* **19**, 31–38 doi:10.1016/j.sbi.2008.12.003

72 Dyson, H.J. and Wright, P.E. (2002) Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* **12**, 54–60 doi:10.1016/S0959-440X(02)00289-0

73 Arai, M., Sugase, K., Dyson, H.J. and Wright, P.E. (2015) Conformational propensities of intrinsically disordered proteins influence the mechanism of binding and folding. *Proc. Natl Acad. Sci. USA* **112**, 9614–9619 doi:10.1073/pnas.1512799112

74 Liu, B., Chia, D., Csizmok, V., Farber, P., Forman-Kay, J.D. and Gradinaru, C.C. (2014) The effect of intrachain electrostatic repulsion on conformational disorder and dynamics of the Sic1 protein. *J. Phys. Chem. B* **118**, 4088–4097 doi:10.1021/jp500776v

75 Fuxreiter, M., Simon, I. and Bondos, S. (2011) Dynamic protein–DNA recognition: beyond what can be seen. *Trends Biochem. Sci.* **36**, 415–423 doi:10.1016/j.tibs.2011.04.006

76 Fuxreiter, M. (2012) Fuzziness: linking regulation to protein dynamics. *Mol. Biosyst.* **8**, 168–177 doi:10.1039/C1MB05234A

77 Tompa, P. and Fuxreiter, M. (2008) Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions. *Trends Biochem. Sci.* **33**, 2–8 doi:10.1016/j.tibs.2007.10.003

78 Sigalov, A.B., Zhuravleva, A.V. and Orekhov, V.Y. (2007) Binding of intrinsically disordered proteins is not necessarily accompanied by a structural transition to a folded form. *Biochimie* **89**, 419–421 doi:10.1016/j.biochi.2006.11.003

79 Dunker, A.K., Garner, E., Guilliot, S., Romero, P., Albrecht, K., Hart, J. et al. (1998) Protein disorder and the evolution of molecular recognition: theory, predictions and observations. *Pac. Symp. Biocomput.* 473–484 PMID: 9697205

80 Zhou, H.-X. (2012) Intrinsic disorder: signaling via highly specific but short-lived association. *Trends Biochem. Sci.* **37**, 43–48 doi:10.1016/j.tibs.2011.11.002

81 Shammas, S.L., Crabtree, M.D., Dahal, L., Wicky, B.I.M. and Clarke, J. (2016) Insights into coupled folding and binding mechanisms from kinetic studies. *J. Biol. Chem.* **291**, 6689–6695 doi:10.1074/jbc.R115.692715

82 Umezawa, K., Ohnuki, J., Higo, J. and Takano, M. (2016) Intrinsic disorder accelerates dissociation rather than association. *Proteins: Struct., Funct., Bioinf.* **84**, 1124–1133 doi:10.1002/prot.25057

83 Shoemaker, B.A., Portman, J.J. and Wolynes, P.G. (2000) Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc. Natl Acad. Sci. USA* **97**, 8868–8873 doi:10.1073/pnas.160259697

84 Shammas, S.L., Travis, A.J. and Clarke, J. (2014) Allostery within a transcription coactivator is predominantly mediated through dissociation rate constants. *Proc. Natl Acad. Sci. USA* **111**, 12055–12060 doi:10.1073/pnas.1405815111

85 Rogers, J.M., Wong, C.T. and Clarke, J. (2014) Coupled folding and binding of the disordered protein PUMA does not require particular residual structure. *J. Am. Chem. Soc.* **136**, 5197–5200 doi:10.1021/ja4125065

86 Wu, H. and Fuxreiter, M. (2016) The structure and dynamics of higher-order assemblies: amyloids, signalosomes, and granules. *Cell* **165**, 1055–1066 doi:10.1016/j.cell.2016.05.004

87 Toretsky, J.A. and Wright, P.E. (2014) Assemblages: functional units formed by cellular phase separation. *J. Cell Biol.* **206**, 579–588 doi:10.1083/jcb.201404124

88 Mitrea, D.M. and Kriwacki, R.W. (2016) Phase separation in biology; functional organization of a higher order. *Cell Commun. Signal.* **14**, 1 doi:10.1186/s12964-015-0125-7

89 Holehouse, A.S. and Pappu, R.V. (2015) Protein polymers: encoding phase transitions. *Nat. Mater.* **14**, 1083–1084 doi:10.1038/nmat4459

90 Banjade, S., Wu, Q., Mittal, A., Peeples, W.B., Pappu, R.V. and Rosen, M.K. (2015) Conserved interdomain linker promotes phase separation of the multivalent adaptor protein Nck. *Proc. Natl Acad. Sci. USA* **112**, E6426–E6435 doi:10.1073/pnas.1508778112

91 Su, X., Ditlev, J.A., Hui, E., Xing, W., Banjade, S. and Okrut, J. (2016) Phase separation of signaling molecules promotes T cell receptor signal transduction. *Science* **352**, 595–599 doi:10.1126/science.aad9964

92 Li, P., Banjade, S., Cheng, H.-C., Kim, S., Chen, B., Guo, L. et al. (2012) Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336–340 doi:10.1038/nature10879

93 Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A. et al. (2015) Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell* **57**, 936–947 doi:10.1016/j.molcel.2015.01.013

94 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 doi:10.1021/acs.chemrev.5b00548

95 Feric, M., Vaidya, N., Harmon, T.S., Mitrea, D.M., Zhu, L., Richardson, T.M. et al. (2016) Coexisting liquid phases underlie nucleolar subcompartments. *Cell* **165**, 1686–1697 doi:10.1016/j.cell.2016.04.047

96 Bergeron-Sandoval, L.-P., Safaee, N. and Michnick, S.W. (2016) Mechanisms and consequences of macromolecular phase separation. *Cell* **165**, 1067–1079 doi:10.1016/j.cell.2016.05.026

97 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 doi:10.1042/BJ20130545

98   Pak, C.W., Kosno, M., Holehouse, A.S., Padrick, S.B., Mittal, A., Ali, R. et al. (2016) Sequence determinants of intracellular phase separation by complex coacervation of a disordered protein. *Mol. Cell* **63**, 72–85 doi:10.1016/j.molcel.2016.05.042

99   Lin, Y., Protter, D.S.W., Rosen, M.K. and Parker, R. (2015) Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. *Mol. Cell* **60**, 208–219 doi:10.1016/j.molcel.2015.08.018

100 Frey, S., Richter, R.P. and Gorlich, D. (2006) FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science* **314**, 815–817 doi:10.1126/science.1132516

101 Xiang, S., Kato, M., Wu, L.C., Lin, Y., Ding, M., Zhang, Y. et al. (2015) The LC domain of hnRNPA2 adopts similar conformations in hydrogel polymers, liquid-like droplets, and nuclei. *Cell* **163**, 829–839 doi:10.1016/j.cell.2015.10.040

102 Sanchez de Groot, N., Torrent, M., Villar-Piqué, A., Lang, B., Ventura, S., Gsponer, J. et al. (2012) Evolutionary selection for protein aggregation. *Biochem. Soc. Trans.* **40**, 1032–1037 doi:10.1042/BST20120160

103 Motlagh, H.N., Wrabl, J.O., Li, J. and Hilser, V.J. (2014) The ensemble nature of allostery. *Nature* **508**, 331–339 doi:10.1038/nature13001

104 Hilser, V.J. and Thompson, E.B. (2007) Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. *Proc. Natl Acad. Sci. USA* **104**, 8311–8315 doi:10.1073/pnas.0700329104

105 Motlagh, H.N., Li, J., Thompson, E.B. and Hilser, V.J. (2012) Interplay between allostery and intrinsic disorder in an ensemble. *Biochem. Soc. Trans.* **40**, 975–980 doi:10.1042/BST20120163

106 Buljan, M., Chalancon, G., Dunker, A.K., Bateman, A., Balaji, S., Fuxreiter, M. et al. (2013) Alternative splicing of intrinsically disordered regions and rewiring of protein interactions. *Curr. Opin. Struct. Biol.* **23**, 443–450 doi:10.1016/j.sbi.2013.03.006

107 Buljan, M., Chalancon, G., Eustermann, S., Wagner, G.P., Fuxreiter, M., Bateman, A. et al. (2012) Tissue-specific splicing of disordered segments that embed binding motifs rewires protein interaction networks. *Mol. Cell* **46**, 871–883 doi:10.1016/j.molcel.2012.05.039

108 Weatheritt, R.J., Gibson, T.J. and Babu, M.M. (2014) Asymmetric mRNA localization contributes to fidelity and sensitivity of spatially localized systems. *Nat. Struct. Mol. Biol.* **21**, 833–839 doi:10.1038/nsmb.2876

109 Gsponer, J., Futschik, M.E., Teichmann, S.A. and Babu, M.M. (2008) Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. *Science* **322**, 1365–1368 doi:10.1126/science.1163581

110 Keren, H., Lev-Maor, G. and Ast, G. (2010) Alternative splicing and evolution: diversification, exon definition and function. *Nat. Rev. Genet.* **11**, 345–355 doi:10.1038/nrg2776

111 Merkin, J., Russell, C., Chen, P. and Burge, C.B. (2012) Evolutionary dynamics of gene and isoform regulation in mammalian tissues. *Science* **338**, 1593–1599 doi:10.1126/science.1228186

112 de Klerk, E. and 't Hoen, P.A. (2015) Alternative mRNA transcription, processing, and translation: insights from RNA sequencing. *Trends Genet.* **31**, 128–139 doi:10.1016/j.tig.2015.01.001

113 Pan, Q., Shai, O., Lee, L.J., Frey, B.J. and Blencowe, B.J. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **40**, 1413–1415 doi:10.1038/ng.259

114 Hao, Y., Colak, R., Teyra, J., Corbi-Verge, C., Ignatchenko, A., Hahne, H. et al. (2015) Semi-supervised learning predicts approximately one third of the alternative splicing isoforms as functional proteins. *Cell Rep.* **12**, 183–189 doi:10.1016/j.celrep.2015.06.031

115 Abascal, F., Ezkurdia, I., Rodriguez-Rivas, J., Rodriguez, J.M., del Pozo, A., Vázquez, J. et al. (2015) Alternatively spliced homologous exons have ancient origins and are highly expressed at the protein level. *PLoS Comput. Biol.* **11**, e1004325 doi:10.1371/journal.pcbi.1004325

116 Tress, M.L., Bodenmiller, B., Aebersold, R. and Valencia, A. (2008) Proteomics studies confirm the presence of alternative protein isoforms on a large scale. *Genome Biol.* **9**, R162 doi:10.1186/gb-2008-9-11-r162

117 Tress, M.L., Martelli, P.L., Frankish, A., Reeves, G.A., Wesselink, J.J., Yeats, C. et al. (2007) The implications of alternative splicing in the ENCODE protein complement. *Proc. Natl Acad. Sci. USA* **104**, 5495–5500 doi:10.1073/pnas.0700800104

118 Hegyi, H., Kalmar, L., Horvath, T. and Tompa, P. (2011) Verification of alternative splicing variants based on domain integrity, truncation length and intrinsic protein disorder. *Nucleic Acids Res.* **39**, 1208–1219 doi:10.1093/nar/gkq843

119 Romero, P.R., Zaidi, S., Fang, Y.Y., Uversky, V.N., Radivojac, P., Oldfield, C.J. et al. (2006) Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms. *Proc. Natl Acad. Sci. USA* **103**, 8390–8395 doi:10.1073/pnas.0507916103

120 Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C. et al. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470–476 doi:10.1038/nature07509

121 Barbosa-Morais, N.L., Irimia, M., Pan, Q., Xiong, H.Y., Gueroussov, S., Lee, L.J. et al. (2012) The evolutionary landscape of alternative splicing in vertebrate species. *Science* **338**, 1587–1593 doi:10.1126/science.1230612

122 Ungewitter, E. and Scrable, H. (2010) Δ40p53 controls the switch from pluripotency to differentiation by regulating IGF signaling in ESCs. *Genes Dev.* **24**, 2408–2419 doi:10.1101/gad.1987810

123 Ellis, J.D., Barrios-Rodiles, M., Çolak, R., Irimia, M., Kim, T., Calarco, J.A. et al. (2012) Tissue-specific alternative splicing remodels protein-protein interaction networks. *Mol. Cell* **46**, 884–892 doi:10.1016/j.molcel.2012.05.037

124 Yang, X., Coulombe-Huntington, J., Kang, S., Sheynkman, G.M., Hao, T., Richardson, A. et al. (2016) Widespread expansion of protein interaction capabilities by alternative splicing. *Cell* **164**, 805–817 doi:10.1016/j.cell.2016.01.029

125 Corominas, R., Yang, X., Lin, G.N., Kang, S., Shen, Y., Ghamsari, L. et al. (2014) Protein interaction network of alternatively spliced isoforms from brain links genetic risk factors for autism. *Nat. Commun.* **5**, 3650 doi:10.1038/ncomms4650

126 Colak, R., Kim, T.H., Michaut, M., Sun, M., Irimia, M., Bellay, J. et al. (2013) Distinct types of disorder in the human proteome: functional implications for alternative splicing. *PLoS Comput. Biol.* **9**, e1003030 doi:10.1371/journal.pcbi.1003030

127 Weatheritt, R.J. and Gibson, T.J. (2012) Linear motifs: lost in (pre)translation. *Trends Biochem. Sci.* **37**, 333–341 doi:10.1016/j.tibs.2012.05.001

128 Weatheritt, R.J., Davey, N.E. and Gibson, T.J. (2012) Linear motifs confer functional diversity onto splice variants. *Nucleic Acids Res.* **40**, 7123–7131 doi:10.1093/nar/gks442

129 Trudeau, T., Nassar, R., Cumberworth, A., Wong, E.T.C., Woollard, G. and Gsponer, J. (2013) Structure and intrinsic disorder in protein autoinhibition. *Structure* **21**, 332–341 doi:10.1016/j.str.2012.12.013

130 Davis, M.J., Shin, C.J., Jing, N. and Ragan, M.A. (2012) Rewiring the dynamic interactome. *Mol. Biosyst.* **8**, 2054–2066, 2013 doi:10.1039/c2mb25050k

131 Gueroussov, S., Gonatopoulos-Pournatzis, T., Irimia, M., Raj, B., Lin, Z.-Y., Gingras, A.-C. et al. (2015) An alternative splicing event amplifies evolutionary differences between vertebrates. *Science* **349**, 868–873 doi:10.1126/science.aaa8381

132 Papasaikas, P. and Valcarcel, J. (2012) Splicing in 4D. *Science* **338**, 1547–1548 doi:10.1126/science.1233219

133 Martin, K.C. and Ephrussi, A. (2009) mRNA localization: gene expression in the spatial dimension. *Cell* **136**, 719–730 doi:10.1016/j.cell.2009.01.044

134 Holt, C.E. and Bullock, S.L. (2009) Subcellular mRNA localization in animal cells and why it matters. *Science* **326**, 1212–1216 doi:10.1126/science.1176488

135 Scott, J.D. and Pawson, T. (2009) Cell signaling in space and time: where proteins come together and when they're apart. *Science* **326**, 1220–1224 doi:10.1126/science.1175668

136 Ephrussi, A., Dickinson, L.K. and Lehmann, R. (1991) Oskar organizes the germ plasm and directs localization of the posterior determinant nanos. *Cell* **66**, 37–50 doi:10.1016/0092-8674(91)90137-N

137 Jung, H., Gkogkas C.G., Sonenberg, N. and Holt, C.E. (2014) Remote control of gene function by local translation. *Cell* **157**, 26–40 doi:10.1016/j.cell.2014.03.005

138 Lécuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T. et al. (2007) Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* **131**, 174–187 doi:10.1016/j.cell.2007.08.003

139 Pertz, O.C., Wang, Y., Yang, F., Wang, W., Gay, L.J., Gristenko, M.A. et al. (2008) Spatial mapping of the neurite and soma proteomes reveals a functional Cdc42/Rac regulatory network. *Proc. Natl Acad. Sci. USA* **105**, 1931–1936 doi:10.1073/pnas.0706545105

140 Wang, Y., Ding, S.-J., Wang, W., Jacobs, J.M., Qian, W.-J., Moore, R.J. et al. (2007) Profiling signaling polarity in chemotactic cells. *Proc. Natl Acad. Sci. USA* **104**, 8328–8333 doi:10.1073/pnas.0701103104

141 Cajigas, I.J., Tushev, G., Will, T.J., tom Dieck, S., Fuerst, N. and Schuman, E.M. (2012) The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. *Neuron* **74**, 453–466 doi:10.1016/j.neuron.2012.02.036

142 Oda, Y. (1999) Choline acetyltransferase: the structure, distribution and pathologic changes in the central nervous system. *Pathol. Int.* **49**, 921–937 doi:10.1046/j.1440-1827.1999.00977.x

143 Xiong, X., Huang, S., Zhang, H., Li, J., Shen, J., Xiong, J. et al. (2009) Enrichment and proteomic analysis of plasma membrane from rat dorsal root ganglions. *Proteome Sci.* **7**, 41 doi:10.1186/1477-5956-7-41

144 Gumy, L.F., Yeo, G.S.H., Tung, Y.-C.L., Zivraj, K.H., Willis, D., Coppola, G. et al. (2011) Transcriptome analysis of embryonic and adult sensory axons reveals changes in mRNA repertoire localization. *RNA* **17**, 85–98 doi:10.1261/rna.2386111

145 Nozumi, M., Togano, T., Takahashi-Niki, K., Lu, J., Honda, A., Taoka, M. et al. (2009) Identification of functional marker proteins in the mammalian growth cone. *Proc. Natl Acad. Sci. USA* **106**, 17211–17216 doi:10.1073/pnas.0904092106

146 Mili, S., Moissoglu, K. and Macara, I.G. (2008) Genome-wide screen reveals APC-associated RNAs enriched in cell protrusions. *Nature* **453**, 115–119 doi:10.1038/nature06888

147 Van Roey, K., Dinkel, H., Weatheritt, R.J., Gibson, T.J. and Davey, N.E. (2013) The switches.ELM resource: a compendium of conditional regulatory interaction interfaces. *Sci. Signal.* **6**, rs7 doi:10.1126/scisignal.2003345

148 Calabretta, S. and Richard, S. (2015) Emerging roles of disordered sequences in RNA-binding proteins. *Trends Biochem. Sci.* **40**, 662–672 doi:10.1016/j.tibs.2015.08.012

149 Aguzzi, A. and Altmeyer, M. (2016) Phase separation: linking cellular compartmentalization to disease. *Trends Cell Biol.* **26**, 547–558 doi:10.1016/j.tcb.2016.03.004

150 Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J. et al. (2015) Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123–133 doi:10.1016/j.cell.2015.09.015

151 Knowles, T.P.J., Vendruscolo, M. and Dobson, C.M. (2014) The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **15**, 384–396 doi:10.1038/nrm3810

152 Chiti, F. and Dobson, C.M. (2006) Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**, 333–366 doi:10.1146/annurev.biochem.75.101304.123901

153 Tantos, A., Kalmar, L. and Tompa, P. (2015) The role of structural disorder in cell cycle regulation, related clinical proteomics, disease development and drug targeting. *Expert Rev. Proteomics* **12**, 221–233 doi:10.1586/14789450.2015.1042866

154 Hegyi, H., Buday, L. and Tompa, P. (2009) Intrinsic structural disorder confers cellular viability on oncogenic fusion proteins. *PLoS Comput. Biol.* **5**, e1000552 doi:10.1371/journal.pcbi.1000552

155 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 doi:10.1016/j.cell.2009.04.029

156 Davey, N.E., Cyert, M.S. and Moses, A.M. (2015) Short linear motifs – ex nihilo evolution of protein regulation. *Cell Commun. Signal.* **13**, 43 doi:10.1186/s12964-015-0120-z

157 Soranno, A., Koenig, I., Borgia, M.B., Hofmann, H., Zosel, F., Nettels, D. et al. (2014) Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments. *Proc. Natl Acad. Sci. USA* **111**, 4874–4879 doi:10.1073/pnas.1322611111

158 Ou, L., Ferreira, A.M., Otieno, S., Xiao, L., Bashford, D. and Kriwacki, R.W. (2011) Incomplete folding upon binding mediates Cdk4/cyclin D complex activation by tyrosine phosphorylation of inhibitor p27 protein. *J. Biol. Chem.* **286**, 30142–30151 doi:10.1074/jbc.M111.244095

159 Rogers, J.M., Steward, A. and Clarke, J. (2013) Folding and binding of an intrinsically disordered protein: fast, but not 'diffusion-limited'. *J. Am. Chem. Soc.* **135**, 1415–1422 doi:10.1021/ja309527h

160 Teixeira, D., Sheth, U., Valencia-Sanchez, M.A., Brengues, M. and Parker, R. (2005) Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA* **11**, 371–382 doi:10.1261/rna.7258505

161 Hardy, J. and Gwinn-Hardy, K. (1998) Genetic classification of primary neurodegenerative disease. *Science* **282**, 1075–1079 doi:10.1126/science.282.5391.1075