

# Advantages of proteins being disordered

Zhirong Liu,<sup>1,2,3,4\*</sup> and Yongqi Huang<sup>1,4,5</sup>

<sup>1</sup>College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

<sup>2</sup>Beijing National Laboratory for Molecular Sciences (BNLMS), Peking University, Beijing 100871, China

<sup>3</sup>State Key Laboratory for Structural Chemistry of Unstable and Stable Species, Peking University, Beijing 100871, China

<sup>4</sup>Center for Quantitative Biology, Peking University, Beijing 100871, China

<sup>5</sup>Department of Structural Biology, St. Jude Children's Research Hospital, Memphis, Tennessee 38105

Received 22 January 2014; Accepted 10 February 2014

DOI: 10.1002/pro.2443

Published online 15 February 2014 proteinscience.org

**Abstract:** The past decade has witnessed great advances in our understanding of protein structure-function relationships in terms of the ubiquitous existence of intrinsically disordered proteins (IDPs) and intrinsically disordered regions (IDRs). The structural disorder of IDPs/IDRs enables them to play essential functions that are complementary to those of ordered proteins. In addition, IDPs/IDRs are persistent in evolution. Therefore, they are expected to possess some advantages over ordered proteins. In this review, we summarize and survey nine possible advantages of IDPs/IDRs: economizing genome/protein resources, overcoming steric restrictions in binding, achieving high specificity with low affinity, increasing binding rate, facilitating posttranslational modifications, enabling flexible linkers, preventing aggregation, providing resistance to non-native conditions, and allowing compatibility with more available sequences. Some potential advantages of IDPs/IDRs are not well understood and require both experimental and theoretical approaches to decipher. The connection with protein design is also briefly discussed.

**Keywords:** intrinsically disordered proteins; protein function; flexibility; protein-protein interaction; molecular recognition; protein design; drug design

## Introduction

Being a challenge or amendment to the conventional sequence-structure-function paradigm for proteins, intrinsically disordered proteins (IDPs) have attracted ever-increasing attention during the past decade.<sup>1–5</sup> IDPs are intriguing because they do not have ordered structures in the free state under physiological conditions but still possess biological functions. The structural disorder of IDPs and intrinsically disordered regions (IDRs) lies in their

distinct amino-acid sequences, that is, they usually have a low hydrophobicity content combined with a high net charge content.<sup>1,6,7</sup> Without folded structures, IDPs/IDRs exist in an ensemble of rapidly-changing conformations with a flat free-energy landscape<sup>8–10</sup> and exhibit almost unlimited structural heterogeneity.<sup>1,11</sup> On the other hand, most IDPs/IDRs undergo a disorder-to-order transition upon binding to their biological partners (i.e., coupled folding and binding),<sup>12,13</sup> although some remain disordered even in their bound state.<sup>14–17</sup>

The structural disorder of IDPs/IDRs enables them to play essential functions that are complementary to those of ordered proteins. Roughly, IDPs/IDRs can be classified into six broad functional classes,<sup>18,19</sup> including effectors, scavengers, assemblers, entropic

Grant sponsor: National Natural Science Foundation of China; Grant number: 20973016.

\*Correspondence to: Zhirong Liu, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China. E-mail: LiuZhiRong@pku.edu.cn

chains, display sites, and chaperones. A detailed bioinformatics analysis of the Swiss-Prot database revealed a positive correlation between IDPs/IDRs and 238 function keywords, the majority of which were related to signaling and regulation of key cellular processes.<sup>20</sup> Due to the essential functions of IDPs/IDRs, it is not unexpected that they are abundant in all species.<sup>21–24</sup> In particular, the average fraction of disordered residues predicted in eukaryotes is higher than that in prokaryotes, suggesting the importance of IDPs/IDRs in evolution. The abundant existence of IDPs/IDRs and their vital functions, as well as the fact that usual techniques used in characterizing conventional proteins may be not applicable for IDPs/IDRs, make IDPs/IDRs an important subfield of molecular structural biology. In fact, IDPs/IDRs are taking a due place in mainstream studies. For example, predicting disordered regions has become a part of the critical assessment of structure prediction (CASP).<sup>25</sup>

As IDPs/IDRs perform functions complementary to those of ordered proteins and are persistent in evolution, they might possess some advantages over ordered proteins (and at the same time possess some disadvantages). In this article, we give a brief review of the advantages of IDPs/IDRs provided by their conformational flexibility. Nine possible advantages are summarized and surveyed (Fig. 1): (1) economizing genome and protein resources; (2) overcoming steric restrictions in binding; (3) achieving high specificity with low affinity; (4) increasing binding rate; (5) facilitating posttranslational modifications; (6) enabling flexible linkers; (7) preventing aggregation; (8) providing resistance to non-native conditions; (9) allowing compatibility with more available sequences. Some advantages are self-evident, for example, the potency of acting as flexible linkers is incompatible with ordered structures. Some other advantages, however, are less clear and knowledge on them is still far from complete. It is acknowledged that the current review addresses only a narrow aspect of IDPs' properties. Readers are directed to several excellent reviews<sup>1,2,26–30</sup> for a more comprehensive understanding of IDPs.

### **Advantage 1: Economizing Genome and Protein Resources**

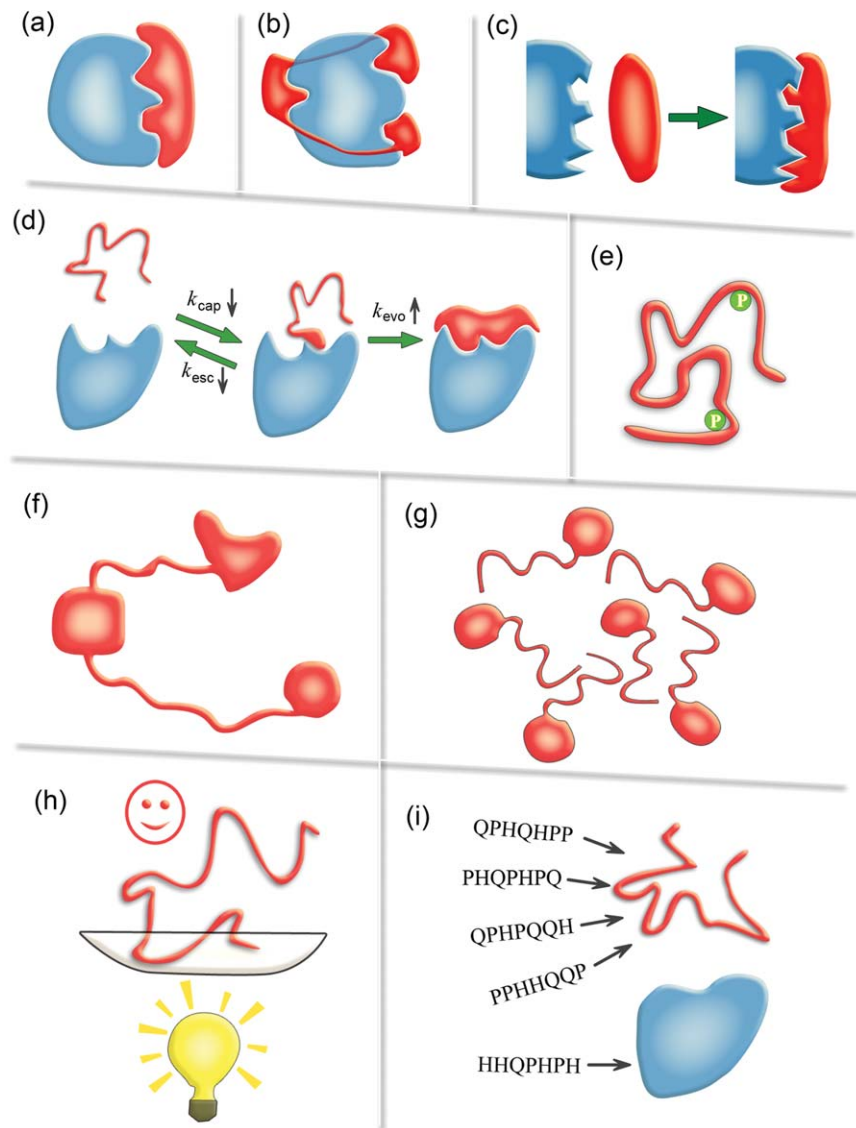
The interface area of IDPs in protein-protein complexes is similar in size to that of ordered proteins, but the sequence used to create the same interface area is much shorter for IDPs [Fig. 1(a)].<sup>31,32</sup> It originates from the fact that IDPs have extended structures which are stable only in the complexes but not in monomers. As a result, IDPs possess greater interface area per residue than ordered proteins. It has been estimated that the protein size would be two to three times larger if IDPs were required to be as stable as monomers, which would unavoidably

exacerbate cellular crowding and increase the sequence size.<sup>31</sup> Therefore, IDPs/IDRs are advantageous in saving genome and protein resources. This feature may be more critical for species with small genome size. For example, viruses were shown to have the widest spread of IDPs/IDRs content compared with three domains of life, and Avian carcinoma virus possesses the highest disorder content of 77.3% among 3500 proteomes.<sup>21</sup>

Intrinsic disorder may also help to reduce the genome size in other mechanisms.<sup>27</sup> The number of genes in the human genome is significantly smaller than that of the diverse proteins necessary in higher organisms.<sup>27</sup> Alternative splicing is a way to avoid over-large genomes, in which multiple proteins can be produced from a single gene. Regions affected by alternative splicing are frequently biased to be disordered,<sup>33</sup> which helps to avoid structural disruption in the spliced proteins. “Moonlighting” is another solution to control the genome size, in which a single protein is capable of carrying out more than one function. IDPs/IDRs provide an important mechanism for moonlighting since they can use the same or overlapping regions to fulfill distinct functions by adopting different conformations upon binding.<sup>34</sup> For example, the intrinsically disordered domain of the sulfhydryl oxidase ALR performs dual functions: as a mitochondrial targeting signal in the cytosol and as a crucial recognition site in the disulfide relay system of intermembrane space.<sup>17</sup> Another example of multifunctional IDPs is anhydrin which acts as a chaperone and an endonuclease.<sup>35</sup>

### **Advantage 2: Overcoming Steric Restrictions in Binding**

The high chain flexibility and conformational disorder of IDPs/IDRs enables them to form complementary binding interfaces with their targets more easily. Via coupled folding and binding, IDPs/IDRs can overcome steric restrictions by protruding into the concavities of partners or wrapping around partners in ways difficult for ordered proteins [Fig. 1(b)]. It was found that the partner interfaces of MoRFs (the short segments of IDPs that perform molecular recognition and usually undergo disorder-to-order transitions upon binding) are significantly less flat than those of ordered proteins.<sup>36</sup> When measured by the RMSD of interface atoms to the fitting plane, the interface planarity of IDPs partners is 3.76 Å while that of ordered proteins is 2.98 Å.<sup>36</sup> The binding modes of IDPs/IDRs are highly diverse and create multifarious unusual complexes: wrappers, chameleons, penetrators, huggers, intertwined strings, long cylindrical containers, connectors, armature, etc.<sup>37</sup> For example, the intrinsically disordered calpastatin (endogenous inhibitor of calpain) wraps around and binds to its target on three surfaces to form a flexible wrapper.<sup>38</sup>



**Figure 1.** Schematic diagrams illustrating nine possible advantages of IDPs/IDRs. (a) Economizing genome and protein resources: IDPs/IDRs use smaller protein size to afford the same interface area as ordered proteins. (b) Overcoming steric restrictions in binding: IDPs/IDRs can overcome steric restrictions in binding complexes by protruding into partners or wrapping around them in ways difficult for ordered proteins. (c) Achieving high specificity with low affinity: the highly complementary binding interfaces and the unfavorable conformational-entropy changes result in high specificity and low affinity. (d) Increasing binding rate: the kinetic advantage of IDPs in a binding process stems from a faster evolution step from the encounter complex and a slower escaping rate. (e) Facilitating posttranslational modifications: the conformational flexibility of IDPs/IDRs greatly facilitates the exposure of modification sites and their binding to modifying enzymes. (f) Enabling flexible linkers: the lack of ordered structures makes IDPs/IDRs dominant in flexible linkers of proteins which are obviously out of reach of ordered proteins. (g) Preventing aggregation: IDPs/IDRs possess favorable interactions with water and are inherently advantageous in preventing aggregation. (h) Providing resistance to non-native conditions: you cannot unfold what is already unfolded. (i) Allowing compatibility with more available sequences: the sequence space of IDPs/IDRs is expected to be larger than that of ordered proteins.

After folding coupled with binding, the interface structure of IDPs/IDRs is still more flexible than that of ordered proteins. The average crystallographic B-factor of interfaces for IDPs ( $51 \text{ \AA}^2$ ) in complex structures was much larger than that of ordered proteins ( $21 \text{ \AA}^2$ ), indicating interface atoms of IDPs remain highly dynamic.<sup>39</sup> IDPs/IDRs also have a high content of the poly-L-proline type II (PPII) helix which is markedly more flexible in com-

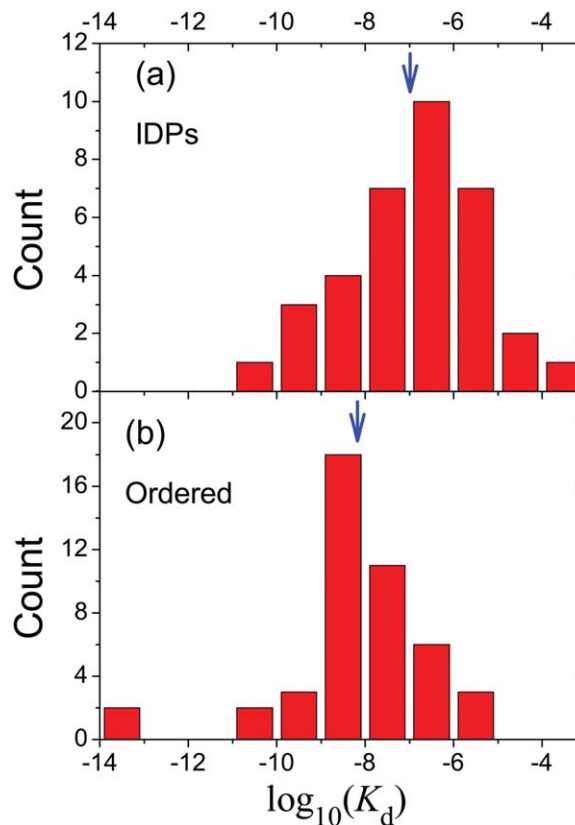
parison with  $\alpha$ -helix and  $\beta$ -sheet.<sup>40,41</sup> In addition to the complete folding upon binding where IDPs/IDRs adopt ordered structures, some IDPs/IDRs experience incomplete folding where a significant part remains disordered in the bound state, forming dynamic or fuzzy complexes.<sup>42–44</sup> In the extreme case (random fuzziness), IDPs remain entirely disordered in the bound state and resemble a “binding cloud” where multiple binding sites are dynamically

distributed.<sup>1,42</sup> An example is the Sic1-Cdc4 complex, where a few binding motifs of Sic1 interact with Cdc4 in a dynamic equilibrium, one at a time, and the parts of Sic1 not interacting with Cdc4 remain disordered.<sup>14,45</sup> For the binding with small molecule ligands, IDPs remain disordered and the ligands bind at different sites along the protein chains, so it may be described as ligand clouds around protein clouds.<sup>10</sup>

### Advantage 3: Achieving High Specificity With Low Affinity

Combining high specificity with low affinity is a widely-mentioned advantage of IDPs/IDRs,<sup>4,12,46,47</sup> which is a useful pair of properties for a reversible signaling transduction by enabling rapid association/dissociation with the partner without excessive binding strength.

The low affinity of IDPs/IDRs originates from the coupled folding and binding. The ordered (folded) state is unstable for IDPs/IDRs when they are free in solution. In a coupled folding and binding process, the folding increases the free energy while the binding of IDPs/IDRs to their targets decreases the free energy. As a result, the net free-energy change in such a coupled folding-binding process is smaller than that in a pure binding process, leading to a lower affinity. The low affinity allows IDPs to have a high dissociation rate, which can be essential for regulatory and signaling functions. Based on a dataset of protein-protein associations for 35 IDPs and 45 ordered proteins,<sup>48</sup> we showed that the distribution of the equilibrium dissociation constants for IDPs is slightly different from that for ordered proteins (Fig. 2), where  $\log K_d = -7.0 \pm 1.6$  for IDPs and  $\log K_d = -8.2 \pm 1.7$  for ordered proteins. In another benchmark dataset, it was found that 62% of the complexes with lower binding affinity contain regions formed via coupled folding and binding.<sup>49</sup> The relatively low affinity of IDPs/IDRs also suggested that high-affinity binding proteins can tolerate more structural disorder, consistent with the difference in the content of IDPs/IDRs of prokaryotic and eukaryotic genomes.<sup>50</sup> It is noted that although the disorder-to-order transition of IDPs/IDRs decreases the conformational entropy, IDPs/IDRs usually possess more favorable interface interactions than ordered proteins.<sup>32,39</sup> Furthermore, the binding free energy is affected by many other factors.<sup>51,52</sup> The affinity of either IDPs/IDRs or ordered proteins is highly variable and covers a wide range. So the affinity difference between IDPs/IDRs and ordered proteins applies to average values, not individual cases. Actually, some IDPs exhibit extremely high affinities. For example, the binding affinity of the intrinsically disordered antitoxin CcdA to the gyrase poison CcdB is in the pico-molar range, much higher than that of most ordered proteins.<sup>53,54</sup>



**Figure 2.** Distribution of binding affinity for (a) IDPs and (b) ordered proteins. The average values of  $\log K_d$  are indicated by arrows. The analysis was based on the dataset compiled in Ref. 48.

The high specificity of IDPs/IDRs is generally believed to be related to their interface characteristics.<sup>3,4,12,27</sup> The binding specificity is mainly determined by the size and complementarity of the binding interface. Due to the chain flexibility and the coupled folding-and-binding, as mentioned above, the interaction interfaces of IDPs/IDRs to their targets are highly complementary [Fig. 1(c)]. Although IDPs/IDRs usually use short segments in molecular recognitions, the extended structures formed in the complexes enable them to achieve large binding interfaces. Therefore, IDPs/IDRs are expected to exhibit high specificity in molecular recognitions. High specificity is necessary for their critical functions in signal recognition, transduction, and regulation.

However, there is still some debate on the specificity of IDPs/IDRs.<sup>47,55</sup> The above reasoning is not consistent with the conventional opinion that more flexible proteins are more promiscuous,<sup>47,55</sup> that is, if structural flexibility enables IDPs/IDRs to form complementary interfaces with their cognate targets, the same property could also enable them to fit the surfaces of noncognate targets and thus IDPs/IDRs would lose their interaction specificity. Under interface-interaction perturbations (to mimic the



structural difference between cognate and noncognate targets), IDPs/IDRs can either adjust their conformations to maintain the interface complementarity to stabilize the enthalpy or become conformationally more dynamic to stabilize the entropy, and as a result, IDPs/IDRs have a more complete enthalpy–entropy compensation to reduce the influence of the perturbations.<sup>55</sup> Indeed, an analysis on the thermodynamic data of interface mutation verified that IDPs had higher enthalpy–entropy compensation than ordered proteins,<sup>55</sup> suggesting IDPs to be more malleable.

To clarify the specificity difference between IDPs and ordered proteins, a coarse-grained molecular dynamics (MD) study was conducted to calculate the mutation free energy ( $\Delta\Delta G$ ) of IDPs and ordered proteins under the same perturbations.<sup>55</sup> It was found that  $\Delta\Delta G$  for IDPs is smaller than that for ordered proteins by an averaged difference of  $0.16 \pm 0.26$  kcal/mol. This difference is smaller than what would be expected from the enthalpy–entropy compensation mentioned above. The reason is the same perturbation caused larger mutation enthalpy ( $\Delta\Delta H$ ) in IDPs.<sup>55</sup> Recently, an analysis of protein–protein complexes from the Protein Data Bank revealed that polar interface interactions make a larger contribution to the specificity in IDPs, and the follow-up computational alanine scanning of both hydrophobic and charged residues confirmed that these mutations caused larger  $\Delta\Delta H$  in IDPs.<sup>56</sup> Therefore, IDPs are likely to possess high specificity, although maybe not as high as ordered proteins.

High specificity may be critical for the biological functions of IDPs/IDRs. On the other hand, since the specificity of IDPs/IDRs may not be as high as that of ordered proteins, IDPs/IDRs may be more promiscuous in molecular recognitions and need to be tightly controlled in cells. Indeed, the abundance of IDPs/IDRs is regulated so that they are available in appropriate amounts and at appropriate times as needed.<sup>28,57</sup> The promiscuity of IDPs/IDRs also makes them a major determinant of dosage sensitivity where promiscuous interactions drive pathological changes in response to gene over-expression.<sup>58</sup> In the absence of tight regulation of IDPs/IDRs, severe outcomes may result, for example, the occurrence of various diseases.<sup>59,60</sup>

A closely related characteristic is multiple specificity—the capacity of interacting with many partners.<sup>61</sup> The high plasticity and malleability of IDPs/IDRs enable them to bind to multiple partners more readily by changing conformations or interaction regions according to the templates provided by different targets.<sup>62,63</sup> The cyclin-dependent kinase (CDK) inhibitor p21, which is a very early example of IDPs, binds to different CDK/cyclin complexes during the cell cycle and the binding diversity is mediated by the adaptability of its LH region.<sup>64</sup> The

induced structures of the same IDP/IDRs may be different in binding with different partners, for example, the C-terminus of p53 becomes a helix, a sheet, or a coil in different complexes.<sup>65</sup> Even when the induced structures of the same IDP/IDRs in different complexes are similar, the utilized anchor residues may be totally different which are important in controlling the specificity.<sup>66</sup> Multispecificity is useful for signaling and regulation. IDPs/IDRs are commonly involved in the hubs of protein–protein interaction networks,<sup>67–70</sup> where one IDP/IDR may bind to many partners or many IDPs/IDRs bind to one partner. The chaperone function of IDPs is also related to the multispecificity where it is vital to bind multiple aggregation-sensitive client proteins.<sup>71</sup> In addition, multispecificity is also a basis of the moonlighting mentioned above.

#### Advantage 4: Increasing Binding Rate

The binding rate of proteins to their targets is important for diverse processes ranging from enzyme catalysis/inhibition to cellular signaling. Compared with ordered proteins, IDPs/IDRs possess a kinetic advantage, that is, their binding rate is relatively higher.

The first insight into the kinetic advantage of IDPs/IDRs was provided by Shoemaker *et al.* in terms of the “fly-casting mechanism.”<sup>72</sup> By virtue of extended conformations, IDPs/IDRs could bind to their targets weakly at a larger distance (capture radius) from the actual binding site. Then they “reel in” the targets to complete the binding process while simultaneously folding.<sup>72</sup> Later, Huang and Liu performed a critical assessment of this mechanism by investigating the inherent influence of chain disorder on the binding kinetics of IDPs/IDRs.<sup>48</sup> Via Langevin dynamics simulations with a Gō-like coarse-grained model, it was found that although IDPs/IDRs possess a larger capture radius, the larger capture radius inevitably leads to a slower translational diffusion. As a result, the rate constant for forming the encounter complex actually decreased modestly as the chain flexibility increases, rather than increased as originally proposed. The real source for the kinetic advantage of IDPs/IDRs lies in the second step, where the encounter complex evolves into the final bound state faster and escapes to the unbound state slower than ordered proteins [Fig. 1(d)].<sup>48</sup> In addition, the available experimental data on the binding kinetics of IDPs and ordered proteins were compared to show the general difference in real systems.<sup>48</sup> On average, IDPs bind two to three times as fast as ordered proteins under the same affinity. It was also noted that the binding rate spans a large range of magnitude so that comparison of a single datum is meaningless.

Protein binding rates are influenced by various factors,<sup>73</sup> for example, nonspecific binding and

electrostatic interactions. Due to their high chain flexibility, IDPs/IDRs may possess more nonnative interactions in the folding and binding processes than ordered proteins. It was shown that nonnative hydrophobic interactions greatly amplify the kinetic advantages of IDPs/IDRs.<sup>74</sup> IDPs/IDRs usually contain a high content of charge residues, and electrostatic interactions were also found to contribute to their kinetic advantage via the electrostatic steering mechanism.<sup>75</sup> Evidently, the rate of the intrinsically disordered WASP GBD (the GTPase binding domain of the Wiskott-Aldrich syndrome protein) upon binding its target (Cdc42) is highly dependent on ionic strength in the experimental study.<sup>76</sup> In the interactions between IDPs and DNA, it was shown that electrostatic steering and protein flexibility synergistically couple to achieve rapidity in recognition.<sup>77</sup> Recently, Ganguly *et al.* showed that long-range electrostatic interactions not only enhance the encounter rate via the electrostatic steering but also promote the folding-competent topologies in the encounter complexes, allowing rapid formation of short-range native interactions to form the final bound state.<sup>78,79</sup>

Binding kinetics is important for functions at molecular and system levels.<sup>73,80</sup> Association between signaling proteins and their cellular targets should be both fast and transient. The enhanced binding rate of IDPs/IDRs, together with the low affinity, provides an advantageous solution to these needs.<sup>47</sup> For the intrinsically disordered translocation domain of Colicin E9, although its binding with TolB is weaker than the competitors, it wins in the competitive recruitment of TolB by its much higher binding rate.<sup>81</sup> In nonsense-mediated mRNA decay (NMD), it was suggested that a fly-casting mechanism enabled by long disordered regions in NMD complexes is exploited for effective long-range communication.<sup>82</sup>

#### **Advantage 5: Facilitating Post-translational Modifications**

Posttranslational modifications (PTMs) are an important means to regulate functions of proteins. The conformational flexibility of IDPs/IDRs greatly facilitates exposure of their modification sites and binding to the modifying enzymes [Fig. 1(e)]. The flexibility of IDPs/IDRs also makes it possible for a single enzyme to bind to and modify sites in a wide variety of proteins. In contrast, PTMs on ordered proteins may be restricted by site accessibility. Therefore, IDPs/IDRs facilitate the regulation of cellular processes by PTMs.

Phosphorylation is one of the most important and well-studied PTMs. It dominates the number of experimentally-identified PTMs by an order of magnitude.<sup>83</sup> It was found that sequence attributes of regions adjacent to phosphorylation sites are very

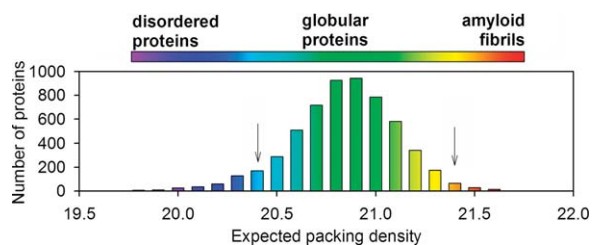
similar to those of IDPs/IDRs, and such a correlation between the structure disorder and the occurrence of phosphorylation has been employed to improve the prediction of protein phosphorylation sites.<sup>84</sup> For *Arabidopsis* plasma membrane proteins, 30% of the phosphorylation sites are located in long intrinsically disordered regions and 28% of sites are located in shorter disordered regions.<sup>85</sup> Phosphorylation of IDPs/IDRs serves as an essential control mechanism in signaling and regulation. A study on the phosphorylation variation during the cell cycle indicated that intrinsically disordered regions tend to contain sites with dynamically varying levels, while ordered regions retain more constant phosphorylation levels.<sup>86</sup> Typical examples are eukaryotic cyclin-Cdk inhibitors Sic1 and p27 whose inhibitory activity, stability, and subcellular localization are regulated by phosphorylation on different sites.<sup>87,88</sup>

Many other PTMs are also associated with IDPs/IDRs. A large-scale analysis of the Swiss-Prot database showed that 17 PTMs keywords (i.e., phosphorylation, amidation, ubiquitination, glycosylation, sulfation, and methylation, etc) are strongly correlated with predicted disorder.<sup>89</sup> In comparison with ordered proteins, IDPs/IDRs contain more ubiquitination sites and this observation was used in developing a predictor program for such sites.<sup>90,91</sup> The preference of ubiquitination in IDPs/IDRs provides a basis for their efficient regulation via rapid degradation.

#### **Advantage 6: Enabling Flexible Linkers**

Acting as entropic chains such as flexible linkers is one of the six broad functional classes of IDPs/IDRs.<sup>18,92</sup> The lack of ordered structures of IDPs/IDRs enables them to act as flexible linkers which are obviously out of reach of ordered proteins.

In multidomain proteins, modular domains are often connected by flexible linkers [Fig. 1(f)].<sup>12</sup> The primary function of these linkers is to restrict the distance and to enable an orientational freedom of the attached domains. IDRs carry out such functions without undergoing a disorder-to-order transition and can be mimicked reasonably well by the behavior of low-complexity polypeptides. The role of IDRs as flexible linkers was nicely demonstrated in an experimental study on the *Escherichia coli* tubulin homologue FtsZ.<sup>93</sup> When the linker in the *E. coli* FtsZ was replaced by those from other bacteria, or even from an unrelated IDP (human  $\alpha$ -adducin), its function in cell division was found to be unaffected if the length of the new linker was similar to that of the original.<sup>93</sup> In the voltage-activated potassium channels Kv, it was found that the intrinsically disordered C-terminal domain behaves as an ideal flexible chain in binding with intracellular scaffold proteins.<sup>94</sup>



**Figure 3.** Schematic distribution of the expected packing density for IDPs/IDRs, ordered proteins and amyloid fibrils (adopted from Ref. 101).

Linking two modular domains or segments via a flexible linker increases the local concentration of one relative to the other as has been widely adopted in biological systems to increase efficiency. This can be used to connect two different functions, for example, the functions of the DNA-binding domain and the activation domain in a DNA transcription factor. In some other cases, this provides a simple way to increase the binding affinity of two linked domains (segments) to a common target.<sup>75,95</sup> Furthermore, the binding affinity can be tuned by varying the length of the linker. For example, the variation of linker length in ratiometric fluorescent sensor proteins was shown to tune their Zn(II) binding affinity from picomolar to femtomolar where the effective local concentration was well described by a worm-like chain model.<sup>96</sup>

#### Advantage 7: Preventing Aggregation

Uncontrolled aggregation is a constant threat for proteins, and proteins have to acquire sequence and structural adaptations to avoid undesired amyloid-like aggregation.<sup>97</sup> The physicochemical properties of the sequences of IDPs/IDRs are generally negatively correlated with those of aggregation-prone sequences.<sup>7</sup> For example, most aggregation-prone sequences have high hydrophobicity and low net charge,<sup>98,99</sup> while IDPs/IDRs have low hydrophobicity and high net charge.<sup>46</sup> So IDP/IDRs sequences are distinct from amyloidogenic peptides,<sup>100</sup> and tend to prevent aggregation. The advantage of IDPs/IDRs relative to ordered proteins in preventing amyloid-like aggregation is apparent in examining the packing density of sequences,<sup>101</sup> that is, the expected number of neighbor contacts per residue within a given distance (similar to the concept of liganacy in chemistry). A residue with high packing density is more sticking to contact with other residues. It was shown that amyloid proteins possess high expected packing density, and IDPs/IDRs possess low values, while ordered proteins possess moderate values locating between those of amyloids and IDPs/IDRs (Fig. 3).<sup>101</sup> It was also shown that ordered proteins contain almost three times as much aggregation nucleating regions as IDPs.<sup>102</sup>

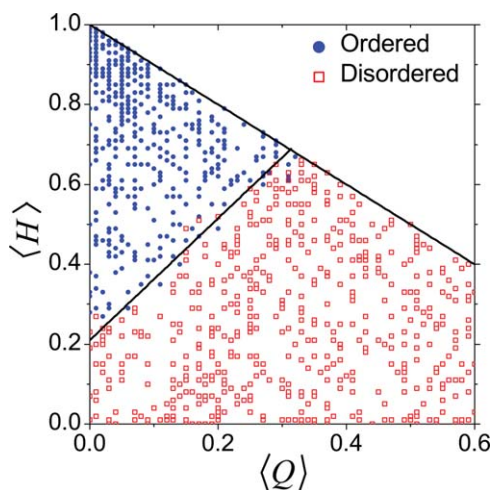
Attaching disordered regions to ordered proteins or aggregation-prone sequences could protect the latter from aggregation [Fig. 1(g)]. Bioinformatics analysis demonstrated that interaction linear motifs are on average embedded in locally disordered regions, where a typical motif contains about six residues and is surrounded by approximately 20 residues that are intrinsically disordered.<sup>103</sup> A simulation with a lattice model showed that small hydrophobic peptides with disordered flanks remain stable under conditions where peptides without flanks tend to aggregate.<sup>104</sup> Such a principle was also adopted in artificial protein fusions to effectively prevent aggregation and achieve high soluble protein expression.<sup>105</sup> Recently, de novo design was conducted to develop a stable monomeric peptide targeting to the tumor necrosis factor (TNF)- $\alpha$  by increasing the propensity of intrinsic disorder.<sup>106</sup>

The capability of IDPs/IDRs to prevent aggregation is also essential for their chaperone function. For example, the *E. coli* protein HdeA experiences a stress-induced order-to-disorder transition to display chaperone activity and binds to other unfolded proteins to prevent their aggregation under the stress conditions.<sup>107</sup>

#### Advantage 8: Providing Resistance to Non-Native Conditions

Ordered proteins are usually vulnerable to unfolding. Their structures and functions may be destroyed by various external stresses such as acid, base, inorganic salt, organic solvent, heating and cooling. In contrast, IDPs are more likely to keep stable with respect to various extreme conditions due to their lack of structures. For example, prothymosin  $\alpha$  can be boiled for a few days without losing its activity.<sup>11</sup> It even displays an atypical “turn out” response where partial structure is induced by heating.<sup>108</sup> A positive correlation between disorder content and the resistance to heat shock was shown in a survey on 11 proteins.<sup>109</sup> A study on a freezing-induced loss-of-function model of globular-disordered functional protein pairs also confirmed that IDPs are more resistant to cold treatment than ordered proteins.<sup>110</sup> Late embryogenesis abundant proteins, which play crucial roles in cellular dehydration tolerance, are mostly IDPs/IDRs.<sup>111</sup> The super stability of IDPs/IDRs was exactly summarized by Uversky as “you cannot break what is already broken” [Fig. 1(h)].<sup>11</sup>

The structural disorder of IDPs/IDRs not only helps them to survive under extreme conditions,<sup>112</sup> but is also beneficial for their resistance to environmental perturbations under physiological conditions.<sup>39</sup> In a real biological system, proteins perform their functions in complicated environments, which can be affected by various factors. To maintain their activity, proteins should not be overly sensitive to



**Figure 4.** Schematic difference of the areas for IDPs and ordered proteins in the CH plot, where  $\langle H \rangle$  and  $\langle Q \rangle$  represent the fraction of hydrophobic and charged residues, respectively (modified from Ref. 7).

perturbations in cellular conditions.<sup>113,114</sup> A coarse-grained simulation revealed that the binding affinity and kinetics of IDPs were less sensitive to the perturbations of temperature and intermolecular interactions than ordered proteins.<sup>39</sup> The origin of such robustness was attributed to the capacity of IDPs to adjust their bound conformations to compensate for various perturbations, which was supported by an analysis of protein complex structures. This ability was termed a “buffer effect,”<sup>39</sup> which, together with the kinetic advantage discussed above, enables IDPs to conduct their functions in cellular signaling and regulation.

#### Advantage 9: Allowing Compatibility With More Available Sequences

The protein sequences span only a very small fraction of the whole available sequence space. For ordered proteins, the number of available sequence is greatly reduced by the need to form a unique folded structure while avoiding insoluble aggregates.<sup>100,115</sup> For IDPs/IDRs, due to the removal of restrictions posed by forming ordered structure, the available sequence space is expected to be much larger than that of ordered proteins [Fig. 1(i)].<sup>1,11</sup> The sequence difference between IDPs/IDRs and ordered proteins can be graphically depicted in the two-dimensional charge-hydrophathy (CH) plot where ordered proteins reside in the region with high hydrophathy and low net charge (Fig. 4). A comparison of the areas for IDPs and ordered proteins in the CH plot showed that the sequence space of extended IDPs is at least fivefold greater than that of compact soluble proteins (including both “molten globule”-like IDPs and well-folded ordered proteins).<sup>11</sup> The difference in available sequence space between IDPs/IDRs and ordered proteins can also be discussed in terms of their sequence

designability,<sup>4,116</sup> which was defined as the number of sequences coding a protein.<sup>117</sup> Without the structural restrictions, IDPs/IDRs may possess more sequence redundancy, resulting in high sequence designability.

Proteins were selected by evolution in nature. The greater number of available sequences may contribute some benefit to IDPs/IDRs in the evolution process. During evolution, many mutations are neutral and do not affect the activity of proteins, while rare mutations are advantageous (positive) and will presumably be preserved because they improve or change the protein function in a desired way. An IDP with a given function may on average have more neutral mutations than an ordered protein due to its sequence redundancy, and it may also have more positive mutations to achieve an advantageous phenotype due to the sequence redundancy of the resulting new IDP. The involvement of IDPs/IDRs in organisms may be closely related to their evolution. For flaviviruses, it was shown that the rapid evolutionary dynamics of structural disorder is a potential driving force for their phenotypic divergence.<sup>118</sup> In mitochondria, it was shown that the IDPs/IDRs contents are markedly different between those descending from a bacterial ancestor and those being added to the mitochondria more recently, suggesting that the frequency of IDPs/IDRs in mitochondria was due to the evolutionary origin rather than the functional difference of the protein.<sup>119</sup>

The sequence redundancy of IDPs/IDRs sheds light on the design of IDPs/IDRs. At present, both IDPs design and IDPs-targeted drug design are in their infancy.<sup>10,120,121</sup> Structure-based design of ordered proteins has been extensively studied, resulting in some well developed strategies.<sup>122</sup> But it is unclear whether these approaches can be applied to design IDPs/IDRs, and rare attempts were conducted in this direction. The lack of a unique folded structure also brings extra hindrance. With an ordered protein, one knows the location and geometry of the active/binding site, which is the basis of conventional rational design. For an IDP, it may be difficult to obtain such knowledge in order to achieve a desired function. On the other hand, IDPs/IDRs possess greater available sequence space than ordered proteins, which would be favorable for the rational IDPs/IDRs design. Recently, Shen *et al.* have successfully designed a small IDP to inhibiting TNF- $\alpha$ , where the coupled-folding-and-binding complex structure was adopted as a basis in the design.<sup>106</sup> It delivered an optimistic hint on the possibility of IDPs/IDRs design.

#### Some Remarks

The subject of this review is the possible advantages of IDPs, and inevitable bias exists in emphasizing them. In fact, no benefit is free. IDPs indeed possess disadvantages (and even harmfulness). IDPs/IDRs



often engage in promiscuous interactions and are associated with various human diseases.<sup>58,59</sup> It was found that 79% of cancer-associated proteins contain predicted IDRs of 30 residues or longer.<sup>123</sup> In the Swiss-Prot disease category, it was shown that 11 disease-related keywords were strongly correlated with IDPs/IDRs while none correlated with ordered proteins.<sup>89</sup> To suppress their harmfulness, the availability of IDPs/IDRs is tightly regulated in cells, which would require extra costs.<sup>57</sup> Another apparent disadvantage of IDPs is their susceptibility to proteolysis.<sup>27,92,124</sup> To avoid unwanted proteolytic degradation as can occur in many biological environments, some mechanisms may exist to protect IDPs/IDRs *in vivo*.<sup>27,124</sup> Binding with partners can provide effective protection for IDPs/IDRs by hiding the protease-sensitive sites. “Nanny” chaperones were also proposed to protect newly synthesized IDPs/IDRs from degradation by specific interactions.<sup>125</sup> In addition, proteases are usually compartmentalized and sequestered in the cell, and their activity is tightly regulated.<sup>27,92</sup>

Advantages of IDPs/IDRs have been widely discussed in the literature. In the above we have summarized and surveyed nine possible advantages. It is not a complete list, for example, Uversky gave a list with 21 advantages in a recent review.<sup>1</sup> It should be noted that many of them lack direct proof. Therefore, they are working concepts and hypotheses rather than established facts. Some advantages seem self-evident, for example, economizing genome/protein resources, overcoming steric restrictions in binding, facilitating posttranslational modifications, and enabling flexible linkers. Some other advantages, however, are complicated and even contradictory, for example, high specificity and compatibility with more available sequences. More work is needed to illuminate these properties of IDPs/IDRs. In this aspect, molecular modeling, especially with coarse-grained models, are powerful tools and extremely useful.<sup>126</sup>

The discussed advantages are not orthogonal. Many are connected in nature. For example, the ability of overcoming steric restrictions in binding would also facilitate posttranslational modifications.

The possible advantages of IDPs/IDRs over ordered proteins differ in magnitude. For example, the interface area per residue in binding can vary substantially, whereas the average binding rate of IDPs is only two to three times faster than that of ordered proteins when the binding rate itself covers a few orders of magnitude. Therefore, it may be difficult for experimental studies to clearly demonstrate some advantages of disorder. Some conflicting studies on IDPs/IDRs, for example, the role in hubs, the evolution rate, and the role of structural disorder in p53-related diseases, may relate to this.

## Conclusion

IDPs/IDRs are distinct from conventional ordered proteins in sequence, structure and function. IDPs/IDRs possess various advantages over ordered proteins, enabling them to perform vital functions in cells and to persist during evolution. Some possible advantages of IDPs/IDRs are not well understood and require both experimental and theoretical approaches to decipher. Such studies would advance our understanding of IDPs/IDRs, and facilitate the development of their application.

## Acknowledgments

The authors gratefully acknowledge Dr. Fan Jin and Huaqing Cao for their insightful discussions. The authors also thank Prof. Dr. Brian Matthews and the anonymous reviewer for their helpful and constructive comments that greatly helped to improve the article.

## References

1. Uversky VN (2013) A decade and a half of protein intrinsic disorder: Biology still waits for physics. *Protein Sci* 22:693–724.
2. Tompa P (2012) Intrinsically disordered proteins: a 10-year recap. *Trends Biochem Sci* 37:509–516.
3. Huang YQ, Liu ZR (2010) Intrinsically disordered proteins: the new sequence-structure-function relations. *Acta Phys Chim Sin* 26:2061–2072.
4. Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CR, Higgs KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z (2001) Intrinsically disordered protein. *J Mol Graph Model* 19:26–59.
5. Wright PE, Dyson HJ (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol* 293:321–331.
6. Uversky VN, Gillespie JR, Fink AL (2000) Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins* 41:415–427.
7. Jin F, Liu ZR (2013) Inherent relationships among different biophysical prediction methods for intrinsically disordered proteins. *Biophys J* 104:488–495.
8. Marsh JA, Forman-Kay JD (2012) Ensemble modeling of protein disordered states: experimental restraint contributions and validation. *Proteins* 80:556–572.
9. Fisher CK, Stultz CM (2011) Constructing ensembles for intrinsically disordered proteins. *Curr Opin Struct Biol* 21:426–431.
10. Jin F, Yu C, Lai LH, Liu ZR (2013) Ligand clouds around protein clouds: a scenario of ligand binding with intrinsically disordered proteins. *PLoS Comput Biol* 9:e1003249.
11. Uversky VN (2013) Unusual biophysics of intrinsically disordered proteins. *Biochim Biophys Acta* 1834:932–951.
12. Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6: 197–208.
13. Knott M, Best RB (2012) A preformed binding interface in the unbound ensemble of an intrinsically disordered protein: evidence from molecular simulations. *PLoS Comput Biol* 8:e1002605.

14. Borg M, Mittag T, Pawson T, Tyers M, Forman-Kay JD, Chan HS (2007) Polyelectrostatic interactions of disordered ligands suggest a physical basis for ultrasensitivity. *Proc Natl Acad Sci USA* 104:9650–9655.
15. Sigalov AB, Kim WM, Saline M, Stem LJ (2008) The intrinsically disordered cytoplasmic domain of the T cell receptor xi chain binds to the Nef protein of simian immunodeficiency virus without a disorder-to-order transition. *Biochemistry* 47:12942–12944.
16. Sigalov AB, Zhuravleva AV, Orekhov VY (2007) Binding of intrinsically disordered proteins is not necessarily accompanied by a structural transition to a folded form. *Biochimie* 89:419–421.
17. Banci L, Bertini I, Cefaro C, Ciofi-Baffoni S, Gajda K, Felli IC, Gallo A, Pavelkova A, Kallergi E, Andreadaki M, Katrakili N, Pozidis C, Tokatlidis K (2013) An intrinsically disordered domain has a dual function coupled to compartment-dependent redox control. *J Mol Biol* 425:594–608.
18. Tompa P (2002) Intrinsically unstructured proteins. *Trends Biochem Sci* 27:527–533.
19. Tompa P, Csermely P (2004) The role of structural disorder in the function of RNA and protein chaperones. *FASEB J* 18:1169–1175.
20. Xie HB, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Uversky VN, 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.
21. Xue B, Dunker AK, Uversky VN (2012) Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. *J Biomol Struct Dyn* 30:137–149.
22. Pushker R, Mooney C, Davey NE, Jacque JM, Shields DC (2013) Marked variability in the extent of protein disorder within and between viral families. *PLoS One* 8:e60724.
23. Oates ME, Romero P, Ishida T, Ghalwash M, Mizianty MJ, Xue B, Dosztanyi Z, Uversky VN, Obradovic Z, Kurgan L, Dunker AK, Gough J (2013) D<sup>2</sup>P<sup>2</sup>: database of disordered protein predictions. *Nucleic Acids Res* 41: D508–D516.
24. Di Domenico T, Walsh I, Tosatto SCE (2013) Analysis and consensus of currently available intrinsic protein disorder annotation sources in the MobiDB database. *BMC Bioinformatics* 14:S3.
25. Monastyrskyy B, Fidelis K, Moulton J, Tramontano A, Kryshtafovych A (2011) Evaluation of disorder predictions in CASP9. *Proteins* 79:107–118.
26. Uversky VN (2011) Intrinsically disordered proteins from A to Z. *Int J Biochem Cell Biol* 43:1090–1103.
27. Dyson HJ (2011) Expanding the proteome: disordered and alternatively folded proteins. *Q Rev Biophys* 44:467–518.
28. Babu MM, van der Lee R, de Groot NS, Gsponer J (2011) Intrinsically disordered proteins: regulation and disease. *Curr Opin Struct Biol* 21:432–440.
29. Mitrea DM, Kriwacki RW (2013) Regulated unfolding of proteins in signaling. *FEBS Lett* 587:1081–1088.
30. Mittag T, Forman-Kay JD (2007) Atomic-level characterization of disordered protein ensembles. *Curr Opin Struct Biol* 17:3–14.
31. Gunasekaran K, Tsai CJ, Kumar S, Zanuy D, Nussinov R (2003) Extended disordered proteins: targeting function with less scaffold. *Trends Biochem Sci* 28:81–85.
32. Meszaros B, Tompa P, Simon I, Dosztanyi Z (2007) Molecular principles of the interactions of disordered proteins. *J Mol Biol* 372:549–561.
33. Romero PR, Zaidi S, Fang YY, Uversky VN, Radivojac P, Oldfield CJ, Cortese MS, Sickmeier M, LeGall T, Obradovic Z, Dunker AK (2006) Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms. *Proc Natl Acad Sci USA* 103:8390–8395.
34. Tompa P, Szasz C, Buday L (2005) Structural disorder throws new light on moonlighting. *Trends Biochem Sci* 30:484–489.
35. Chakrabortee S, Meersman F, Schierle GSK, Bertocini CW, McGee B, Kaminski CF, Tunnacliffe A (2010) Catalytic and chaperone-like functions in an intrinsically disordered protein associated with desiccation tolerance. *Proc Natl Acad Sci USA* 107:16084–16089.
36. Vacic V, Oldfield CJ, Mohan A, Radivojac P, Cortese MS, Uversky VN, Dunker AK (2007) Characterization of molecular recognition features, MoRFs, and their binding partners. *J Proteome Res* 6:2351–2366.
37. Uversky VN (2011) Multitude of binding modes attainable by intrinsically disordered proteins: a portrait gallery of disorder-based complexes. *Chem Soc Rev* 40: 1623–1634.
38. Kiss R, Bozoky Z, Kovacs D, Rona G, Friedrich P, Dvortsak P, Weisemann R, Tompa P, Perczel A (2008) Calcium-induced tripartite binding of intrinsically disordered calpastatin to its cognate enzyme, calpain. *FEBS Lett* 582:2149–2154.
39. Huang YQ, Liu ZR (2010) Smoothing molecular interactions: The "kinetic buffer" effect of intrinsically disordered proteins. *Proteins* 78:3251–3259.
40. Mohan A, Oldfield CJ, Radivojac P, Vacic V, Cortese MS, Dunker AK, Uversky VN (2006) Analysis of molecular recognition features (MoRFs). *J Mol Biol* 362: 1043–1059.
41. Adzhubei AA, Sternberg MJE, Makarov AA (2013) Polypyrrolone-II helix in proteins: structure and function. *J Mol Biol* 425:2100–2132.
42. Tompa P, Fuxreiter M (2008) Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem Sci* 33:2–8.
43. Fuxreiter M (2012) Fuzziness: linking regulation to protein dynamics. *Mol Biosyst* 8:168–177.
44. Mittag T, Kay LE, Forman-Kay JD (2009) Protein dynamics and conformational disorder in molecular recognition. *J Mol Recognit* 23:105–116.
45. Mittag T, Orlicky S, Choy WY, Tang XJ, Lin H, Sicheri F, Kay LE, Tyers M, Forman-Kay JD (2008) Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor. *Proc Natl Acad Sci USA* 105: 17772–17777.
46. Uversky VN (2002) Natively unfolded proteins: A point where biology waits for physics. *Protein Sci* 11:739–756.
47. Zhou HX (2012) Intrinsic disorder: signaling via highly specific but short-lived association. *Trends Biochem Sci* 37:43–48.
48. Huang YQ, Liu ZR (2009) Kinetic advantage of intrinsically disordered proteins in coupled folding-binding process: a critical assessment of the "fly-casting" mechanism. *J Mol Biol* 393:1143–1159.
49. Vreven T, Hwang H, Pierce BG, Weng ZP (2012) Prediction of protein-protein binding free energies. *Protein Sci* 21:396–404.
50. Liu JT, Faeder JR, Camacho CJ (2009) Toward a quantitative theory of intrinsically disordered proteins and their function. *Proc Natl Acad Sci USA* 106:19819–19823.
51. Gilson MK, Zhou HX (2007) Calculation of protein-ligand binding affinities. *Annu Rev Biophys Biomol Struct* 36:21–42.
52. Wang RX, Lai LH, Wang SM (2002) Further development and validation of empirical scoring functions for

- structure-based binding affinity prediction. *J Comput-Aided Mol Des* 16:11–26.
53. De Jonge N, Garcia-Pino A, Buts L, Haesaerts S, Charlier D, Zangger K, Wyns L, De Greve H, Loris R (2009) Rejuvenation of CcdB-poisoned gyrase by an intrinsically disordered protein domain. *Mol Cell* 35: 154–163.
  54. Drobnak I, De Jonge N, Haesaerts S, Vesnaver G, Loris R, Lah J (2013) Energetic basis of uncoupling folding from binding for an intrinsically disordered protein. *J Am Chem Soc* 135:1288–1294.
  55. Huang YQ, Liu ZR (2013) Do intrinsically disordered proteins possess high specificity in protein-protein interactions? *Chem-Eur J* 19:4462–4467.
  56. Wong ETC, Na D, Gsponer J (2013) On the importance of polar interactions for complexes containing intrinsically disordered proteins. *PLoS Comput Biol* 9: e1003192.
  57. Gsponer J, Futschik ME, Teichmann SA, Babu MM (2008) Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. *Science* 322:1365–1368.
  58. Vavouri T, Semple JI, Garcia-Verdugo R, Lehner B (2009) Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity. *Cell* 138:198–208.
  59. Midic U, Oldfield CJ, Dunker AK, Obradovic Z, Uversky VN (2009) Protein disorder in the human diseaseome: unfoldomics of human genetic diseases. *BMC Genomics* 10:12.
  60. Mahmoudabadi G, Rajagopalan K, Getzenberg RH, Hannehalli S, Rangarajan G, Kulkarni P (2013) Intrinsically disordered proteins and conformational noise implications in cancer. *Cell Cycle* 12:26–31.
  61. Schreiber G, Keating AE (2011) Protein binding specificity versus promiscuity. *Curr Opin Struct Biol* 21:50–61.
  62. Bhattacharjee A, Wallin S (2013) Exploring protein-peptide binding specificity through computational peptide screening. *PLoS Comput Biol* 9:e1003277.
  63. Staneva I, Huang YQ, Liu ZR, Wallin S (2012) Binding of two intrinsically disordered peptides to a multi-specific protein: a combined Monte Carlo and molecular dynamics study. *PLoS Comput Biol* 8:e1002682.
  64. Kriwacki RW, Hengst L, Tennant L, Reed SI, Wright PE (1996) Structural studies of p21<sup>Waf1/Cip1/Sdi1</sup> in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proc Natl Acad Sci USA* 93: 11504–11509.
  65. Oldfield CJ, Meng J, Yang JY, Yang MQ, Uversky VN, Dunker AK (2008) Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. *BMC Genomics* 9:S1.
  66. Huang YQ, Liu ZR (2011) Anchoring intrinsically disordered proteins to multiple targets: Lessons from N-terminus of the p53 protein. *Int J Mol Sci* 12:1410–1430.
  67. Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN (2005) Flexible nets - the roles of intrinsic disorder in protein interaction networks. *FEBS J* 272: 5129–5148.
  68. Patil A, Nakamura H (2006) Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. *FEBS Lett* 580:2041–2045.
  69. Singh GP, Ganapathi M, Dash D (2007) Role of intrinsic disorder in transient interactions of hub proteins. *Proteins* 66:761–765.
  70. Bertolazzi P, Bock ME, Guerra C (2013) On the functional and structural characterization of hubs in protein-protein interaction networks. *Biotechnol Adv* 31:274–286.
  71. Bardwell JCA, Jakob U (2012) Conditional disorder in chaperone action. *Trends Biochem Sci* 37:517–525.
  72. Shoemaker BA, Portman JJ, Wolynes PG (2000) Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc Natl Acad Sci USA* 97:8868–8873.
  73. Schreiber G, Haran G, Zhou HX (2009) Fundamental aspects of protein-protein association kinetics. *Chem Rev* 109:839–860.
  74. Huang YQ, Liu ZR (2010) Nonnative interactions in coupled folding and binding processes of intrinsically disordered proteins. *PLoS One* 5:e15375.
  75. Zhou HX, Pang XD, Lu C (2012) Rate constants and mechanisms of intrinsically disordered proteins binding to structured targets. *Phys Chem Chem Phys* 14: 10466–10476.
  76. Hemsath L, Dvorsky R, Fiegen D, Carlier MF, Ahmadian MR (2005) An electrostatic steering mechanism of Cdc42 recognition by Wiskott-Aldrich syndrome proteins. *Mol Cell* 20:313–324.
  77. Levy Y, Onuchic JN, Wolynes PG (2007) Fly-casting in protein-DNA binding: frustration between protein folding and electrostatics facilitates target recognition. *J Am Chem Soc* 129:738–739.
  78. Ganguly D, Otieno S, Waddell B, Iconaru L, Kriwacki RW, Chen JH (2012) Electrostatically accelerated coupled binding and folding of intrinsically disordered proteins. *J Mol Biol* 422:674–684.
  79. Ganguly D, Zhang WH, Chen JH (2013) Electrostatically accelerated encounter and folding for facile recognition of intrinsically disordered proteins. *PLoS Comput Biol* 9:e1003363.
  80. Yin N, Pei JF, Lai LH (2013) A comprehensive analysis of the influence of drug binding kinetics on drug action at molecular and systems levels. *Mol Biosyst* 9:1381–1389.
  81. Papadakos G, Housden NG, Lilly KJ, Kaminska R, Kleantous C (2012) Kinetic basis for the competitive recruitment of TolB by the intrinsically disordered translocation domain of Colicin E9. *J Mol Biol* 418:269–280.
  82. Kalmar L, Acs V, Silhavy D, Tompa P (2012) Long-range interactions in nonsense-mediated mRNA decay are mediated by intrinsically disordered protein regions. *J Mol Biol* 424:125–131.
  83. Khoury GA, Baliban RC, Floudas CA (2011) Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci Rep* 1:90.
  84. Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res* 32:1037–1049.
  85. Nespoulous C, Rofidal V, Sommerer N, Hem S, Rossignol M (2012) Phosphoproteomic analysis reveals major default phosphorylation sites outside long intrinsically disordered regions of Arabidopsis plasma membrane proteins. *Proteome Sci* 10:62.
  86. Tyanova S, Cox J, Olsen J, Mann M, Frishman D (2013) Phosphorylation variation during the cell cycle scales with structural propensities of proteins. *PLoS Comput Biol* 9:e1002842.
  87. Yang XJ, Lau KY, Sevim V, Tang C (2013) Design principles of the yeast G1/S switch. *PLoS Biol* 11:e1001673.
  88. Ou L, Waddell MB, Kriwacki RW (2012) Mechanism of cell cycle entry mediated by the intrinsically disordered protein p27<sup>Kip1</sup>. *ACS Chem Biol* 7:678–682.
  89. Xie HB, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Obradovic Z, Uversky VN (2007)



- Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. *J Proteome Res* 6:1917–1932.
90. Edwards YJK, Lobley AE, Pentony MM, Jones DT (2009) Insights into the regulation of intrinsically disordered proteins in the human proteome by analyzing sequence and gene expression data. *Genome Biol* 10:50.
  91. Radivojac P, Vacic V, Haynes C, Cocklin RR, Mohan A, Heyen JW, Goebel MG, Iakoucheva LM (2010) Identification, analysis, and prediction of protein ubiquitination sites. *Proteins* 78:365–380.
  92. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z (2002) Intrinsic disorder and protein function. *Biochemistry* 41:6573–6582.
  93. Gardner K, Moore DA, Erickson HP (2013) The C-terminal linker of *Escherichia coli* FtsZ functions as an intrinsically disordered peptide. *Mol Microbiol* 89:264–275.
  94. Magidovich E, Orr I, Fass D, Abdu U, Yifrach O (2007) Intrinsic disorder in the C-terminal domain of the Shaker voltage-activated K<sup>+</sup> channel modulates its interaction with scaffold proteins. *Proc Natl Acad Sci USA* 104:13022–13027.
  95. Zhou HX (2001) The affinity-enhancing roles of flexible linkers in two-domain DNA-binding proteins. *Biochemistry* 40:15069–15073.
  96. van Dongen E, Evers TH, Dekkers LM, Meijer EW, Klomp LWJ, Merx M (2007) Variation of linker length in ratiometric fluorescent sensor proteins allows rational tuning of Zn(II) affinity in the picomolar to femtomolar range. *J Am Chem Soc* 129:3494–3495.
  97. Monsellier E, Chiti F (2007) Prevention of amyloid-like aggregation as a driving force of protein evolution. *EMBO Rep* 8:737–742.
  98. Pawar AP, DuBay KF, Zurdo J, Chiti F, Vendruscolo M, Dobson CM (2005) Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. *J Mol Biol* 350:379–392.
  99. Toombs JA, McCarty BR, Ross ED (2010) Compositional determinants of prion formation in yeast. *Mol Cell Biol* 30:319–332.
  100. Naranjo Y, Pons M, Konrat R (2012) Meta-structure correlation in protein space unveils different selection rules for folded and intrinsically disordered proteins. *Mol Biosyst* 8:411–416.
  101. Galzitskaya OV, Garbuzynskiy SO, Lobanov MY (2006) Prediction of amyloidogenic and disordered regions in protein chains. *PLoS Comput Biol* 2:1639–1648.
  102. Linding R, Schymkowitz J, Rousseau F, Diella F, Serrano L (2004) A comparative study of the relationship between protein structure and beta-aggregation in globular and intrinsically disordered proteins. *J Mol Biol* 342:345–353.
  103. Fuxreiter M, Tompa P, Simon I (2007) Local structural disorder imparts plasticity on linear motifs. *Bioinformatics* 23:950–956.
  104. Abeln S, Frenkel D (2008) Disordered flanks prevent peptide aggregation. *PLoS Comput Biol* 4:e1000241.
  105. Santner AA, Croy CH, Vasanwala FH, Uversky VN, Van YYJ, Dunker AK (2012) Sweeping away protein aggregation with entropic bristles: intrinsically disordered protein fusions enhance soluble expression. *Biochemistry* 51:7250–7262.
  106. Shen Q, Zhang CS, Zhang XL, Liu ZR, Lai LH (in press) Computational design of disordered peptides targeting TNF $\alpha$ .
  107. Cumberworth A, Lamour G, Babu MM, Gsponer J (2013) Promiscuity as a functional trait: intrinsically disordered regions as central players of interactomes. *Biochem J* 454:361–369.
  108. Uversky VN (2009) Intrinsically disordered proteins and their environment: effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *Protein J* 28:305–325.
  109. Tsvetkov P, Myers N, Moscovitz O, Sharon M, Prilusky J, Shaul Y (2012) Thermo-resistant intrinsically disordered proteins are efficient 20S proteasome substrates. *Mol Biosyst* 8:368–373.
  110. Tantos A, Friedrich P, Tompa P (2009) Cold stability of intrinsically disordered proteins. *FEBS Lett* 583:465–469.
  111. Hundertmark M, Hincha DK (2008) LEA (Late Embryogenesis Abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics* 9:118.
  112. Zhang M, Lin SX, Song XW, Liu J, Fu Y, Ge X, Fu XM, Chang ZY, Chen PR (2011) A genetically incorporated crosslinker reveals chaperone cooperation in acid resistance. *Nat Chem Biol* 7:671–677.
  113. Lee L, Stollar E, Chang JF, Grossmann JG, O'Brien R, Ladbury J, Carpenter B, Roberts S, Luisi B (2001) Expression of the Oct-1 transcription factor and characterization of its interactions with the Bob1 coactivator. *Biochemistry* 40:6580–6588.
  114. Dyson HJ, Wright PE (2002) Coupling of folding and binding for unstructured proteins. *Curr Opin Struct Biol* 12:54–60.
  115. Lau KF, Dill KA (1990) Theory for protein mutability and biogenesis. *Proc Natl Acad Sci USA* 87:638–642.
  116. Williams PD, Pollock DD, Goldstein RA (2001) Evolution of functionality in lattice proteins. *J Mol Graph Model* 19:150–156.
  117. Li H, Helling R, Tang C, Wingreen N (1996) Emergence of preferred structures in a simple model of protein folding. *Science* 273:666–669.
  118. Ortiz JF, MacDonald ML, Masterson P, Uversky VN, Siltberg-Liberles J (2013) Rapid evolutionary dynamics of structural disorder as a potential driving force for biological divergence in flaviviruses. *Genome Biol Evol* 5:504–513.
  119. Ito M, Tohsato Y, Sugisawa H, Kohara S, Fukuchi S, Nishikawa I, Nishikawa K (2012) Intrinsically disordered proteins in human mitochondria. *Genes Cells* 17:817–825.
  120. Cheng Y, LeGall T, Oldfield CJ, Mueller JP, Van YYJ, Romero P, Cortese MS, Uversky VN, Dunker AK (2006) Rational drug design via intrinsically disordered protein. *Trends Biotechnol* 24:435–442.
  121. Metallo SJ (2010) Intrinsically disordered proteins are potential drug targets. *Curr Opin Chem Biol* 14:481–488.
  122. Zhang CS, Lai LH (2011) Towards structure-based protein drug design. *Biochem Soc Trans* 39:1382–1386.
  123. Iakoucheva LM, Brown CJ, Lawson JD, Obradovic Z, Dunker AK (2002) Intrinsic disorder in cell-signaling and cancer-associated proteins. *J Mol Biol* 323:573–584.
  124. Suskiewicz MJ, Sussman JL, Silman I, Shaul Y (2011) Context-dependent resistance to proteolysis of intrinsically disordered proteins. *Protein Sci* 20:1285–1297.
  125. Tsvetkov P, Reuven N, Shaul Y (2009) The nanny model for IDPs. *Nat Chem Biol* 5:778–781.
  126. Chan HS, Zhang ZQ, Wallin S, Liu ZR (2011) Cooperativity, local-nonlocal coupling, and nonnative interactions: Principles of protein folding from coarse-grained models. *Annu Rev Phys Chem* 62:301–326.