

Intrinsically Disordered Proteins in Chronic Diseases

Prakash Kulkarni^{1,*} and Vladimir N. Uversky^{2,3,*}

¹ Department of Medical Oncology and Therapeutics Research, City of Hope National Medical Center, Duarte, CA 91010, USA

² Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA

³ Laboratory of New methods in Biology, Institute for Biological Instrumentation, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia

* Correspondence: pkulkarni@coh.org (P.K.); vuvversky@health.usf.edu (V.N.U.)

Received: 3 April 2019; Accepted: 3 April 2019; Published: 11 April 2019

It is now increasingly evident that a large fraction of the human proteome comprises proteins that, under physiological conditions, lack fixed, ordered 3D structures as a whole or have segments that are not likely to form a defined 3D structure [1–7]. These proteins and regions are referred to as intrinsically disordered proteins (IDPs) and intrinsically disordered protein regions (IDPRs), respectively. Despite their lack of a stable structure, IDPs/IDPRs are involved in a multitude of crucial biological functions related to regulation, recognition, signaling, and control, where binding to multiple partners and high-specificity/low-affinity interactions plays a crucial role [8–14]. Furthermore, intrinsic disorder is a unique structural feature that enables IDPs/IDPRs to participate in both one-to-many and many-to-one signaling [9,10]. Since they serve as general regulators of various cellular processes, IDPs/IDPRs themselves are tightly controlled [15,16]. However, when misexpressed, misprocessed, mismodified, or dysregulated, IDPs/IDPRs are prone to engage in promiscuous, often unwanted interactions and, thus, are associated with the development of various pathological states. In fact, the majority of human cancer-related proteins [8], as well as many proteins associated with neurodegeneration [17,18], diabetes [19], cardiovascular disease [20], amyloidosis [21], and genetic diseases [22], are either intrinsically disordered or contain long IDPRs. This broad involvement of misbehaving IDPs/IDPRs in human diseases is known as “disorder in disorders” (or D²) concept [23,24].

It is generally believed that IDPs/IDPRs stochastically sample all possible configurations driven by thermal fluctuations. While this may be true for many extended IDPs, which behave as “random” coils, it is likely that due to the variability in interaction energy between different amino acid sequences, some configurations may be strongly preferred while others are forbidden. Furthermore, unlike folded proteins that exhibit a high degree of structural order and undergo collective motions, albeit fairly constrained, IDPs/IDPRs also exhibit some structural and dynamical ordering, being much less constrained in their motions than folded proteins. Thus, the larger structural plasticity of IDPs emphasizes the importance of entropically driven motions. Because of their simplified spatial organization and globally reduced structural content, IDPs/IDPRs are characterized by the exceptional spatiotemporal heterogeneity, where different parts of a protein are ordered (or disordered) to a different degree, and this distribution is constantly changing over time [25,26]. Therefore, IDPs/IDPRs are not homogeneous, but have a very complex mosaic architecture reflecting their highly heterogeneous spatiotemporal structural organization that includes foldons (independently foldable units of a protein), inducible foldons (disordered regions that can fold at least in part due to their interaction with binding partners), non-foldons (non-foldable protein regions), semi-foldons (regions that are always in a semi-folded form), and unfoldons (ordered regions that have to undergo an order-to-disorder transition to become functional) [25,26]. Since these differently (dis)ordered structural elements might have well-defined and specific functions, an IDP/IDPR can be multifunctional, being involved in interaction with, regulation of, and be controlled by a multitude of structurally unrelated partners [27].

Although high-specificity/low-affinity interactions are considered a functional hallmark of IDPs/IDPRs, these proteins can also interact with ultrahigh (picomolar) affinity, but fully retain their structural disorder, long-range flexibility, and highly dynamic character [28]. Such structures, wherein bound IDPs/IDPRs retain structural plasticity, have been defined as fuzzy complexes [29,30]. While a large number of fuzzy complexes have been characterized, many of such interactions are between an IDP/IDPR and a structured protein. There are also examples of direct interactions between IDPs/IDPRs leading to the formation of a fuzzy complex without disorder-to-order transition [31–33]. With recent advances in experimental techniques and better integration with computational simulations, the picture that is emerging suggests that the modest structural ordering and large amplitude collective motions of IDPs/IDPRs enable them to mediate multiple interactions with different partners in the cell.

This Special Issue of *Biomolecules* is dedicated to elucidating how conformational dynamics affects interactions of IDP/IDPR with partner proteins, especially in those IDPs that are implicated in chronic diseases. Because of the unique properties they possess, and also because they are not amenable to classical tools and methods used to study ordered proteins, the IDPs/IDPRs have attracted attention across disciplines and scales. In fact, increasing curiosity in understanding how these fascinating molecules that lack unique structure but carry out the myriad functions resulted in an explosive expansion of the disorder-related literature, and scientists from various backgrounds as diverse as protein biochemists, classical physics, polymer physics, biophysics, theoretical physics, bioengineering, and computational and information science have begun to unravel many mysteries of IDPs/IDPRs. Technological advances, and the availability of more and more sophisticated computational tools, have further contributed to a deeper understanding of how IDPs are able carry out their functions. Therefore, it is not surprising that the various contributions to this Special Issue are from experts in biology, bioinformatics, biochemistry, biophysics, structural biology, and computational physics.

The contributions by Bignon et al. ‘Modulation of measles virus N_{TAIL} interactions through fuzziness and sequence features of disordered binding sites’ [34], and by Kumar and Thompson ‘Role of phosphorylation in the modulation of glucocorticoid receptor’s intrinsically disordered domain’ [35] provide an overview of the disorder-based functionality mostly from a biochemical perspective. Bignon et al. review recent findings on the different interaction mechanisms of the C-terminal domain of the intrinsically disordered nucleoprotein (N) of measles virus (MeV) N_{TAIL}, with two of its known binding partners, namely, the C-terminal X domain of the phosphoprotein of MeV XD (a globular viral protein) and the heat-shock protein 70 (hsp70, a globular cellular protein). Kumar and Thompson, on the other hand, delve in the important role of phosphorylation in gene regulation by glucocorticoid receptor (GR) [34]. The GR N-terminal domain is highly disordered and undergoes disorder-to-order transition following site-specific phosphorylation. And this transition from disorder to order is critical for AF1’s efficient interaction with several coregulatory proteins and subsequent AF1-mediated GR activity [35].

Three articles comprise the structural biology section. Of course, protein NMR remains the work horse for gaining insight into IDP/IDPR (un)structure/function relationships. Levy et al., in their article ‘p53 phosphomimetics preserve transient secondary structure but reduce binding to Mdm2 and MdmX’ [36], examined the disordered p53 transactivation domain (p53TAD) using a combination of biophysical techniques including NMR. The p53TAD contains transient helical structures that are necessary for its binding to the negative regulators, mouse double minute 2 (Mdm2) and MdmX. The second paper in this section, ‘Conserved glycines control disorder and function in the Cold-Regulated Protein, COR15A’ by Sowemimo et al. [37], analyzes the plant cold-regulated protein COR15A from *Arabidopsis* that is important for freeze tolerance. During freezing-induced cellular dehydration, COR15A transitions from a disordered to mostly α -helical structure. The authors tested whether mutations that increase the helicity of COR15A also increase its protective function using a combination of NMR, circular dichroism, and fluorescence spectroscopy. The results of these experiments showed that the mutants with higher content of α -helical structure were characterized by an increased membrane stabilization potential during freezing [37]. Finally, the

paper by Melková et al. 'Structure and functions of microtubule associated proteins tau and MAP2c: similarities and differences' [38] used a combination of NMR and cryo-electron microscopy to examine the propensities of MAPs, tau40 and MAP2, to form transient local structures and long-range contacts in the free state, and conformations adopted by these proteins in complexes with microtubules and filamentous actin, as well as in pathological aggregates. For both molecules, the authors identified transient structural motifs by conformational analysis of the experimental data and observed that many of the short sequence motifs that exhibit transient structural features are linked to functional properties, manifested by specific interactions. Therefore, this detailed structure–function analysis may help in explaining the observed differences between biological activities of tau40 and MAP2c [38].

Single-molecule fluorescence resonance energy transfer (or smFRET) is yet another powerful tool used to study IDPs/IDPRs. It is one of the few approaches that are sensitive to transient populations of substrates within molecular ensembles, and therefore, is ideally suited to discern conformational preferences of IDP/IDPR ensembles. The first paper in the section, 'Spontaneous switching among conformational ensembles in intrinsically disordered proteins', by Choi et al. [39] describes a growing number of proteins that appear intrinsically disordered by biochemical and bioinformatics characterization but switch between restricted regions of conformational space. Such switching between disparate corners of conformational space could bias ligand binding and regulate the volume of IDPs acting as structural or entropic elements. Therefore, mapping the accessible energy landscape and capturing dynamics across a wide range of timescales are essential for recognition of when an IDP/IDPR is acting as such a switch [39]. In the paper entitled 'Extreme fuzziness: Direct interactions between two IDPs', Wang and Wang [40] examined whether two IDPs can interact directly to form a fuzzy complex without disorder-to-order transition. Using a combination of smFRET, NMR, and molecular dynamics (MD) simulation, the authors demonstrate that direct interactions between the two pairs of IDPs, 4.1G-CTD/NuMA and H1/ProTα, do form fuzzy complexes while retaining high conformational dynamics of the isolated proteins, which they name as the extremely fuzzy complexes. Therefore, extreme fuzziness completes the full spectrum of protein–protein interaction modes, suggesting that a more generalized model beyond existing binding mechanisms is required [40].

Using a different biophysical approach, confocal fluorescence microscopy, Kaur et al. [41] in their paper entitled 'Molecular crowding tunes material states of ribonucleoprotein condensates', study how molecular crowding impacts ribonucleoprotein (RNP) liquid condensation using an archetypal disordered RNP, called fused in sarcoma (FUS), as an example. RNP condensation is largely governed by promiscuous attractive inter-chain interactions mediated by low-complexity domains (LCDs). The authors demonstrate that the liquid–liquid coexistence boundary of FUS is lowered by polymer crowders, consistent with an excluded volume model. With increasing bulk crowder concentration, the RNP partition increases and the diffusion rate decreases in the condensed phase. These results reveal that the impact of crowding is largely independent of LCD charge and sequence patterns. These results are consistent with a thermodynamic model of crowder-mediated depletion interaction, which suggests that inter-RNP attraction is enhanced by molecular crowding [41].

Characterizing the structure–function relationship of IDPs/IDPRs is no doubt an essential but daunting task as they can adapt transient structure. Molecular dynamics simulations (MDS) has emerged as a natural complement to various experimental approaches for atomic-level characterizations and mechanistic investigations of this intriguing class of proteins. This SI includes three interesting papers that have exploited this computational technique in conjunction with experimental data to gain new insight into the IDP/IDPR structure/function paradigm. In the first article in this section, 'Electrostatics of Tau protein by molecular dynamics', Castro et al. [42] employed MDS to study the structure of a microtubule associated protein Tau that promotes microtubule assembly and stability. To date, the 3D structure of Tau has not been fully solved, experimentally. This is the first MDS study of full-length Tau in conjunction with a region from the microtubule tubulin with which it interacts. The results bring a new insight into Tau and tubulin

proteins, their characteristics and structure-function relationship, and highlights the fact that Tau is a disordered protein with discrete portions of well-defined secondary structure mostly at the microtubule binding region [42].

The second paper, 'Structural and dynamical order of a disordered protein: Molecular insights into conformational switching of PAGE4 at the systems level', describes another MDS study by Lin et al. [43]. Using prostate-associated gene 4 (PAGE4), an IDP implicated in prostate cancer (PCa), as an example, the authors describe the quantitative reproduction of experimental observations and reveal how structural and dynamic ordering are encoded in the sequence of PAGE4, and how these features can be modulated by different extents of phosphorylation by different kinases. This ordering is reflected in changing populations of certain secondary structural elements, as well as in the regularity of its collective motions, and correlate with the functional interactions of the different conformational ensembles of PAGE4 to give rise to repeated transitions between cellular phenotypes with important physiological consequences [43].

The third paper in this section on MDS and biophysical computation, 'Recent advances in computational protocols addressing intrinsically disordered proteins' by Bhattacharya and Lin [44], argues that to understand the conformational dynamics of IDPs/IDPRs and how their structural ensembles recognize multiple binding partners and small molecule inhibitors, knowledge-based and physics-based sampling techniques, guided by the experimental structural data, can be utilized for the comprehensive and focused *in silico* analyses. However, efficient sampling of the IDP/IDPR conformational ensemble requires traversing the numerous degrees of freedom in the IDP/IDPR energy landscape, as well as force-fields that accurately model the protein and solvent interactions. Therefore, these authors provide an overview of the current state of computational methods for studying IDP/IDPR structure and dynamics and discuss the major challenges in this field.

Finally, the paper 'Functional segments on intrinsically disordered regions in disease-related proteins' by Anbo et al. [45] describes a bioinformatics approach to study IDPs/IDPRs. IDPRs, which are often found in the ordered proteins, are known to play important roles in signaling pathways and transcriptional regulation. Therefore, the authors performed a bioinformatics analysis and found more than a thousand potential functional IDPR segments in disease-related proteins, which are found in cancers, congenital disorders, digestive system diseases, and reproductive system diseases. A detailed analysis of some of these regions showed that the functional segments are located on experimentally verified IDPRs. Since IDPs involved in disease pathology tend to have numerous protein-protein interactors, these data suggest that, by occupying hub positions in the protein-protein interaction networks, IDPs can have huge impacts on human diseases. This study highlights the utility of bioinformatics approaches in conjunction with experimental data to in casting new light on the IDPs/IDPRs [45].

In summary, we trust these articles on the IDPs/IDPRs will not only serve as excellent references, but will also stimulate a flurry of activity toward gaining a deeper insight into these fascinating molecules that constitute a large fraction of the proteomes across all three kingdoms of life and engage in myriad biological activities in ways that seem to challenge conventional wisdom. With new advances in experimental techniques, theoretical concepts, and computational capabilities that permit observations across spatiotemporal scales, it is likely that we may have a better understanding of the IDPs and be able to design strategies to target them for therapeutic purposes.

References

- Peng, Z.; Yan, J.; Fan, X.; Mizianty, M.J.; Xue, B.; Wang, K.; Hu, G.; Uversky, V.N.; Kurgan, L. Exceptionally abundant exceptions: Comprehensive characterization of intrinsic disorder in all domains of life. *Cell Mol. Life Sci* **2015**, *72*, 137–151.
- Colak, R.; Kim, T.; Michaut, M.; Sun, M.; Irimia, M.; Bellay, J.; Myers, C.L.; Blencowe, B.J.; Kim, P.M. Distinct types of disorder in the human proteome: Functional implications for alternative splicing. *PLoS Comput Biol* **2013**, *9*, e1003030.
- Xue, B.; Dunker, A.K.; Uversky, V.N. Orderly order in protein intrinsic disorder distribution: Disorder in 3500 proteomes from viruses and the three domains of life. *J. Biomol Struct Dyn* **2012**, *30*, 137–149.
- Uversky, V.N. The mysterious unfoldome: Structureless, underappreciated, yet vital part of any given proteome. *J. Biomed. Biotechnol* **2010**, *2010*, 568068.
- Shimizu, K.; Toh, H. Interaction between intrinsically disordered proteins frequently occurs in a human protein-protein interaction network. *J. Mol. Biol* **2009**, *392*, 1253–1265.
- Ward, J.J.; Sodhi, J.S.; McGuffin, L.J.; Buxton, B.F.; Jones, D.T. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol* **2004**, *337*, 635–645.
- Dunker, A.K.; Obradovic, Z.; Romero, P.; Garner, E.C.; Brown, C.J. Intrinsic protein disorder in complete genomes. *Genome Inform. Ser. Workshop Genome Inform.* **2000**, *11*, 161–171.
- Iakoucheva, L.M.; Brown, C.J.; Lawson, J.D.; Obradovic, Z.; Dunker, A.K. Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol* **2002**, *323*, 573–584.
- Dunker, A.K.; Cortese, M.S.; Romero, P.; Iakoucheva, L.M.; Uversky, V.N. Flexible nets: The roles of intrinsic disorder in protein interaction networks. *FEBS Journal* **2005**, *272*, 5129–5148.
- Uversky, V.N.; Oldfield, C.J.; Dunker, A.K. Showing your id: Intrinsic disorder as an id for recognition, regulation and cell signaling. *J. Mol. Recognit* **2005**, *18*, 343–384.
- Radivojac, P.; Iakoucheva, L.M.; Oldfield, C.J.; Obradovic, Z.; Uversky, V.N.; Dunker, A.K. Intrinsic disorder and functional proteomics. *Biophys J.* **2007**.
- Vucetic, S.; Xie, H.; Iakoucheva, L.M.; Oldfield, C.J.; Dunker, A.K.; Obradovic, Z.; Uversky, V.N. Functional anthology of intrinsic disorder. 2. Cellular components, domains, technical terms, developmental processes, and coding sequence diversities correlated with long disordered regions. *J. Proteome Res.* **2007**, *6*, 1899–1916.
- Xie, H.; Vucetic, S.; Iakoucheva, L.M.; Oldfield, C.J.; Dunker, A.K.; Obradovic, Z.; Uversky, V.N. Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. *J. Proteome Res.* **2007**, *6*, 1917–1932.
- Xie, H.; Vucetic, S.; Iakoucheva, L.M.; Oldfield, C.J.; Dunker, A.K.; Uversky, V.N.; Obradovic, Z. Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. *J. Proteome Res.* **2007**, *6*, 1882–1898.
- Gsponer, J.; Futschik, M.E.; Teichmann, S.A.; Babu, M.M. Tight regulation of unstructured proteins: From transcript synthesis to protein degradation. *Science* **2008**, *322*, 1365–1368.
- Uversky, V.N.; Dunker, A.K. Biochemistry. Controlled chaos. *Science* **2008**, *322*, 1340–1341.
- Uversky, V.N. The triple power of d(3): Protein intrinsic disorder in degenerative diseases. *Front. Biosci (Landmark Ed.)* **2014**, *19*, 181–258.
- Uversky, V.N. Intrinsic disorder in proteins associated with neurodegenerative diseases. *Front. Biosci (Landmark Ed.)* **2009**, *14*, 5188–5238.
- Du, Z.; Uversky, V.N. A comprehensive survey of the roles of highly disordered proteins in type 2 diabetes. *Int J. Mol. Sci* **2017**, *18*.
- Cheng, Y.; LeGall, T.; Oldfield, C.J.; Dunker, A.K.; Uversky, V.N. Abundance of intrinsic disorder in protein associated with cardiovascular disease. *Biochemistry* **2006**, *45*, 10448–10460.
- Uversky, V.N. Amyloidogenesis of natively unfolded proteins. *Curr Alzheimer Res.* **2008**, *5*, 260–287.
- Midic, U.; Oldfield, C.J.; Dunker, A.K.; Obradovic, Z.; Uversky, V.N. Protein disorder in the human diseasome: Unfoldomics of human genetic diseases. *BMC Genomics* **2009**, *10 Suppl 1*, S12.
- Uversky, V.N.; Oldfield, C.J.; Dunker, A.K. Intrinsically disordered proteins in human diseases: Introducing the d2 concept. *Annu Rev. Biophys* **2008**, *37*, 215–246.
- Uversky, V.N.; Dave, V.; Iakoucheva, L.M.; Malaney, P.; Metallo, S.J.; Pathak, R.R.; Joerger, A.C. Pathological unfoldomics of uncontrolled chaos: Intrinsically disordered proteins and human diseases. *Chem Rev.* **2014**, *114*, 6844–6879.

25. Uversky, V.N. Unusual biophysics of intrinsically disordered proteins. *Biochim Biophys Acta* **2013**, *1834*, 932–951.
26. Uversky, V.N. Paradoxes and wonders of intrinsic disorder: Complexity of simplicity. *Intrinsically Disord Proteins* **2016**, *4*, e1135015.
27. Uversky, V.N. Functional roles of transiently and intrinsically disordered regions within proteins. *FEBS J.* **2015**, *282*, 1182–1189.
28. Borgia, A.; Borgia, M.B.; Bugge, K.; Kissling, V.M.; Heidarsson, P.O.; Fernandes, C.B.; Sottini, A.; Soranno, A.; Buholzer, K.J.; Nettels, D., et al. Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **2018**, *555*, 61–66.
29. Fuxreiter, M. Fold or not to fold upon binding - does it really matter? *Curr Opin Struct Biol* **2018**, *54*, 19–25.
30. Tompa, P.; Fuxreiter, M. Fuzzy complexes: Polymorphism and structural disorder in protein-protein interactions. *Trends Biochem Sci* **2008**, *33*, 2–8.
31. Shen, Q.; Shi, J.; Zeng, D.; Zhao, B.; Li, P.; Hwang, W.; Cho, J.H. Molecular mechanisms of tight binding through fuzzy interactions. *Biophys J.* **2018**, *114*, 1313–1320.
32. Teixeira, J.M.C.; Fuentes, H.; Bielskute, S.; Gairi, M.; Zerko, S.; Kozminski, W.; Pons, M. The two isoforms of lyn display different intramolecular fuzzy complexes with the sh3 domain. *Molecules* **2018**, *23*.
33. Sigalov, A.B.; Zhuravleva, A.V.; Orekhov, V.Y. Binding of intrinsically disordered proteins is not necessarily accompanied by a structural transition to a folded form. *Biochimie* **2007**, *89*, 419–421.
34. Bignon, C.; Troilo, F.; Gianni, S.; Longhi, S. Modulation of measles virus ntail interactions through fuzziness and sequence features of disordered binding sites. *Biomolecules* **2018**, *9*, 8.
35. Kumar, R.; Thompson, E.B. Role of phosphorylation in the modulation of the glucocorticoid receptor's intrinsically disordered domain. *Biomolecules* **2019**, *9*, 77.
36. Levy, R.; Gregory, E.; Borchers, W.; Daughdrill, G. P53 phosphomimetics preserve transient secondary structure but reduce binding to mdm2 and mdmx. *Biomolecules* **2019**, *9*, 83.
37. Sowemimo, O.T.; Knox-Brown, P.; Borchers, W.; Rindfleisch, T.; Thalhammer, A.; Daughdrill, G.W. Conserved glycines control disorder and function in the cold-regulated protein, cor15a. *Biomolecules* **2019**, *9*, 84.
38. Melkova, K.; Zapletal, V.; Narasimhan, S.; Jansen, S.; Hritz, J.; Skrabana, R.; Zweckstetter, M.; Ringkjøbing Jensen, M.; Blackledge, M.; Zidek, L. Structure and functions of microtubule associated proteins tau and map2c: Similarities and differences. *Biomolecules* **2019**, *9*, 105.
39. Choi, U.B.; Sanabria, H.; Smirnova, T.; Bowen, M.E.; Weninger, K.R. Spontaneous switching among conformational ensembles in intrinsically disordered proteins. *Biomolecules* **2019**, *9*, 114.
40. Wang, W.; Wang, D. Extreme fuzziness: Direct interactions between two idps. *Biomolecules* **2019**, *9*, 81.
41. Kaur, T.; Alshareedah, I.; Wang, W.; Ngo, J.; Moosa, M.M.; Banerjee, P.R. Molecular crowding tunes material states of ribonucleoprotein condensates. *Biomolecules* **2019**, *9*, 71.
42. Castro, T.G.; Munteanu, F.D.; Cavaco-Paulo, A. Electrostatics of tau protein by molecular dynamics. *Biomolecules* **2019**, *9*, 116.
43. Lin, X.; Kulkarni, P.; Bocci, F.; Schafer, N.P.; Roy, S.; Tsai, M.Y.; He, Y.; Chen, Y.; Rajagopalan, K.; Mooney, S.M., et al. Structural and dynamical order of a disordered protein: Molecular insights into conformational switching of page4 at the systems level. *Biomolecules* **2019**, *9*, 77.
44. Bhattacharya, S.; Lin, X. Recent Advances in Computational Protocols Addressing Intrinsically Disordered Proteins. *Biomolecules* **2019**, *9*, 146.
45. Anbo, H.; Sato, M.; Okoshi, A.; Fukuchi, S. Functional segments on intrinsically disordered regions in disease-related proteins. *Biomolecules* **2019**, *9*, 88.

