On the Importance of Computational Biology and Bioinformatics to the Origins and Rapid Progression of the Intrinsically Disordered Proteins Field

Chapter · January 2020

DOI: 10.1142/9789811215636_0014

4 authors, including:

Lukasz Kurgan
Virginia Commonwealth University
255 PUBLICATIONS **8,470** CITATIONS

SEE PROFILE

Pacific Symposium on Biocomputing 25:149-158(2020)

On the Importance of Computational Biology and Bioinformatics to the Origins and Rapid Progression of the Intrinsically Disordered Proteins Field

Lukasz Kurgan
Department of Computer Science, Virginia Commonwealth University
Richmond, VA USA
Email: lkurgan@vcu.edu

Predrag Radivojac
Khoury College of Computer Sciences, Northeastern University
Boston MA, USA
Email: predrag@northeastern.edu

Joel L. Sussman
Department of Structural Biology, Weizmann Institute of Science
Rehovot, Israel
Email: Joel.Sussman@weizmann.ac.il

A. Keith Dunker
Center for Computational Biology and Bioinformatics, Indiana University School of Medicine
Indianapolis, IN USA
Email: kedunker@iupui.edu

The concept that an unfolded protein could have biological function was proposed by Linus Pauling in 1940. The first experimental evidence for the existence of unfolded proteins, which are now often called intrinsically disordered proteins and regions (IDPs and IDRs), was published in the 1950s. From the 1950s to the 1990s, many additional articles describing IDPs or IDRs were published, and even several Nobel Prizes have been awarded for research on these proteins. The first computational biology/bioinformatics papers on these proteins were reported between 1998 and 2002, and these computational efforts rapidly increased shortly thereafter. In the late 1980s and mid-1990s, a few IDPs and IDRs involved with signaling and regulation were structurally characterized by nuclear magnet resonance (NMR), and these studies demonstrated a likely involvement of IDPs or IDRs in the associated biological regulation. Altogether, the computational biology/bioinformatics and NMR investigations were synergistic in stimulating the rapid increase in the research on IDPs and IDRs. Herein, we describe this history and recent developments in the IDP research arena. We also highlight a recent collection of papers on IDPs and IDRs that are driven by computational biology and bioinformatics efforts.

Keywords: Intrinsic disorder; bioinformatics; unfolded protein; protein function; IDP; IDR.

1. Introduction

The earliest proposal found so far that intrinsically disordered proteins and regions (IDPs and IDRs) can carry out biological function is based on the 1939 suggestion by Rothen and Landsteiner that antibodies can bind differently structured antigens and do so by starting out in the unfolded state [1]. This suggestion was followed up by Pauling in 1940 with the proposal that the unstructured antibodies fold into different structures when templated by interactions with differently shaped antigens [2]. X-ray diffraction analysis in 2003 of an antibody that binds to two differently shaped antigens was purported to show that Pauling's 1940 hypothesis is confirmed because the same antibody showed structural differences when associated with the differently shaped antigens [3]. Our examination of these crystal structures, however, showed that the backbone structures were essentially identical [4]. Thus, the structural changes were confined to the side chains without the backbone structural differences that are at the core of Pauling's hypothesis. However, numerous other IDPs and IDRs do indeed show backbone structural differences upon binding to differently shaped partners [4, 5].

Starting in the 1950s a number of proteins and protein regions were found to lack structure, including casein (1952), phosvitin (1956), trypsinogen (1976), TMV coat protein (1978), and several other proteins that we briefly reviewed in [4, 6]. In the late 1980s and mid-1990s, reports that relied on NMR structural studies of three regulatory proteins (antennapedia [7], p21waf/Cip/Sdi1 [8], and flgM [9]) showed them to be functional IDPs or to contain functional IDRs. These reports and follow-up publications show that IDPs or IDRs directly participate in protein-protein or protein-DNA interactions *in vitro*, and that these interactions bring about significant biological regulation *in vivo*.

Several Nobel prizes have been awarded for studies on proteins that are IDPs or that contain IDRs, including studies on the following: gene regulation (1965, Physiology or Medicine), genetic control of early embryonic development (1995, Physiology or Medicine), prion disease (1997, Physiology or Medicine), transcription (2006, Chemistry), induced pluripotent stem cells (2012, Physiology or Medicine), circadian rhythms (2017, Physiology or Medicine), and phage display (2018, Chemistry). About ten additional Nobel Prizes may also involve discoveries based on studies of IDPs or IDRs. Work is in progress to illuminate the IDPs or IDRs underlying all of these Nobel Prizes, and, where possible, to present the underlying IDP- or IDR-associated molecular mechanisms.

Computational studies performed in 1990s showed that IDPs and IDRs have amino acid compositions that differ significantly from the compositions of structured proteins, and that these differences can account for the lack of structure formation by IDPs and IDRs [10]. Specifically, compared to structured proteins, the IDPs and IDRs are depleted in aromatic residues, are significantly reduced in other hydrophobic residues, are measurably enriched in polar residues, and contain increased amounts of proline. In addition, IDPs and IDRs often have a significantly increased net charge and reduced sequence complexity [11]. These characteristics inhibit protein folding *in vitro*.

2. Computational prediction of IDPs and IDRs and their functions

Compositional differences between structured proteins and IDPs or IDRs were used to construct computational predictors of protein disorder [11-13]. A number of more sophisticated methods that combine the compositional information with various types of sequence-derived information, such as sequence alignment profiles and conservation, were published in recent years [14-22]. Recent empirical analyses demonstrate that some of these methods generate highly accurate predictions of IDPs and IDRs [23-26]. In one of the early computational studies, one such predictor [12] was applied to protein sequence databases revealing that IDRs are unexpectedly common [27]. Soon after another predictor [11] was applied to a collection of whole proteomes from archaea, bacteria, and eukaryotes, showing that IDPs and IDRs are common across all three domains of life, and are especially enriched in eukaryotes [28]. Several follow-up studies that relied on more accurate predictors and that investigated much larger datasets composed of hundreds or even thousands of proteomes have confirmed these results [29-34].

NMR studies and a selection of informatics investigations on protein structure and function were cited in a seminal review on IDPs, a review that called for the reassessment of the protein structure-function paradigm [35]. The search term given in the legend to Figure 1 along with manual review yield more than 900 reviews on IDPs or IDRs. A perusal of the titles of these reviews gives the impression that IDPs and IDRs participate in very wide ranging biological activities. More than 25 distinct, mostly signaling-related, functions were identified as being directly associated with a large collection of experimentally verified IDPs or IDRs [36]. DisProt (www.DisProt.org), a manually curated database of IDPs and IDRs, has significantly enlarged this number to over 40 IDP- and IDR-associated functions [37-39]. Bioinformatics approaches has helped to identify on the order of 200 functions that are possibly associated with IDPs or IDRs [30, 40-45], but inherent uncertainties in these studies mean that these bioinformatics-based associations need confirmation by means of laboratory experiments. Moreover, over a dozen of computational tools that predict functions of IDPs and IDRs were developed in recent years [19, 46, 47]. They predict molecular partners that interact with IDRs [48-52], disordered linker regions [53] and moonlighting regions [54].

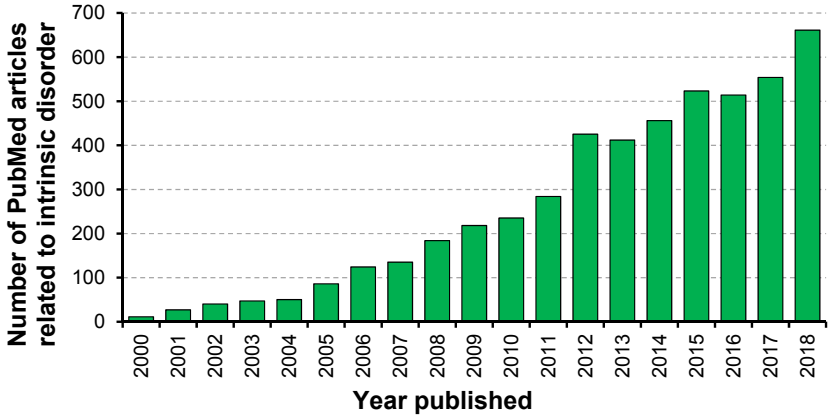

Fig. 1. Annual counts of IDP and IDR papers between 2000 and 2018. PubMed was searched using "(intrinsically OR inherently OR natively) AND (disordered OR unstructured OR unfolded) AND protein" query, which returned 4,986 hits.

3. Popularization of research on IDPs and IDRs

The NMR studies in the late 1980's to mid-1990s and the bioinformatics studies in the late 1990's to early 2000s were soon followed by an increase in the number of publications on IDPs or IDRs (Figure 1). Many events that were organized in the last two decades have contributed to the rapid growth of research in this area including:

- Sessions on IDPs in Pacific Symposium on Biocomputing (PSB) in January 1999, 2001, 2012, and 2020.
- Sessions on IDPs in the Albany Conversation, a structural biology meeting, in June 2003, 2009, and 2019.
- Half-day symposia hosted by the IDP Subgroup of the Biophysical Society in late February or early March each year since 2007.
- An EMBO Meeting on IDPs in Budapest in May 2007. This was the first multiday conference devoted to these proteins.
- IDP Gordon Research Conferences during even numbered years in late June or early July since 2010.
- A multi-day meeting on IDPs in Barcelona in October 2010.
- A meeting on high resolution methods for the investigation of IDPs in Riva del Sole, Castiglione della Pescaia, Grosseto, Italy, in September 2014. This meeting served as the capstone for a training grant that supported more than 15 young researchers during their Ph.D. and postdoctoral research.

Another important stimulant have been the special issues of journals and edited books devoted to IDPs, which include:

- Two IDP and IDR-related issues of *Methods in Molecular Biology* (2012) [55]
- An IDP and IDR issue of *Chemical Reviews* (2014) [56]
- Special issue on "In-Silico Prediction and Characterization of Intrinsic Disorder in Proteins" that was published in *International Journal of Molecular Sciences* (2015)
- Edited book entitled "Intrinsically Disordered Proteins Studied by NMR Spectroscopy" published in the *Advances in Experimental Medicine and Biology* book series (2015) [57]
- Special issue on "Intrinsically Disordered Proteins in the Norm and Pathology: In-Silico Perspective" that appeared in the *International Journal of Molecular Sciences* (2017)
- Special issue on "Intrinsically Disordered Proteins: Structure, Function and Therapeutics" in the *Journal of Molecular Biology* (2018) [58]
- Special issue on "Intrinsically Disordered Proteins in Chronic Diseases" published by the *Biomolecules* journal (2019) [59]

Moreover, the *Cell Communication and Signaling* journal, the *Entropy* journal (https://www.mdpi.com/journal/entropy/special_issues/Information_Disordered_Networks), and the *Biomolecules* journal (https://www.mdpi.com/journal/biomolecules/special_issues/Disorder-Based_Functionality) have pending special issues devoted to IDPs and IDRs. Besides the above-mentioned journal issues, the research on IDPs and IDRs was published in a wide range of journals.

The nearly 5000 articles considered in Figure 1 were published in 648 different journals. Figure 2 provides a sorted list of the top 40 journals that published the largest number of articles related to the disordered proteins and regions. This list includes journals in numerous fields of research that cover both experimental and computational area, attesting to the broad reach of research on the IDPs and IDRs.

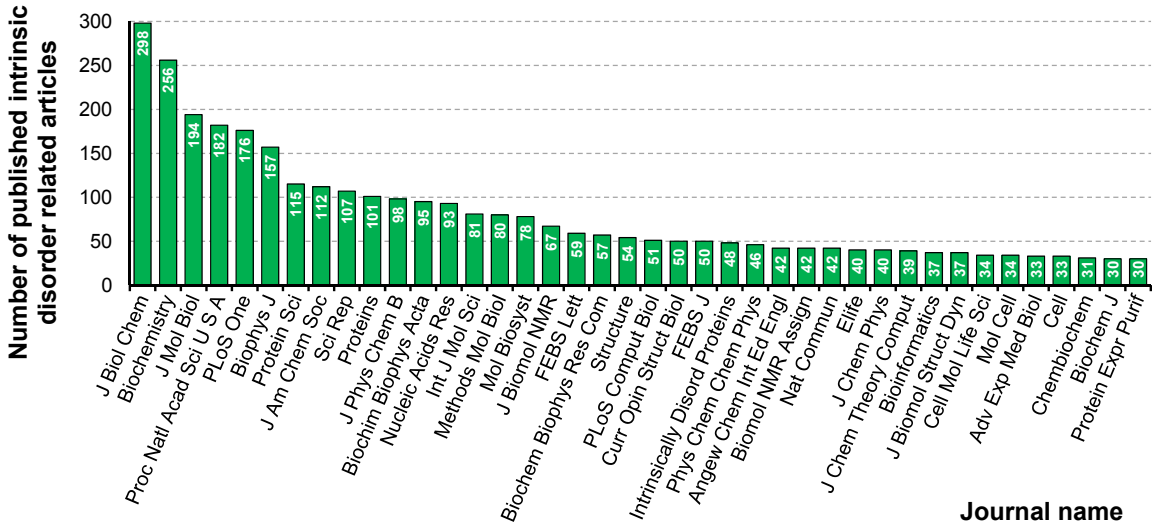

Figure 2. List of the top 40 journals that published the largest number of the disorder related articles. These data are based on the PubMed query listed in Figure 1.

4. A Collection of Recent Papers on IDPs and IDRs

The "Intrinsically Disordered Proteins and Their Functions" session at the 2020 PSB meeting includes five articles that focus on the computational biology and bioinformatics efforts.

4.1. *"Many-to-one binding by intrinsically disordered regions" by W.L. Alterovitz, B. Xue, F. Huang, P. Romero, A. Kloczkowski, V.N. Uversky, and A.K. Dunker*

Previously this group studied individual IDRs that bind to many non-identical protein partners (called one-to-many binding). Here the focus is on different IDRs that bind to the same protein partner (called many-to-one binding [5]). Often the different IDRs simply bind to different sites on the surface of the partner protein (independent binding). In other cases, the different IDRs have binding sites that overlap, either being similar where the binding surfaces are highly overlapping over their entireties, or being intersecting in which the binding surfaces diverge significantly. These patterns suggest that disorder flexibility increases its potential to find binding sites on the surfaces of structured proteins.

4.2. *"Disordered function conjunction: On the in-silico function annotation of intrinsically disordered regions" by S. Ghadermarzi, A. Katuwawala, C.J. Oldfield, A. Barik, and L. Kurgan*

This article investigates feasibility of a computational prediction of functions of IDRs using the results produced by the current residue-level predictors of disorder functions, including ANCHOR [52, 60], , DisoRDPbind [50, 61], fMoRFpred [48], DFLpred [53] and DMRpred [54]. The authors provide empirical evidence that the IDR-level prediction could be done accurately and that information extracted directly from the sequences of these IDRs can be used to improve the predictive performance. This observation opens a new research area that concerns the development of computational methods that predict functions for experimentally annotated IDRs.

4.3. *"De novo ensemble modeling suggests that AP2-binding to disordered regions can increase steric volume of Epsin but not Eps15" by N.S. Jagannathan, C.W.V. Hogue and L. Tucker-Kellogg*

This article describes a theoretical analysis of the radius of gyration and end to end distances of an IDR of AP2 in the presence and absence of two of receptors, Epsin and Eps15. Both receptors are thought to play a part in Cathrin-mediated endocytosis, and the authors test whether the interaction between the IDR and the receptor may actively contribute to membrane bending by inducing an extended conformation of AP2. The FoldTraj program was used to sample millions of conformations of the AP2 disordered domain based on steric considerations, superimpose 1-4 copies of the Epsins, and conclude that AP2 binding leads to increased radius of gyration and end-to-end distance of Epsin, but not Eps15.

4.4. *"Modulation of p53 transactivation domain conformations by ligand binding and cancer-associated mutations" by X. Liu and J. Chen*

This contribution centers on computational characterization of the effects of four cancer-associated mutations on the conformational ensemble of the disordered N-terminal transactivation domain of the p53 protein. The authors conclude that the resulting induced conformational collapse of this IDP could be considered as a general mechanism for shielding functional sites and inhibiting recognition of their targets. More generally, this work provides foundations to develop novel computational tools for drug design approaches targeting regulatory IDPs.

4.5. *"Exploring relationships between the density of charge tracks within disordered regions and phase separation" by R. Somjee, D.U. Mitrae, and R.W. Kriwacki*

Several publications point out that IDPs are often involved in biomolecular condensates that arise via liquid-liquid phase separations. This publication explores the contributions of charged tracks within IDRs. An algorithm, called ABT density, quantifies the density of charged tracks and reveals that higher densities of charged tracks cluster into distinct biomolecular condensates. These data suggest that multivalent electrostatic interactions within IDRs help to organize certain biomolecular condensates via phase separation.

References

1. Rothen, A. and K. Landsteiner, *Adsorption of Antibodies by Egg Albumin Films.* Science, 1939. **90**(2325): p. 65-6.
2. Pauling, L., *A theory of the structure and process of formation of antibodies.* J Am Chem Soc, 1940. **62**: p. 2643-2657.
3. James, L.C. and D.S. Tawfik, *Conformational diversity and protein evolution--a 60-year-old hypothesis revisited.* Trends Biochem Sci, 2003. **28**(7): p. 361-8.
4. Dunker, A.K. and C.J. Oldfield, *Back to the Future: Nuclear Magnetic Resonance and Bioinformatics Studies on Intrinsically Disordered Proteins.* Adv Exp Med Biol, 2015. **870**: p. 1-34.
5. Hsu, W.L., C.J. Oldfield, B. Xue, J. Meng, F. Huang, P. Romero, V.N. Uversky and A.K. Dunker, *Exploring the binding diversity of intrinsically disordered proteins involved in one-to-many binding.* Protein Sci, 2013. **22**(3): p. 258-73.
6. Dunker, A.K., J.D. Lawson, C.J. Brown, R.M. Williams, P. Romero, J.S. Oh, C.J. Oldfield, A.M. Campen, C.M. Ratliff, K.W. Hipps, J. Ausio, M.S. Nissen, R. Reeves, C. Kang, C.R. Kissinger, R.W. Bailey, M.D. Griswold, W. Chiu, E.C. Garner and Z. Obradovic, *Intrinsically disordered protein.* J Mol Graph Model, 2001. **19**(1): p. 26-59.
7. Qian, Y.Q., M. Billeter, G. Otting, M. Muller, W.J. Gehring and K. Wuthrich, *The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors.* Cell, 1989. **59**(3): p. 573-80.
8. Kriwacki, R.W., L. Hengst, L. Tennant, S.I. Reed and P.E. Wright, *Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity.* Proc Natl Acad Sci U S A, 1996. **93**(21): p. 11504-11509.
9. Daughdrill, G.W., M.S. Chadsey, J.E. Karlinsey, K.T. Hughes and F.W. Dahlquist, *The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, sigma 28.* Nat Struct Biol, 1997. **4**(4): p. 285-91.
10. Xie, Q., G.E. Arnold, P. Romero, Z. Obradovic, E. Garner and A.K. Dunker, *The Sequence Attribute Method for Determining Relationships Between Sequence and Protein Disorder.* Genome Inform Ser Workshop Genome Inform, 1998. **9**: p. 193-200.
11. Romero, P., Z. Obradovic, X.H. Li, E.C. Garner, C.J. Brown and A.K. Dunker, *Sequence complexity of disordered protein.* Proteins-Structure Function and Genetics, 2001. **42**(1): p. 38-48.
12. Romero, P., Z. Obradovic, C. Kissinger, J.E. Villafranca and A.K. Dunker, *Identifying disordered regions in proteins from amino acid sequence.* 1997 Ieee International Conference on Neural Networks, Vols 1-4, 1997: p. 90-95.
13. Uversky, V.N., J.R. Gillespie and A.L. Fink, *Why are "natively unfolded" proteins unstructured under physiologic conditions?* Proteins, 2000. **41**(3): p. 415-427.
14. He, B., K. Wang, Y. Liu, B. Xue, V.N. Uversky and A.K. Dunker, *Predicting intrinsic disorder in proteins: an overview.* Cell Res, 2009. **19**(8): p. 929-949.
15. Dosztányi, Z., B. Mészáros and I. Simon, *Bioinformatical approaches to characterize intrinsically disordered/unstructured proteins.* Briefings in Bioinformatics, 2010. **11**(2): p. 225-243.
16. Pentony, M., J. Ward and D. Jones, *Computational Resources for the Prediction and Analysis of Native Disorder in Proteins*, in *Proteome Bioinformatics*, S.J. Hubbard and A.R. Jones, Editors. 2010, Humana Press. p. 369-393.

17. Atkins, J., S. Boateng, T. Sorensen and L. McGuffin, *Disorder Prediction Methods, Their Applicability to Different Protein Targets and Their Usefulness for Guiding Experimental Studies.* International Journal of Molecular Sciences, 2015. **16**(8): p. 19040.
18. Deng, X., J. Eickholt and J. Cheng, *A comprehensive overview of computational protein disorder prediction methods.* Mol Biosyst, 2012. **8**(1): p. 114-21.
19. Meng, F., V.N. Uversky and L. Kurgan, *Comprehensive review of methods for prediction of intrinsic disorder and its molecular functions.* Cell Mol Life Sci, 2017. **74**(17): p. 3069-3090.
20. Meng, F., V. Uversky and L. Kurgan, *Computational Prediction of Intrinsic Disorder in Proteins.* Curr Protoc Protein Sci, 2017. **88**: p. 2 16 1-2 16 14.
21. Oldfield, C.J., V.N. Uversky, A.K. Dunker and L. Kurgan, *Chapter 1 - Introduction to intrinsically disordered proteins and regions*, in *Intrinsically Disordered Proteins*, N. Salvi, Editor. 2019, Academic Press. p. 1-34.
22. Prilusky, J., C.E. Felder, T. Zeev-Ben-Mordehai, E.H. Rydberg, O. Man, J.S. Beckmann, I. Silman and J.L. Sussman, *FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded.* Bioinformatics, 2005. **21**(16): p. 3435-8.
23. Necci, M., D. Piovesan, Z. Dosztanyi, P. Tompa and S.C.E. Tosatto, *A comprehensive assessment of long intrinsic protein disorder from the DisProt database.* Bioinformatics, 2017.
24. Walsh, I., M. Giollo, T. Di Domenico, C. Ferrari, O. Zimmermann and S.C. Tosatto, *Comprehensive large-scale assessment of intrinsic protein disorder.* Bioinformatics, 2015. **31**(2): p. 201-8.
25. Monastyrskyy, B., A. Kryshtafovych, J. Moult, A. Tramontano and K. Fidelis, *Assessment of protein disorder region predictions in CASP10.* Proteins, 2014. **82 Suppl 2**: p. 127-37.
26. Peng, Z.L. and L. Kurgan, *Comprehensive comparative assessment of in-silico predictors of disordered regions.* Curr Protein Pept Sci, 2012. **13**(1): p. 6-18.
27. Romero, P., Z. Obradovic, C.R. Kissinger, J.E. Villafranca, E. Garner, S. Guilliot and A.K. Dunker, *Thousands of proteins likely to have long disordered regions.* Pac Symp Biocomput, 1998: p. 437-48.
28. Dunker, A.K., Z. Obradovic, P. Romero, E.C. Garner and C.J. Brown, *Intrinsic protein disorder in complete genomes.* Genome Inform Ser Workshop Genome Inform, 2000. **11**: p. 161-71.
29. Xue, B., A.K. Dunker and V.N. Uversky, *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**(2): p. 137-49.
30. Peng, Z., J. Yan, X. Fan, M.J. Mizianty, B. Xue, K. Wang, G. Hu, V.N. Uversky and L. Kurgan, *Exceptionally abundant exceptions: comprehensive characterization of intrinsic disorder in all domains of life.* Cell Mol Life Sci, 2015. **72**(1): p. 137-51.
31. Hu, G., K. Wang, J. Song, V.N. Uversky and L. Kurgan, *Taxonomic Landscape of the Dark Proteomes: Whole-Proteome Scale Interplay Between Structural Darkness, Intrinsic Disorder, and Crystallization Propensity.* Proteomics, 2018: p. e1800243.
32. Peng, Z., M.J. Mizianty and L. Kurgan, *Genome-scale prediction of proteins with long intrinsically disordered regions.* Proteins, 2014. **82**(1): p. 145-58.

33. Yan, J., M.J. Mizianty, P.L. Filipow, V.N. Uversky and L. Kurgan, *RAPID: fast and accurate sequence-based prediction of intrinsic disorder content on proteomic scale.* Biochim Biophys Acta, 2013. **1834**(8): p. 1671-80.
34. Oates, M.E., P. Romero, T. Ishida, M. Ghalwash, M.J. Mizianty, B. Xue, Z. Dosztanyi, V.N. Uversky, Z. Obradovic, L. Kurgan, A.K. Dunker and J. Gough, *D(2)P(2): database of disordered protein predictions.* Nucleic Acids Res, 2013. **41**(Database issue): p. D508-16.
35. Wright, P.E. and H.J. Dyson, *Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm.* J Mol Biol, 1999. **293**(2): p. 321-331.
36. Dunker, A.K., C.J. Brown, J.D. Lawson, L.M. Iakoucheva and Z. Obradovic, *Intrinsic disorder and protein function.* Biochemistry, 2002. **41**(21): p. 6573-82.
37. Piovesan, D., F. Tabaro, I. Micetic, M. Necci, F. Quaglia, C.J. Oldfield, M.C. Aspromonte, N.E. Davey, R. Davidovic, Z. Dosztanyi, A. Elofsson, A. Gasparini, A. Hatos, A.V. Kajava, L. Kalmar, E. Leonardi, T. Lazar, S. Macedo-Ribeiro, M. Macossay-Castillo, A. Meszaros, G. Minervini, N. Murvai, J. Pujols, D.B. Roche, E. Salladini, E. Schad, A. Schramm, B. Szabo, A. Tantos, F. Tonello, K.D. Tsirigos, N. Veljkovic, S. Ventura, W. Vranken, P. Warholm, V.N. Uversky, A.K. Dunker, S. Longhi, P. Tompa and S.C. Tosatto, *DisProt 7.0: a major update of the database of disordered proteins.* Nucleic Acids Res, 2016. **D1**: p. D219-D227.
38. Sickmeier, M., J.A. Hamilton, T. LeGall, V. Vacic, M.S. Cortese, A. Tantos, B. Szabo, P. Tompa, J. Chen, V.N. Uversky, Z. Obradovic and A.K. Dunker, *DisProt: the Database of Disordered Proteins.* Nucleic Acids Res, 2007. **35**(Database issue): p. D786-93.
39. Vucetic, S., Z. Obradovic, V. Vacic, P. Radivojac, K. Peng, L.M. Iakoucheva, M.S. Cortese, J.D. Lawson, C.J. Brown, J.G. Sikes, C.D. Newton and A.K. Dunker, *DisProt: a database of protein disorder.* Bioinformatics, 2005. **21**(1): p. 137-40.
40. Ward, J.J., J.S. Sodhi, L.J. McGuffin, B.F. Buxton and D.T. Jones, *Prediction and functional analysis of native disorder in proteins from the three kingdoms of life.* J Mol Biol, 2004. **337**(3): p. 635-45.
41. Xie, H., S. Vucetic, L.M. Iakoucheva, C.J. Oldfield, A.K. Dunker, V.N. Uversky and Z. Obradovic, *Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions.* J. Proteome Res., 2007. **6**(5): p. 1882-98.
42. Xie, H., S. Vucetic, L.M. Iakoucheva, C.J. Oldfield, A.K. Dunker, Z. Obradovic and V.N. Uversky, *Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins.* J Proteome Res, 2007. **6**(5): p. 1917-32.
43. Vucetic, S., H. Xie, L.M. Iakoucheva, C.J. Oldfield, A.K. Dunker, Z. Obradovic and V.N. Uversky, *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**(5): p. 1899-916.
44. Hu, G., Z. Wu, V.N. Uversky and L. Kurgan, *Functional Analysis of Human Hub Proteins and Their Interactors Involved in the Intrinsic Disorder-Enriched Interactions.* Int J Mol Sci, 2017. **18**(12).
45. Meng, F., I. Na, L. Kurgan and V.N. Uversky, *Compartmentalization and Functionality of Nuclear Disorder: Intrinsic Disorder and Protein-Protein Interactions in Intra-Nuclear Compartments.* Int J Mol Sci, 2016. **17**(1).

46. Katuwawala, A., S. Ghadermarzi and L. Kurgan, *Chapter Nine - Computational prediction of functions of intrinsically disordered regions*, in *Progress in Molecular Biology and Translational Science*, V.N. Uversky, Editor. 2019, Academic Press. p. 341-369.
47. Katuwawala, A., Z. Peng, J. Yang and L. Kurgan, *Computational Prediction of MoRFs, Short Disorder-to-order Transitioning Protein Binding Regions.* Comput Struct Biotechnol J, 2019. **17**: p. 454-462.
48. Yan, J., A.K. Dunker, V.N. Uversky and L. Kurgan, *Molecular recognition features (MoRFs) in three domains of life.* Mol Biosyst, 2016. **12**(3): p. 697-710.
49. Disfani, F.M., W.L. Hsu, M.J. Mizianty, C.J. Oldfield, B. Xue, A.K. Dunker, V.N. Uversky and L. Kurgan, *MoRFpred, a computational tool for sequence-based prediction and characterization of short disorder-to-order transitioning binding regions in proteins.* Bioinformatics, 2012. **28**(12): p. i75-83.
50. Peng, Z. and L. Kurgan, *High-throughput prediction of RNA, DNA and protein binding regions mediated by intrinsic disorder.* Nucleic Acids Res, 2015. **43**(18): p. e121.
51. Malhis, N., M. Jacobson and J. Gsponer, *MoRFchibi SYSTEM: software tools for the identification of MoRFs in protein sequences.* Nucleic Acids Res, 2016.
52. Dosztanyi, Z., B. Meszaros and I. Simon, *ANCHOR: web server for predicting protein binding regions in disordered proteins.* Bioinformatics, 2009. **25**(20): p. 2745-6.
53. Meng, F. and L. Kurgan, *DFLpred: High-throughput prediction of disordered flexible linker regions in protein sequences.* Bioinformatics, 2016. **32**(12): p. i341-i350.
54. Meng, F. and L. Kurgan, *High-throughput prediction of disordered moonlighting regions in protein sequences.* Proteins, 2018.
55. Uversky, V.N. and A.K. Dunker, *Intrinsically disordered protein analysis : methods and experimental tools*. Methods in molecular biology,. 2012, New York: Humana Press.
56. Uversky, V.N., *Introduction to intrinsically disordered proteins (IDPs).* Chem Rev, 2014. **114**(13): p. 6557-60.
57. Felli, I.C. and R. Pierattelli, *Intrinsically disordered proteins studied by NMR spectroscopy*. Advances in experimental medicine and biology,. 2015, Cham: Springer. xiii, 421 pages.
58. Chen, J. and R.W. Kriwacki, *Intrinsically Disordered Proteins: Structure, Function and Therapeutics.* J Mol Biol, 2018. **430**(16): p. 2275-2277.
59. Kulkarni, P. and V.N. Uversky, *Intrinsically Disordered Proteins in Chronic Diseases.* Biomolecules, 2019. **9**(4).
60. Meszaros, B., G. Erdos and Z. Dosztanyi, *IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding.* Nucleic Acids Res, 2018. **46**(W1): p. W329-W337.
61. Peng, Z., C. Wang, V.N. Uversky and L. Kurgan, *Prediction of Disordered RNA, DNA, and Protein Binding Regions Using DisoRDPbind.* Methods Mol Biol, 2017. **1484**: p. 187-203.