Untapped potential of disordered proteins in current druggable human proteome

Gang Hu[1#], Zhonghua Wu[1#], Kui Wang[1], Vladimir N. Uversky*[2,3,4,5], and Lukasz Kurgan*[6]

[1] School of Mathematical Sciences and LPMC, Nankai University, Tianjin, People's Republic of China
[2] Department of Biology, Faculty of Science, King Abdulaziz University, Jeddah 21589, Kingdom of Saudi Arabia
[3] Department of Molecular Medicine and USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, Florida, USA
[4] Institute for Biological Instrumentation, Russian Academy of Sciences, Pushchino, Moscow Region, Russian Federation
[5] Laboratory of Structural Dynamics, Stability and Folding of Proteins, Institute of Cytology, Russian Academy of Sciences, St. Petersburg, Russian Federation
[6] Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta, Canada

Joint first authors
* Corresponding authors
VNU: Department of Molecular Medicine, University of South Florida, 12901 Bruce B. Downs Blvd. MDC07, Tampa, Florida 33612, USA, E-mail: vuversky@health.usf.edu
LK: Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta T6G 2V4, Canada; Phone: (780) 492-5488; Fax: (780) 492-1811; E-mail: lkurgan@ece.ualberta.ca

Abstract

Current efforts in design and characterization of drugs often rely on the structure of their protein targets. However, a large fraction of proteins lack unique 3-D structures and exist as highly dynamic structural ensembles. These intrinsically disordered proteins are involved in pathogenesis of various human diseases and are highly abundant in eukaryotes. Based on a comprehensive analysis of the current druggable human proteome covering 12 drug classes and 18 major classes of drug targets we show a significant bias toward high structural coverage and low abundance of intrinsic disorder. We review reasons for this bias including widespread use of the structural information in various stages of drug development and characterization process and difficulty with attaining structures for the intrinsically disordered proteins. We also discuss future of intrinsically disordered proteins as drug targets. Given the overall high disorder content of the human proteome and current bias of the druggable human proteome toward structural proteins, it is inevitable that disordered proteins will have to raise up on the list of prospective drug targets. The protein disorder-assisted drug design can draw from current rational drug design techniques and would also need novel approaches that no longer rely on a unique protein structure.

Keywords: disorder in disorders; druggable human proteome; intrinsic disorder; intrinsically disordered proteins; human drug target; rational drug design.

1. INTRODUCTION

Current drugs interact with a variety of molecules with proteins being the prime targets (1-3). The process of drug design and characterization often relies on the structure of the underlying protein target(s) (4-12). However, recent research reveals that many biologically functional protein regions and even entire proteins might lack a stable tertiary and/or secondary structure in solution, indicating that a unique 3-D structure is not required for a protein to be biologically active. On the contrary, many crucial biological functions do in fact originate from or critically depend on the lack of ordered structure in a protein molecule (13-22). These intrinsically disordered proteins and protein regions (IDPs and IDPRs, respectively) are known to exist as highly dynamic structural ensembles, which can be classified as collapsed (molten globule-like) or extended (coil- or pre-molten globule-like) conformational ensembles (15, 23-25).

Recent years have observed interest in the use of IDPs and IDPRs as drug targets (26-29). This interest is fueled by the fact that intrinsic disorder is highly abundant in nature (19, 30-34). In particular, the overall levels of disorder increase with the increase in the organism complexity, culminating in highly disordered eukaryotic proteins, over a half of which are predicted to have long disordered regions (LDRs) (19, 30, 31, 33, 34). The LDRs are defined as regions of at least 30 consecutive disordered residues (31, 35-37) and are recognized as biologically functional intrinsically disordered domains (31, 38, 39). The high natural abundance of intrinsic disorder results in an intimate involvement of IDPs/IDPRs in many important biological processes (13-19, 23, 24, 40-48). High conformational flexibility of IDPs/IDPRs makes them perfect targets for various post-translational modifications, thereby providing useful means for the control, regulation, and modulation (47, 49, 50). The biological functions of IDPs/IDPRs are many and they complement functions of ordered proteins (45-47). Many IDPs undergo disorder-to-order transition upon function, e.g., due to the interaction with other proteins, nucleic acids, membranes, or small molecules (13-15, 17, 18, 42, 48, 51-54). The degree of the function-induced structural rearrangements vary over a very wide range. Typically, bound IDPs retain substantial levels of disorder, but some of them could be tightly folded (55, 56). Some IDPs/IDPRs are able to interact with a number of partners and some of these promiscuous binders were shown to fold differently being bound to different partners (53, 57, 58). Furthermore, binding promiscuity defines the frequent involvement of IDPs/IDPRs in functional regulation of their binding partners, assembly of supra-molecular complexes, and in the organization, maintenance, and control of complex protein-protein interaction networks (23, 42, 53, 59-61).

Many proteins associated with the various diseases are involved in recognition, regulation, and cell signaling, and a great number of them are IDPs or hybrid proteins containing ordered domains and functional IDPRs. This common involvement of IDPs in the pathogenesis of numerous human diseases gave rise to the "disorder in disorders" or D^2 concept (62-67). A partial list of human diseases originating from the misbehavior of IDPs includes cancer (51, 68-73), neurodegenerative (64, 74-77) and cardiovascular (78) diseases, type II diabetes (76),

AIDS (79), amyloidosis (80), and cystic fibrosis (81). Often, pathogenic misbehavior of IDPs/IDPRs originates from the point mutation(s) or from a protein exposure to internal or external toxins. Furthermore, this misbehavior can also be caused by impaired post-translational modifications, such as phosphorylation, advanced glycation, deamidation, racemization, etc., an increased probability of degradation, impaired trafficking, loss of binding partners, or oxidative damage (82). All these factors can act independently, additively, or synergistically (67, 82-84).

Given the high levels of abundance of IDPs/IDPRs in higher eukaryotes, their promiscuous binding and involvement in various human diseases we investigate how well they have penetrated into the human drug targets. We show that the druggable human proteome is biased towards structured proteins and characterized by a relatively low abundance of intrinsic disorder. We review reasons for this bias and discuss future of intrinsic disorder in the druggable human proteome.

2. ANALYSIS OF DRUGGABLE HUMAN PROTEOME

We comprehensively analyze the abundance of intrinsic disorder together with structural coverage in the current druggable human proteome for a representative set of drug classes including small molecule-based compounds, antibodies, peptides, proteins, antisense oligonucleotides, vaccines, siRNAs, adenoviral vectors, and aptamers. We also perform this analysis for a wide range of major functional classes of human drug targets, such as GPCRs, various enzyme classes, enzyme regulators, various receptors, transporters, transcription factors, cytokines, and growth factors.

The lists of drugs and human drug targets together with their categorization into classes were collected from a recent comprehensive resource from ref.(1). This resource includes data from the DrugBank database (85) cross-referenced with the Drugs@FDA database (http://www.fda.gov/drugsatfda) to collect FDA approved drugs and obtain list of established drug targets (86). It also incorporates novel drug targets that were obtained from the Drug in Clinical Trials Database that is maintained by CenterWatch (http://www.centerwatch.com/). We mapped target proteins into the complete human proteome (68,820 proteins) from UniProt (87) to collect their sequences and facilitate characterization of intrinsic disorder and structural coverage. The corresponding 1027 human proteins compose the druggable human proteome and include 556 established targets and 476 novel targets. We note that the novel targets were defined with a scope of a given drug and thus the sum of novel and established targets is larger than the number of all human targets. The drugs and target proteins were categorized into classes (1) utilizing the standardized annotations established by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology for enzymes (88), classification used by the Transporter Classification database for transporters (89), and classification extracted for UniProtKB and Swiss-Prot (90) for other types of targets. Our analysis considers 12 major drug classes and 18 major classes of drug targets that have at least

10 protein targets (Fig. 1). This restriction allows us to generate statistically sound estimates of the abundance of disorder and structural coverage for the selected classes of drug targets and druggable human proteome, and to compare them with the values for the human proteome. Figure 1 shows that among the target classes, several targets belonging to the kinase family, such as non-specific serine/threonine protein kinase, cytoplasmic protein tyrosine kinase, and mitogen-activated protein kinase, show a major percentage of novel targets (white bar). This relatively high prevalence of kinases among the novel drug targets is explained by the increasing recognition of kinases as important regulators of various cellular pathways and cascades critically involved in the pathogenesis of various diseases (91-95).

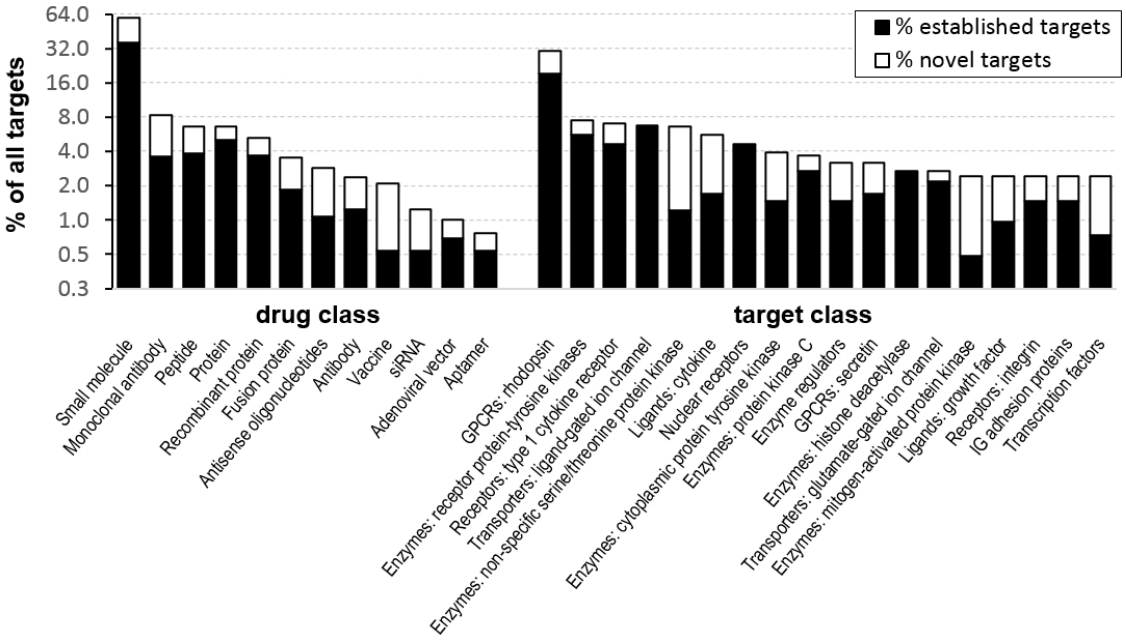

Figure 1. Breakdown of the fraction of targets for the considered 12 drug classes and 18 targets classes. The *y*-axis in in the logarithmic scale. The black and white bars denote established and novel targets, respectively, which are sorted in the descending by the total fraction of targets for drug and targets classes.

Given the large size of our data, we applied five complementary high-throughput disorder predictors including two versions of IUPred (96), which predict long and short disordered regions, and three versions of Espritz (97) that predict three main types of intrinsic disorder annotations including annotations via X-ray crystallography, NMR, and based on experimental evidence from the Disprot database (98). These methods offer good predictive performance based on a recent large-scale assessment (99). Their predictions were combined via majority vote consensus where a given residue is predicted as disordered if majority of the methods (three or more out of the five) predict it as disordered; otherwise the residue is predicted as structured. The same consensus was recently utilized in related studies (34, 100, 101). Furthermore, the use of consensus was previously shown to lead to an increase in the predictive performance (99, 102, 103). We quantified two measures of the abundance of intrinsic disorder. They include disorder content, defined as the fraction of disordered amino acids among all

amino acids in a given set of proteins, and number of long disorder regions (LDRs) per unit of 2000 amino acids, defined as number of LDRs in a given set of proteins divided by the total number of amino acids in these proteins and multiplied by 2000. These measures were computed for the entire human proteome (68,820 proteins), the druggable human proteome (1,027 proteins), and for each class of druggable proteins.

The structural coverage was calculated based on a method from ref. (104), which was also recently used to characterize relation between overall abundance of intrinsic disorder and structural coverage (34). Briefly, a given protein was compared to all sequences from the Protein Data Bank (105, 106) using three rounds of PSI-BLAST (107). The protein was assumed to be structured if PSI-BLAST found a hit with at least 50 amino acids in length and E-value < 0.001. This means that structured proteins have structural information for at least one long fragment of their sequences. The structural coverage of the human genome or a given class of druggable proteins is defined the fraction of structured sequences in this protein set.

3. STRUCTURAL COVERAGE AND INTRINSIC DISORDER IN DRUGGABLE HUMAN PROTEOME

The structural coverage, disorder content and count of LDRs for the whole human proteome, druggable human proteome, all major classes of druggable protein targets, and druggable targets associated with main classes of drugs are summarized in Figure 2. The druggable proteome is characterized by high levels of structural coverage, with about 94% of druggable targets having at least part of their sequences with solved structure (Fig. 2). To compare, the same estimate for the entire human proteomes results in significantly lower structural coverage value of 53%. This clear bias could be attributed to the widespread utilization of the structural information in various stages of drug development process including rational drug design (5-9, 108, 109), investigations of mechanistic insights into protein-drug interactions (110-114), drug repositioning (10, 11), and characterization of side-effects and off-targets of drugs (12, 115, 116). Furthermore, we observe an inverse trend for the abundance of intrinsic disorder. When compared with the human proteome, the druggable human proteome is characterized by almost two-fold lower disorder content (0.13 vs. 0.23) and similarly significantly lower counts of LDRs (0.65 vs 1.2). This could be explained by the fact that structural characterization of proteins containing substantial amounts of disorder is challenging (117, 118) and that such proteins are generally avoided in the target selection pipelines for structural genomics (119-122). This is a particularly important factor in the case of the human proteome, which is characterized by relatively high amounts of disorder compared to the bacterial and archaeal proteomes, and to the proteomes of some eukaryotic species (31, 33, 34, 123).

Perhaps not too surprisingly, Figure 2A reveals that druggable targets for all major types of drugs have universally very high structural coverage and significantly lower abundance of disorder when compared to the entire proteome. Small molecule-based compounds which

account for over 60% of all drugs (Fig. 1) have the structural coverage at 95% and disorder content at 0.12. However, targets from two drug classes, proteins and antisense oligonucleotides, reach disorder content values of 0.16 and 0.19, respectively. This is interesting particularly given the relatively high fraction of novel targets for the oligonucleotides (Fig. 1). Moreover, similarly high values of coverage and disorder are true for both established and novel drug targets (Fig. 2B). This demonstrates that currently there is no major shift in the drug target selection for the compounds that are in clinical trials. In our view, one of the obstacles that prevents deeper penetration of disordered proteins into the druggable human proteomes is the current paradigm that is focused on the structure-assisted drug design and characterization (124-134).

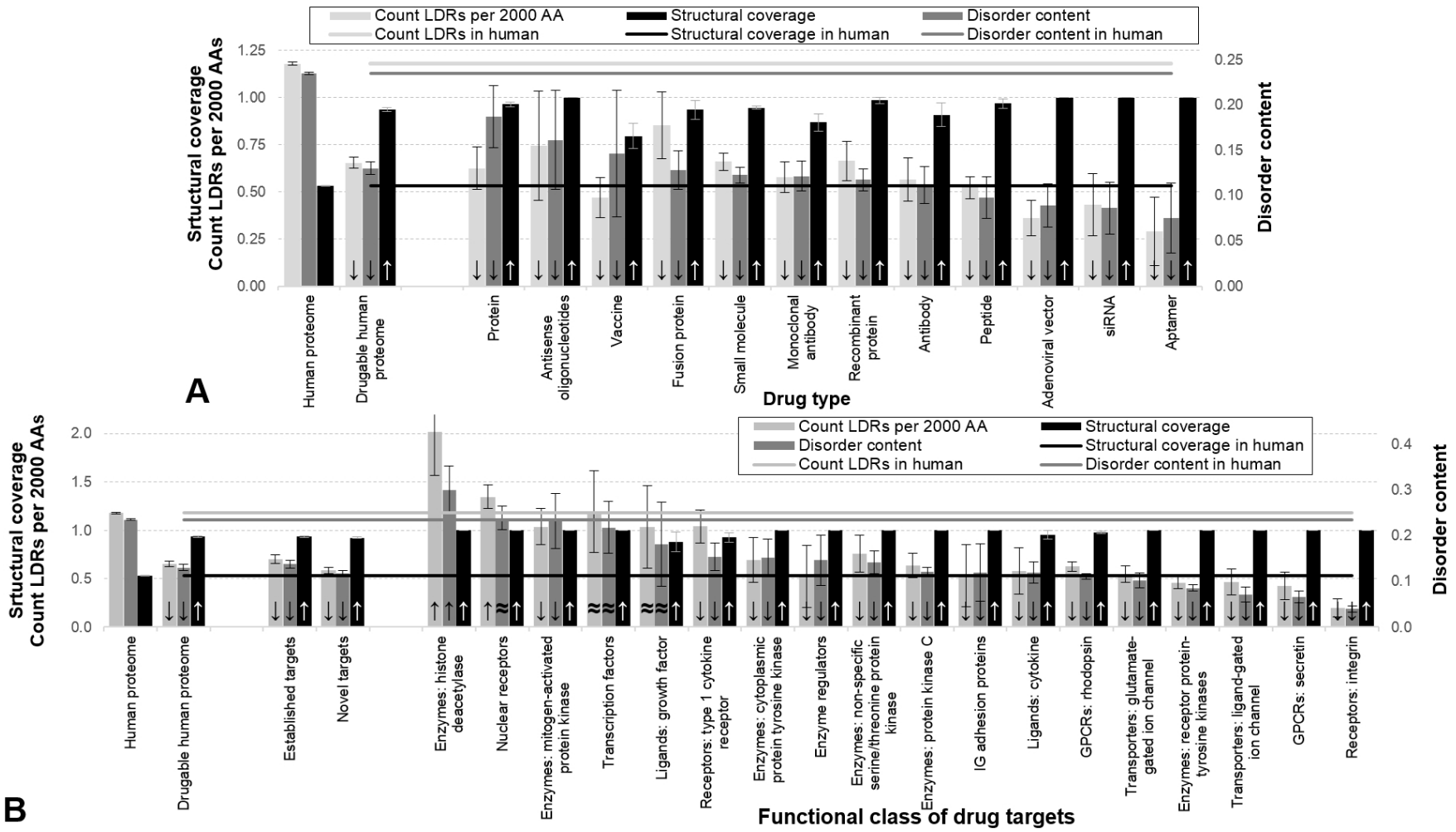

Figure 2. Structural coverage, number of LDRs and disorder content for the human proteome, druggable human proteome, and protein targets associated with major drug classes (Panel A), major classes of targets (Panel B), and novel vs. established targets (Panel B). Drug types and functional protein classes are sorted in the descending order by the disorder content. The measurements are based on 10 repetitions with 50% of randomly chosen proteins from a given dataset; bars denote the mean value with the corresponding standard deviations shown as the error bars. Solid horiontal lines denote the structural coverage (black), number of LDRs (light gray) and disorder content (dark gray) of the entire human proteome. The significance of the differences in the coverage, content and counts between the entire proteome and druggable proteome including each druggable class target is annotated at the base of the bars where ↑, ↓, and ≈ denote that druggable proteins have significantly higher, lower, or comparable values based on *p*-value of 0.05. Given that the measurements were normal, based on the Anderson-Darling test at *p*-value of 0.05, we used the *t*-test to assess the significance; otherwise we used the Wilcoxon rank-sum test.

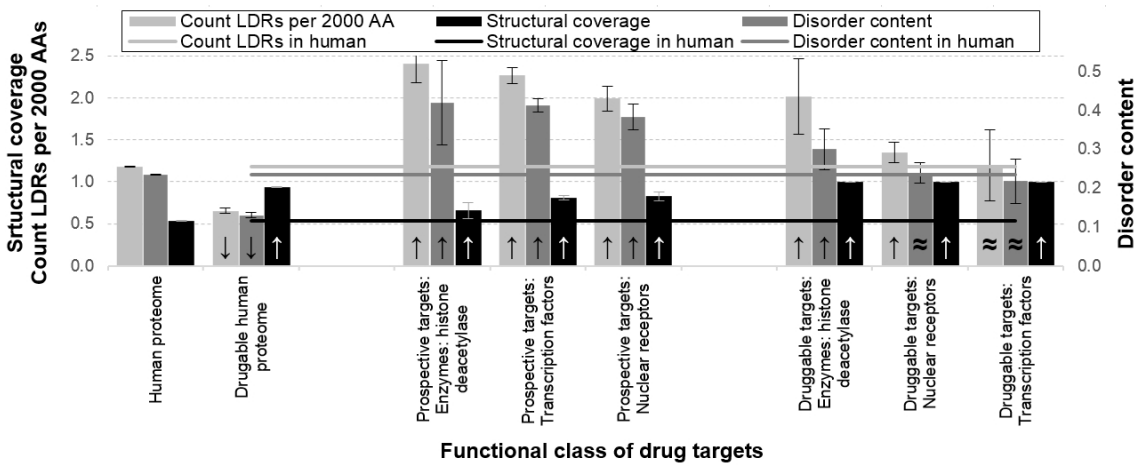

Figure 3. Structural coverage, number of LDRs and disorder content for the human proteome, druggable human proteome, and three classes of proteins characterized by relatively high disorder content: histone deacetylases, nuclear receptors and transcription factors. The analysis is performed for the current druggable targets from these three classes (set of bars on the right) and the remaining proteins from these three classes (set of bars in the middle). The measurements are based on 10 repetitions with 50% of randomly chosen proteins from a given dataset; bars denote the mean value with the corresponding standard deviations shown as the error bars. Solid horiontal lines denote the structural coverage (black), number of LDRs (light gray) and disorder content (dark gray) of the entire human proteome. The significance of the differences in the coverage, content and counts between the entire proteome and each set of proteins is annotated at the base of the bars where ↑, ↓, and ≈ denote that druggable proteins have significantly higher, lower, or comparable values based on p-value of 0.05. Given that the measurements were normal, based on the Anderson-Darling test at p-value of 0.05, we used the t-test to assess the significance; otherwise we used the Wilcoxon rank-sum test.

4. INTRINSICALLY DISORDERED PROTEINS AS DRUG TARGETS

Being involved in crucial protein-protein interactions and linked to the pathogenesis of various human diseases, IDPs/IDPRs represent novel, attractive, but challenging drug targets (26). Research shows that it is feasible to find molecules that target IDPRs (29). Although the best part of major functional classes of druggable targets have significant bias towards low amounts of disorder and high structural coverage, some types of targets have comparable or higher disorder content when compared to the whole proteome (Fig. 2B). The most notable examples are histone deacetylases, nuclear receptors, and transcription factors. In comparison with the whole human proteome, the druggable histone deacetylases are characterized by the significantly higher count of LDRs (2 vs. 1.2 for the whole genome) and significantly higher disorder content (0.30 vs. 0.23). The druggable nuclear receptors and transcription factors also have relative high disorder contents at 0.24 and 0.22, respectively. These three classes of drug targets account for 9.7% of all druggable targets that we considered (Fig. 1). We extracted a complete set of 20 well-annotated (supported by experimental evidence) histone deacetylases, 79 nuclear receptors and 408 transcription factors from the human proteome. Comparison of the druggable proteins from these three functional classes with the remaining proteins reveals that the former are actually depleted in disorder and LDRs when compared with their overall population in the whole genome (Fig. 3). The latter groups of proteins, which constitute prospective druggable targets, are characterized by very high LDR counts between 2 and 2.4 per 2000 residues and very high disorder content ranging between 0.38 and 0.42. These

measurements are significantly higher than for the entire human proteome and (obviously) than for the current druggable proteome. This suggest that the future of the drug design will likely include a stronger emphasis on the disordered protein targets. However, inherent lack of structure in these proteins calls for novel ways for the protein disorder-assisted drug discovery (28).

The protein disorder-assisted drug design depends on current rational drug design techniques, i.e., by targeting the ordered domains of proteins containing disordered regions and the structured partners of disordered proteins, and also requires novel approaches that no longer rely on a unique protein structure. The former approaches rely on the heterogeneous nature of IDPs containing ordered and disordered domains and the ability of IDPs/IDPRs to interact with ordered partners. The ordered domains can be targeted as typical structured proteins, and drugs blocking ligand binding sites in such ordered domains can be found by high-throughput screening of possible compounds or by rational drug design approaches based on the prior knowledge of the domain's 3-D structure (28). Drugs for blocking sites of ordered proteins involved in interactions with IDPs/IDPRs can also be designed based on the information on the 3-D structure of their ordered partner; an example case that involves a compound that inhibits interaction between a disordered domain of p53 and ordered Mdm2 was described in ref. (26).

The latter, novel approaches directly target disordered regions. Identification of such regions benefits from an observation that IDPs/IDPRs often bind their partners via relatively short contiguous motifs that often become ordered upon binding (135, 136) and that such regions can be accurately predicted (137, 138). One option for targeting disordered regions is to find/design drugs that induce misfolding of this region, where the structure induced via the interaction with the drug prevents function of the corresponding protein. An example here is related to disruption of association between two disordered proteins: Max and c-Myc oncoprotein (139). Even though structure-based rational design failed in this case, inhibitors were found via high-throughput screening (140). Using NMR, they were found to bind to the c-Myc monomer in a disordered region located in the c-Myc:Max interface, effectively preventing formation of the corresponding dimer (139). Other examples can be found in refs. (26, 28). Although these examples demonstrate that drugs can modulate function of disordered proteins and regions, the development of corresponding rational drug design methods is in very early stages (141). One recently proposed approach advocates use of molecular dynamics simulations (142).

LIST OF ABBREVIATIONS

IDP (Intrinsically disordered protein)
IDPR (Intrinsically disordered protein region)
D^2 (Disorder in disorders)
FDA (Food and drug administration)
LDR (Long disordered region)

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Lukasz Kurgan was supported by the Discovery grant 298328 from the Natural Sciences and Engineering Research Council (NSERC) Canada and the National Natural Science Foundation of China (NSFC) grants 31050110432 and 31150110577 and. Gang Hu was supported by the National Natural Science Foundation of China (NSFC) grant 11101226. Kui Wang was supported by the National Natural Science Foundation of China (NSFC) grant 11301286.

REFERENCES

1. Rask-Andersen M, Masuram S, Schioth HB. The Druggable Genome: Evaluation of Drug Targets in Clinical Trials Suggests Major Shifts in Molecular Class and Indication. Annu Rev Pharmacol. 2014;54:9-26. PubMed PMID: WOS:000329822200002. English.

2. Arinaminpathy Y, Khurana E, Engelman DM, Gerstein MB. Computational analysis of membrane proteins: the largest class of drug targets. Drug discovery today. 2009 Dec;14(23-24):1130-5. PubMed PMID: 19733256. PMCID: 2796609.

3. Zhang CC, Kast J. Applications of current proteomics techniques in modern drug design. Current computer-aided drug design. 2010 Sep;6(3):147-64. PubMed PMID: 20438442.

4. Lahti JL, Tang GW, Capriotti E, Liu T, Altman RB. Bioinformatics and variability in drug response: a protein structural perspective. Journal of the Royal Society, Interface / the Royal Society. 2012 Jul 7;9(72):1409-37. PubMed PMID: 22552919. PMCID: 3367825.

5. Lounnas V, Ritschel T, Kelder J, McGuire R, Bywater RP, Foloppe N. Current progress in Structure-Based Rational Drug Design marks a new mindset in drug discovery. Computational and structural biotechnology journal. 2013;5:e201302011. PubMed PMID: 24688704. PMCID: 3962124.

6. Mavromoustakos T, Durdagi S, Koukoulitsa C, Simcic M, Papadopoulos MG, Hodoscek M, et al. Strategies in the rational drug design. Current medicinal chemistry. 2011;18(17):2517-30. PubMed PMID: 21568895.

7. Lundstrom K. Structural genomics: the ultimate approach for rational drug design. Mol Biotechnol. 2006 Oct;34(2):205-12. PubMed PMID: 17172666.

8. Patel D, Bauman JD, Arnold E. Advantages of crystallographic fragment screening: functional and mechanistic insights from a powerful platform for efficient drug discovery. Prog Biophys Mol Biol. 2014 Nov-Dec;116(2-3):92-100. PubMed PMID: 25117499.

9. Hung CL, Chen CC. Computational approaches for drug discovery. Drug development research. 2014 Sep;75(6):412-8. PubMed PMID: 25195585.

10. Ma DL, Chan DS, Leung CH. Drug repositioning by structure-based virtual screening. Chemical Society reviews. 2013 Mar 7;42(5):2130-41. PubMed PMID: 23288298.

11. Liu X, Zhu F, Ma XH, Shi Z, Yang SY, Wei YQ, et al. Predicting targeted polypharmacology for drug repositioning and multi- target drug discovery. Current medicinal chemistry. 2013;20(13):1646-61. PubMed PMID: 23410165.

12. Hu G, Gao J, Wang K, Mizianty MJ, Ruan J, Kurgan L. Finding protein targets for small biologically relevant ligands across fold space using inverse ligand binding predictions. Structure. 2012 Nov 7;20(11):1815-22. PubMed PMID: 23141694.
13. Wright PE, Dyson HJ. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J Mol Biol. 1999 Oct 22;293(2):321-31. PubMed PMID: 10550212.
14. Uversky VN, Gillespie JR, Fink AL. Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins. 2000 Nov 15;41(3):415-27. PubMed PMID: 11025552.
15. Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, et al. Intrinsically disordered protein. J Mol Graph Model. 2001;19(1):26-59. PubMed PMID: 11381529.
16. Dunker AK, Obradovic Z. The protein trinity--linking function and disorder. Nat Biotechnol. 2001 Sep;19(9):805-6. PubMed PMID: 11533628.
17. Tompa P. Intrinsically unstructured proteins. Trends Biochem Sci. 2002 Oct;27(10):527-33. PubMed PMID: 12368089.
18. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol. 2005 Mar;6(3):197-208. PubMed PMID: 15738986. Epub 2005/03/02. eng.
19. Uversky VN. The mysterious unfoldome: structureless, underappreciated, yet vital part of any given proteome. J Biomed Biotechnol. 2010;2010:568068. PubMed PMID: 20011072. PMCID: 2789583. Epub 2009/12/17. eng.
20. Dunker AK, Oldfield CJ, Meng J, Romero P, Yang JY, Chen JW, et al. The unfoldomics decade: an update on intrinsically disordered proteins. BMC Genomics. 2008;9 Suppl 2:S1. PubMed PMID: 18831774. PMCID: 2559873. Epub 2008/10/10. eng.
21. Dunker AK, Silman I, Uversky VN, Sussman JL. Function and structure of inherently disordered proteins. Curr Opin Struct Biol. 2008 Dec;18(6):756-64. PubMed PMID: 18952168. Epub 2008/10/28. eng.
22. Habchi J, Tompa P, Longhi S, Uversky VN. Introducing Protein Intrinsic Disorder. Chemical reviews. 2014 Jul 9;114(13):6561-88. PubMed PMID: WOS:000338980600002. English.
23. Daughdrill GW, Pielak GJ, Uversky VN, Cortese MS, Dunker AK. Natively disordered proteins. In: Buchner J, Kiefhaber T, editors. Handbook of Protein Folding. Weinheim, Germany: Wiley-VCH, Verlag GmbH & Co. KGaA; 2005. p. 271-353.
24. Uversky VN. Protein folding revisited. A polypeptide chain at the folding-misfolding-nonfolding cross-roads: which way to go? Cell Mol Life Sci. 2003 Sep;60(9):1852-71. PubMed PMID: 14523548.
25. Turoverov KK, Kuznetsova IM, Uversky VN. The protein kingdom extended: ordered and intrinsically disordered proteins, their folding, supramolecular complex formation, and aggregation. Prog Biophys Mol Biol. 2010 Jun-Jul;102(2-3):73-84. PubMed PMID: 20097220. PMCID: 2916636. Epub 2010/01/26. eng.
26. Uversky VN. Intrinsically disordered proteins and novel strategies for drug discovery. Expert opinion on drug discovery. 2012 Jun;7(6):475-88. PubMed PMID: 22559227.
27. McEwan IJ. Intrinsic disorder in the androgen receptor: identification, characterisation and drugability. Molecular bioSystems. 2012;8(1):82-90. PubMed PMID: WOS:000297561600008. English.
28. Dunker AK, Uversky VN. Drugs for 'protein clouds': targeting intrinsically disordered transcription factors. Current opinion in pharmacology. 2010 Dec;10(6):782-8. PubMed PMID: 20889377.

29. Metallo SJ. Intrinsically disordered proteins are potential drug targets. Curr Opin Chem Biol. 2010 Aug;14(4):481-8. PubMed PMID: 20598937. PMCID: 2918680. Epub 2010/07/06. eng.
30. Dunker AK, Obradovic Z, Romero P, Garner EC, Brown CJ. Intrinsic protein disorder in complete genomes. Genome Inform Ser Workshop Genome Inform. 2000;11:161-71. PubMed PMID: 11700597. Epub 2001/11/09. eng.
31. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. J Mol Biol. 2004 Mar 26;337(3):635-45. PubMed PMID: 15019783. Epub 2004/03/17. eng.
32. Xue B, Williams RW, Oldfield CJ, Dunker AK, Uversky VN. Archaic chaos: intrinsically disordered proteins in Archaea. BMC Syst Biol. 2010;4 Suppl 1:S1. PubMed PMID: 20522251. PMCID: 2880407. Epub 2010/06/11. eng.
33. Xue B, Dunker AK, Uversky VN. Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. Journal of biomolecular structure & dynamics. 2012;30(2):137-49. PubMed PMID: 22702725. Epub 2012/06/19. eng.
34. Peng Z, Yan J, Fan X, Mizianty MJ, Xue B, Wang K, et al. Exceptionally abundant exceptions: comprehensive characterization of intrinsic disorder in all domains of life. Cell Mol Life Sci. 2015 Jan;72(1):137-51. PubMed PMID: 24939692. Epub 2014/06/19. eng.
35. Lobley A, Swindells MB, Orengo CA, Jones DT. Inferring function using patterns of native disorder in proteins. PLoS computational biology. 2007 Aug;3(8):1567-79. PubMed PMID: WOS:000249767100015. English.
36. Pancsa R, Tompa P. Structural Disorder in Eukaryotes. PloS one. 2012 Apr 5;7(4). PubMed PMID: WOS:000305010500022. English.
37. Peng Z, Mizianty MJ, Kurgan L. Genome-scale prediction of proteins with long intrinsically disordered regions. Proteins. 2014 Jan;82(1):145-58. PubMed PMID: 23798504.
38. Pentony MM, Jones DT. Modularity of intrinsic disorder in the human proteome. Proteins-Structure Function and Bioinformatics. 2010 Jan;78(1):212-21. PubMed PMID: WOS:000273155400019. English.
39. Tompa P, Fuxreiter M, Oldfield CJ, Simon I, Dunker AK, Uversky VN. Close encounters of the third kind: disordered domains and the interactions of proteins. BioEssays : news and reviews in molecular, cellular and developmental biology. 2009 Mar;31(3):328-35. PubMed PMID: 19260013.
40. Dunker AK, Brown CJ, Obradovic Z. Identification and functions of usefully disordered proteins. Advances in protein chemistry. 2002;62:25-49. PubMed PMID: 12418100.
41. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z. Intrinsic disorder and protein function. Biochemistry. 2002 May 28;41(21):6573-82. PubMed PMID: 12022860.
42. Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN. Flexible nets. The roles of intrinsic disorder in protein interaction networks. Febs J. 2005 Oct;272(20):5129-48. PubMed PMID: 16218947.
43. Tompa P. The interplay between structure and function in intrinsically unstructured proteins. FEBS Lett. 2005 Jun 13;579(15):3346-54. PubMed PMID: 15943980.
44. Tompa P, Csermely P. The role of structural disorder in the function of RNA and protein chaperones. FASEB J. 2004 Aug;18(11):1169-75. PubMed PMID: 15284216.
45. Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Uversky VN, et al. Functional

anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. J Proteome Res. 2007 May;6(5):1882-98. PubMed PMID: 17391014. Epub 2007/03/30. eng.
46. Vucetic S, Xie H, Iakoucheva LM, Oldfield CJ, Dunker AK, Obradovic Z, et al. 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 May;6(5):1899-916. PubMed PMID: 17391015. Epub 2007/03/30. eng.
47. Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Obradovic Z, et al. Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. J Proteome Res. 2007 May;6(5):1917-32. PubMed PMID: 17391016. Epub 2007/03/30. eng.
48. Uversky VN, Oldfield CJ, Dunker AK. Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. J Mol Recognit. 2005 Aug 10;18(5):343-84. PubMed PMID: 16094605.
49. Uversky VN. Intrinsic disorder-based protein interactions and their modulators. Current pharmaceutical design. 2013;19(23):4191-213. PubMed PMID: 23170892.
50. Pejaver V, Hsu WL, Xin F, Dunker AK, Uversky VN, Radivojac P. The structural and functional signatures of proteins that undergo multiple events of post-translational modification. Protein Sci. 2014 Aug;23(8):1077-93. PubMed PMID: 24888500. PMCID: 4116656.
51. Iakoucheva LM, Brown CJ, Lawson JD, Obradovic Z, Dunker AK. Intrinsic disorder in cell-signaling and cancer-associated proteins. J Mol Biol. 2002;323:573-84.
52. Dyson HJ, Wright PE. Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance. Adv Protein Chem. 2002;62:311-40. PubMed PMID: 12418108.
53. Oldfield CJ, Meng J, Yang JY, Yang MQ, Uversky VN, Dunker AK. Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. BMC Genomics. 2008;9 Suppl 1:S1. PubMed PMID: 18366598.
54. Fuxreiter M, Toth-Petroczy A, Kraut DA, Matouschek AT, Lim RYH, Xue B, et al. Disordered Proteinaceous Machines. Chemical reviews. 2014 Jul 9;114(13):6806-43. PubMed PMID: WOS:000338980600009. English.
55. Tompa P, Fuxreiter M. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. Trends Biochem Sci. 2008 Jan;33(1):2-8. PubMed PMID: 18054235. Epub 2007/12/07. eng.
56. Uversky VN. Multitude of binding modes attainable by intrinsically disordered proteins: a portrait gallery of disorder-based complexes. Chemical Society reviews. 2011 Mar;40(3):1623-34. PubMed PMID: 21049125.
57. Hsu WL, Oldfield C, Meng J, Huang F, Xue B, Uversky VN, et al. Intrinsic protein disorder and protein-protein interactions. Pac Symp Biocomput. 2012:116-27. PubMed PMID: 22174268.
58. Hsu WL, Oldfield CJ, Xue B, Meng J, Huang F, Romero P, et al. Exploring the binding diversity of intrinsically disordered proteins involved in one-to-many binding. Protein science : a publication of the Protein Society. 2013 Mar;22(3):258-73. PubMed PMID: 23233352. PMCID: 3595456.
59. Haynes C, Oldfield CJ, Ji F, Klitgord N, Cusick ME, Radivojac P, et al. Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. PLoS Comput Biol. 2006 Aug 4;2(8):e100. PubMed PMID: 16884331.
60. Dosztanyi Z, Chen J, Dunker AK, Simon I, Tompa P. Disorder and sequence repeats in hub proteins

and their implications for network evolution. J Proteome Res. 2006 Nov;5(11):2985-95. PubMed PMID: 17081050.
61. Singh GP, Ganapathi M, Dash D. Role of intrinsic disorder in transient interactions of hub proteins. Proteins. 2007 Mar 1;66(4):761-5. PubMed PMID: 17154416.
62. Uversky VN, Oldfield CJ, Dunker AK. Intrinsically disordered proteins in human diseases: introducing the D2 concept. Annu Rev Biophys. 2008;37:215-46. PubMed PMID: 18573080. Epub 2008/06/25. eng.
63. Midic U, Oldfield CJ, Dunker AK, Obradovic Z, Uversky VN. Unfoldomics of human genetic diseases: illustrative examples of ordered and intrinsically disordered members of the human diseasome. Protein Pept Lett. 2009;16(12):1533-47. PubMed PMID: 20001916. Epub 2009/12/17. eng.
64. Uversky VN, Oldfield CJ, Midic U, Xie H, Xue B, Vucetic S, et al. Unfoldomics of human diseases: linking protein intrinsic disorder with diseases. BMC Genomics. 2009;10 Suppl 1:S7. PubMed PMID: 19594884. PMCID: 2709268. Epub 2009/07/25. eng.
65. Uversky VN. Intrinsic disorder in proteins associated with neurodegenerative diseases. Front Biosci. 2009;14:5188-238. PubMed PMID: 19482612. Epub 2009/06/02. eng.
66. Uversky VN. Targeting intrinsically disordered proteins in neurodegenerative and protein dysfunction diseases: another illustration of the D(2) concept. Expert Rev Proteomics. 2010 Aug;7(4):543-64. PubMed PMID: 20653509. Epub 2010/07/27. eng.
67. Uversky VN, Dave V, Iakoucheva LM, Malaney P, Metallo SJ, Pathak RR, et al. Pathological unfoldomics of uncontrolled chaos: intrinsically disordered proteins and human diseases. Chem Rev. 2014 Jul 9;114(13):6844-79. PubMed PMID: 24830552. PMCID: 4100540.
68. Wells M, Tidow H, Rutherford TJ, Markwick P, Jensen MR, Mylonas E, et al. Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. Proc Natl Acad Sci U S A. 2008 Apr 15;105(15):5762-7. PubMed PMID: 18391200. PMCID: 2311362. Epub 2008/04/09. eng.
69. Mark WY, Liao JC, Lu Y, Ayed A, Laister R, Szymczyna B, et al. Characterization of segments from the central region of BRCA1: an intrinsically disordered scaffold for multiple protein-protein and protein-DNA interactions? J Mol Biol. 2005 Jan 14;345(2):275-87. PubMed PMID: 15571721.
70. Ng KP, Potikyan G, Savene RO, Denny CT, Uversky VN, Lee KA. Multiple aromatic side chains within a disordered structure are critical for transcription and transforming activity of EWS family oncoproteins. Proc Natl Acad Sci U S A. 2007 Jan 9;104(2):479-84. PubMed PMID: 17202261. PMCID: 1766410. Epub 2007/01/05. eng.
71. Uversky VN, Roman A, Oldfield CJ, Dunker AK. Protein intrinsic disorder and human papillomaviruses: increased amount of disorder in E6 and E7 oncoproteins from high risk HPVs. J Proteome Res. 2006 Aug;5(8):1829-42. PubMed PMID: 16889404. Epub 2006/08/08. eng.
72. Malaney P, Pathak RR, Xue B, Uversky VN, Dave V. Intrinsic Disorder in PTEN and its Interactome Confers Structural Plasticity and Functional Versatility. Sci Rep. 2013;3:2035. PubMed PMID: 23783762. PMCID: 3687229. Epub 2013/06/21. eng.
73. Rajagopalan K, Mooney SM, Parekh N, Getzenberg RH, Kulkarni P. A majority of the cancer/testis antigens are intrinsically disordered proteins. Journal of cellular biochemistry. 2011 Nov;112(11):3256-67. PubMed PMID: 21748782. PMCID: 3214731.
74. Uversky VN. A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered

protein involved in neurodegenerative disorders. J Biomol Struct Dyn. 2003 Oct;21(2):211-34. PubMed PMID: 12956606. Epub 2003/09/06. eng.
75. Uversky VN. Intrinsic disorder in proteins associated with neurodegenerative diseases. Front Biosci (Landmark Ed). 2009;14:5188-238. PubMed PMID: 19482612. Epub 2009/06/02. eng.
76. Uversky VN, Oldfield CJ, Dunker AK. Intrinsically disordered proteins in human diseases: introducing the D2 concept. Annu Rev Biophys. 2008;37:215-46. PubMed PMID: 18573080. Epub 2008/06/25. eng.
77. Dolan PT, Roth AP, Xue B, Sun R, Dunker K, Uversky VN, et al. Intrinsic disorder mediates hepatitis C virus core – host cell protein interactions. Protein Science. 2014. Epub Nov 25.
78. Cheng Y, LeGall T, Oldfield CJ, Dunker AK, Uversky VN. Abundance of intrinsic disorder in protein associated with cardiovascular disease. Biochemistry. 2006 Sep 5;45(35):10448-60. PubMed PMID: 16939197. Epub 2006/08/31. eng.
79. Casu F, Duggan BM, Hennig M. The arginine-rich RNA-binding motif of HIV-1 Rev is intrinsically disordered and folds upon RRE binding. Biophys J. 2013 Aug 20;105(4):1004-17. PubMed PMID: 23972852. PMCID: 3752131. Epub 2013/08/27. eng.
80. Uversky VN. Amyloidogenesis of natively unfolded proteins. Current Alzheimer research. 2008 Jun;5(3):260-87. PubMed PMID: 18537543. PMCID: 2719783.
81. Baker JM, Hudson RP, Kanelis V, Choy WY, Thibodeau PH, Thomas PJ, et al. CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices. Nat Struct Mol Biol. 2007 Aug;14(8):738-45. PubMed PMID: 17660831. Epub 2007/07/31. eng.
82. Kathiriya JJ, Pathak RR, Clayman E, Xue B, Uversky VN, Dave V. Presence and utility of intrinsically disordered regions in kinases. Molecular bioSystems. 2014 Sep 30;10(11):2876-88. PubMed PMID: 25099472.
83. Uversky VN. Intrinsically disordered proteins and novel strategies for drug discovery. Expert Opin Drug Discov. 2012 Jun;7(6):475-88. PubMed PMID: 22559227. Epub 2012/05/09. eng.
84. Vacic V, Markwick PR, Oldfield CJ, Zhao X, Haynes C, Uversky VN, et al. Disease-associated mutations disrupt functionally important regions of intrinsic protein disorder. PLoS Comput Biol. 2012;8(10):e1002709. PubMed PMID: 23055912. PMCID: 3464192.
85. Knox C, Law V, Jewison T, Liu P, Ly S, Frolkis A, et al. DrugBank 3.0: a comprehensive resource for 'Omics' research on drugs. Nucleic Acids Res. 2011 Jan;39:D1035-D41. PubMed PMID: WOS:000285831700163. English.
86. Rask-Andersen M, Almen MS, Schioth HB. Trends in the exploitation of novel drug targets. Nat Rev Drug Discov. 2011 Aug;10(8):579-90. PubMed PMID: WOS:000293326200019. English.
87. Bairoch A, Bougueleret L, Altairac S, Amendolia V, Auchincloss A, Puy GA, et al. The universal protein resource (UniProt). Nucleic Acids Res. 2007 Jan;35:D193-D7. PubMed PMID: WOS:000243494600040. English.
88. International Union of Biochemistry and Molecular Biology. Nomenclature Committee., Webb EC. Enzyme nomenclature 1992 : recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzymes. San Diego: Published for the International Union of Biochemistry and Molecular Biology by Academic Press; 1992. xiii, 862 p. p.
89. Saier MH, Jr., Tran CV, Barabote RD. TCDB: the Transporter Classification Database for

membrane transport protein analyses and information. Nucleic acids research. 2006 Jan 1;34(Database issue):D181-6. PubMed PMID: 16381841. PMCID: 1334385.
90. Boutet E, Lieberherr D, Tognolli M, Schneider M, Bairoch A. UniProtKB/Swiss-Prot. Methods in molecular biology. 2007;406:89-112. PubMed PMID: 18287689.
91. Cohen P. Targeting protein kinases for the development of anti-inflammatory drugs. Current opinion in cell biology. 2009 Apr;21(2):317-24. PubMed PMID: 19217767.
92. Coskun M, Olsen J, Seidelin JB, Nielsen OH. MAP kinases in inflammatory bowel disease. Clinica chimica acta; international journal of clinical chemistry. 2011 Mar 18;412(7-8):513-20. PubMed PMID: 21185271.
93. Vancraenenbroeck R, Lobbestael E, De Maeyer M, Baekelandt V, Taymans JM. Kinases as Targets for Parkinson's Disease: From Genetics to Therapy. Cns Neurol Disord-Dr. 2011 Sep;10(6):724-40. PubMed PMID: WOS:000295404000010. English.
94. Robak P, Robak T. A Targeted Therapy for Protein and Lipid Kinases in Chronic Lymphocytic Leukemia. Current medicinal chemistry. 2012 Nov;19(31):5294-318. PubMed PMID: WOS:000310598500004. English.
95. Kaidanovich-Beilin O, Eldar-Finkelman H. Peptides targeting protein kinases: Strategies and implications. Physiology. 2006 Dec;21:411-8. PubMed PMID: WOS:000242539900006. English.
96. Dosztanyi Z, Csizmok V, Tompa P, Simon I. The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. Journal of Molecular Biology. 2005 Apr 8;347(4):827-39. PubMed PMID: WOS:000227959200012. English.
97. Walsh I, Martin AJM, Di Domenico T, Tosatto SCE. ESpritz: accurate and fast prediction of protein disorder. Bioinformatics. 2012 Feb 15;28(4):503-9. PubMed PMID: WOS:000300490500008. English.
98. Sickmeier M, Hamilton JA, LeGall T, Vacic V, Cortese MS, Tantos A, et al. DisProt: the database of disordered proteins. Nucleic Acids Res. 2007 Jan;35:D786-D93. PubMed PMID: WOS:000243494600154. English.
99. Walsh I, Giollo M, Di Domenico T, Ferrari C, Zimmermann O, Tosatto SC. Comprehensive large-scale assessment of intrinsic protein disorder. Bioinformatics. 2015 Jan 15;31(2):201-8. PubMed PMID: 25246432.
100. Howell M, Green R, Killeen A, Wedderburn L, Picascio V, Rabionet A, et al. Not That Rigid Midgets and Not So Flexible Giants: On the Abundance and Roles of Intrinsic Disorder in Short and Long Proteins. J Biol Syst. 2012 Dec;20(4):471-511. PubMed PMID: WOS:000314608500009. English.
101. Peng Z, Oldfield CJ, Xue B, Mizianty MJ, Dunker AK, Kurgan L, et al. A creature with a hundred waggly tails: intrinsically disordered proteins in the ribosome. Cell Mol Life Sci. 2014 Apr;71(8):1477-504. PubMed PMID: 23942625.
102. Jin L, Wang W, Fang G. Targeting protein-protein interaction by small molecules. Annual review of pharmacology and toxicology. 2014;54:435-56. PubMed PMID: 24160698.
103. Peng Z, Kurgan L. On the complementarity of the consensus-based disorder prediction. Pac Symp Biocomput. 2012:176-87. PubMed PMID: 22174273.
104. Vitkup D, Melamud E, Moult J, Sander C. Completeness in structural genomics. Nature structural biology. 2001 Jun;8(6):559-66. PubMed PMID: 11373627.
105. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res. 2000 Jan 1;28(1):235-42. PubMed PMID: WOS:000084896300069. English.

106. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, et al. The Protein Data Bank. Acta Crystallogr D. 2002 Jun;58:899-907. PubMed PMID: WOS:000176271000003. English.
107. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997 Sep 1;25(17):3389-402. PubMed PMID: WOS:A1997XU79300002. English.
108. Gane PJ, Dean PM. Recent advances in structure-based rational drug design. Curr Opin Struct Biol. 2000 Aug;10(4):401-4. PubMed PMID: 10981625.
109. Pflug A, Guilligay D, Reich S, Cusack S. Structure of influenza A polymerase bound to the viral RNA promoter. Nature. 2014 Dec 18;516(7531):355-60. PubMed PMID: 25409142.
110. Christopoulos A. Advances in G protein-coupled receptor allostery: from function to structure. Molecular pharmacology. 2014 Nov;86(5):463-78. PubMed PMID: 25061106.
111. Ding X, Zhao X, Watts A. G-protein-coupled receptor structure, ligand binding and activation as studied by solid-state NMR spectroscopy. The Biochemical journal. 2013 Mar 15;450(3):443-57. PubMed PMID: 23445222.
112. Pielak RM, Schnell JR, Chou JJ. Mechanism of drug inhibition and drug resistance of influenza A M2 channel. Proc Natl Acad Sci U S A. 2009 May 5;106(18):7379-84. PubMed PMID: 19383794. PMCID: 2678642.
113. Cady SD, Schmidt-Rohr K, Wang J, Soto CS, Degrado WF, Hong M. Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. Nature. 2010 Feb 4;463(7281):689-92. PubMed PMID: 20130653. PMCID: 2818718.
114. Tan Q, Zhu Y, Li J, Chen Z, Han GW, Kufareva I, et al. Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. Science. 2013 Sep 20;341(6152):1387-90. PubMed PMID: 24030490. PMCID: 3819204.
115. Hu G, Wang K, Groenendyk J, Barakat K, Mizianty MJ, Ruan J, et al. Human structural proteome-wide characterization of Cyclosporine A targets. Bioinformatics. 2014 Dec 15;30(24):3561-6. PubMed PMID: 25172926. PMCID: 4253830.
116. Arooj M, Sakkiah S, Cao GP, Kim S, Arulalapperumal V, Lee KW. Finding off-targets, biological pathways & target diseases for chymase inhibitors via structure-based systems biology approach. Proteins. 2014 Aug 21. PubMed PMID: 25143259.
117. Oldfield CJ, Xue B, Van YY, Ulrich EL, Markley JL, Dunker AK, et al. Utilization of protein intrinsic disorder knowledge in structural proteomics. Biochimica et biophysica acta. 2013 Feb;1834(2):487-98. PubMed PMID: 23232152. PMCID: 3951342.
118. Mizianty MJ, Kurgan L. Sequence-based prediction of protein crystallization, purification and production propensity. Bioinformatics. 2011 Jul 1;27(13):i24-33. PubMed PMID: 21685077. PMCID: 3117383.
119. Linding R, Jensen LJ, Diella F, Bork P, Gibson TJ, Russell RB. Protein disorder prediction: implications for structural proteomics. Structure. 2003 Nov;11(11):1453-9. PubMed PMID: 14604535.
120. Slabinski L, Jaroszewski L, Rodrigues AP, Rychlewski L, Wilson IA, Lesley SA, et al. The challenge of protein structure determination--lessons from structural genomics. Protein science : a publication of the Protein Society. 2007 Nov;16(11):2472-82. PubMed PMID: 17962404. PMCID: 2211687.
121. Mizianty MJ, Fan X, Yan J, Chalmers E, Woloschuk C, Joachimiak A, et al. Covering complete

proteomes with X-ray structures: a current snapshot. Acta crystallographica Section D, Biological crystallography. 2014 Nov;70(Pt 11):2781-93. PubMed PMID: 25372670. PMCID: 4220968.
122. Mizianty MJ, Kurgan LA. CRYSpred: accurate sequence-based protein crystallization propensity prediction using sequence-derived structural characteristics. Protein Pept Lett. 2012 Jan;19(1):40-9. PubMed PMID: 21919861.
123. Yan J, Mizianty MJ, Filipow PL, Uversky VN, Kurgan L. RAPID: fast and accurate sequence-based prediction of intrinsic disorder content on proteomic scale. Biochim Biophys Acta. 2013 Aug;1834(8):1671-80. PubMed PMID: 23732563.
124. Mallipeddi PL, Kumar G, White SW, Webb TR. Recent advances in computer-aided drug design as applied to anti-influenza drug discovery. Curr Top Med Chem. 2014;14(16):1875-89. PubMed PMID: 25262796.
125. Hoi PM, Li S, Vong CT, Tseng HH, Kwan YW, Lee SM. Recent advances in structure-based drug design and virtual screening of VEGFR tyrosine kinase inhibitors. Methods. 2015 Jan;71:85-91. PubMed PMID: 25239735.
126. Hilgenfeld R. From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. Febs J. 2014 Sep;281(18):4085-96. PubMed PMID: 25039866.
127. Congreve M, Dias JM, Marshall FH. Structure-based drug design for G protein-coupled receptors. Progress in medicinal chemistry. 2014;53:1-63. PubMed PMID: 24418607.
128. Caliandro R, Belviso DB, Aresta BM, de Candia M, Altomare CD. Protein crystallography and fragment-based drug design. Future medicinal chemistry. 2013 Jun;5(10):1121-40. PubMed PMID: 23795969.
129. Ferreira RS, Andricopulo AD. Structure-based drug design to overcome drug resistance: challenges and opportunities. Current pharmaceutical design. 2014;20(5):687-93. PubMed PMID: 23688077.
130. Mao L, Yang Y. Targeting the nuclear transport machinery by rational drug design. Current pharmaceutical design. 2013;19(12):2318-25. PubMed PMID: 23082981.
131. Du J, Cross TA, Zhou HX. Recent progress in structure-based anti-influenza drug design. Drug discovery today. 2012 Oct;17(19-20):1111-20. PubMed PMID: 22704956. PMCID: 3459301.
132. Du QS, Wei H, Huang RB, Chou KC. Progress in structure-based drug design against influenza A virus. Expert opinion on drug discovery. 2011 Jun;6(6):619-31. PubMed PMID: 22646152.
133. Speck-Planche A, Cordeiro MN. Computer-aided drug design methodologies toward the design of anti-hepatitis C agents. Curr Top Med Chem. 2012;12(8):802-13. PubMed PMID: 22352909.
134. Chen L, Morrow JK, Tran HT, Phatak SS, Du-Cuny L, Zhang S. From laptop to benchtop to bedside: structure-based drug design on protein targets. Current pharmaceutical design. 2012;18(9):1217-39. PubMed PMID: 22316152. PMCID: 3820560.
135. Vacic V, Oldfield CJ, Mohan A, Radivojac P, Cortese MS, Uversky VN, et al. Characterization of molecular recognition features, MoRFs, and their binding partners. J Proteome Res. 2007 Jun;6(6):2351-66. PubMed PMID: 17488107. PMCID: 2570643. Epub 2007/05/10. eng.
136. Mohan A, Oldfield CJ, Radivojac P, Vacic V, Cortese MS, Dunker AK, et al. Analysis of molecular recognition features (MoRFs). J Mol Biol. 2006 Oct 6;362(5):1043-59. PubMed PMID: 16935303.
137. Disfani FM, Hsu WL, Mizianty MJ, Oldfield CJ, Xue B, Dunker AK, et al. MoRFpred, a computational tool for sequence-based prediction and characterization of short disorder-to-order

transitioning binding regions in proteins. Bioinformatics. 2012 Jun 15;28(12):i75-83. PubMed PMID: 22689782. PMCID: 3371841.
138. Meszaros B, Simon I, Dosztanyi Z. Prediction of protein binding regions in disordered proteins. PLoS computational biology. 2009 May;5(5):e1000376. PubMed PMID: 19412530. PMCID: 2671142.
139. Hammoudeh DI, Follis AV, Prochownik EV, Metallo SJ. Multiple independent binding sites for small-molecule inhibitors on the oncoprotein c-Myc. J Am Chem Soc. 2009 Jun 3;131(21):7390-401. PubMed PMID: 19432426. Epub 2009/05/13. eng.
140. Wang H, Hammoudeh DI, Follis AV, Reese BE, Lazo JS, Metallo SJ, et al. Improved low molecular weight Myc-Max inhibitors. Mol Cancer Ther. 2007 Sep;6(9):2399-408. PubMed PMID: 17876039.
141. Marasco D, Scognamiglio PL. Identification of inhibitors of biological interactions involving intrinsically disordered proteins. International journal of molecular sciences. 2015;16(4):7394-412. PubMed PMID: 25849651. PMCID: 4425024.
142. Chen CYC, Tou WL. How to design a drug for the disordered proteins. Drug discovery today. 2013 Oct;18(19-20):910-5. PubMed PMID: WOS:000325510200002. English.