

Intrinsically Unstructured Proteins: Potential Targets for Drug Discovery

Pathan Salma, Chintan Chhatbar and Sriram Seshadri
Department of Biochemistry and Biotechnology, Institute of Science,
Nirma University, Sarkhej Gandhinagar Highway, Ahmedabad Gujarat-382481

Abstract: Problem statement: The function of a protein is dependent on its three-dimensional structure. However, numerous proteins lacking intrinsic globular 3D structure under physiological conditions had been recognized. These proteins are frequently involved in some of the most critical cellular control mechanisms and it appears that their rapid turnover, aided by their unstructured nature in the unbound state, provides a level of control that allows rapid and accurate responses of the cell to changing environmental conditions. **Approach:** A significant number of proteins known to be involved in protein deposition disorders were now considered to be Intrinsically Unstructured Proteins (IUPs). For example, A β peptide and tau protein in Alzheimer's disease, PrP in Prion's disease and α -Synuclein in Parkinson's disease. The disorder of intrinsically unstructured proteins (IUP's) was crucial to their functions. They may adopt defined but extended structures when bound to cognate ligands. Their amino acid compositions were less hydrophobic than those of soluble proteins. They lack hydrophobic cores and hence did not become insoluble when heated. About 40% of eukaryotic proteins had at least one long (>50 residues) disordered region. Roughly 10% of proteins in various genomes had been predicted to be fully disordered. Presently over 100 IUP's had been identified; none are enzymes. Obviously, IUP's were greatly underrepresented in the Protein Data Bank, although there were few cases of an IUP bound to a folded (intrinsically structured) protein. **Results:** The five functional categories for intrinsically unstructured proteins and domains were entropic chains (bristles to ensure spacing, springs, flexible spacers/linkers), effectors (inhibitors and disassemblers), scavengers, assemblers and display sites. These IUPs could serve as potential targets for Structure Based Drug Design (SBDD) which stress on the transition from disordered to ordered confirmation through drug stimulation. Recently an unstructured domain of a regulatory protein had been found to be involved in inhibiting catalytic activity of insulin receptor and targeting this IUP would provide a new approach which can be employed in modifying insulin signaling in treatment of diabetes. IUPs were also involved diseases and disorders such as cardio vascular diseases, cancers and autoimmune disorders. Unstructured proteins had also been shown to be important components of invasion, survival and disguising strategies of pathogens such as *Plasmodium falciparum*. **Conclusion:** New greater focuses on proteins that were in some way unstructured normally would promise to provide a greater understanding of protein function particularly with respect to protein-protein interactions and hence can give new potential targets for future strategies.

Key words: Unstructured proteins, MoRFs, structural transitions, drug targets

INTRODUCTION

Until recently, folded three dimensional structure of protein is necessary for the functionality of the protein was considered to be the classical paradigm of structural biology. However, abundance of unfolded or coiled proteins in nature, mainly in eukaryotes, has led to the discovery of a new group of proteins which are not just structurally unfolded but their functionally is

found to be dependent on their unstructured confirmations. In many of these proteins the unstructured confirmation extends throughout the protein and in some it is partially present, while in others there are long unstructured segments in otherwise ordered folded proteins^[1,2]. These unstructured proteins play essential roles in cell cycle control, transcriptional and translational regulation and modulation of activity and/or assembly of other

Corresponding Author: Sriram Seshadri, Department of Biochemistry and Biotechnology, Institute of Science, Nirma University, Sarkhej Gandhinagar Highway, Ahmedabad Gujarat-382481
Tel: +91-2717-241901/ext 627 Fax: +91-2717-241916

proteins, signal transduction and even regulation of nerve cell functions^[3,4]. Numbers of proteins involved in protein deposition disorders are now considered to belong to the family of unstructured proteins. For example, A β peptide and tau protein in Alzheimer's disease, PrP in Prion's disease and α -Synuclein in Parkinson's disease^[3]. This observed intrinsic disorder has many important implications on the functional and regulatory behavior of IUPs. IUPs show high specificity with low affinity, rapid turnover rate and the ability to overcome barriers such as steric hindrance that arises due to larger surface area and rigid structure of many signaling molecules and thermodynamic energy transfer during molecular interactions^[5]. As depicted by Protein Quartet Model, proteins exist in four different confirmations: ordered forms, molten globules, pre-molten globules and random coils; each one having distinct state of activity and regulation^[4]. However, the ultimate functioning (activity) of protein is dependent on transitions between these structures. It has been confirmed that the unstructured proteins undergo structural transitions between folded and unfolded confirmations whenever they interact with their binding partner such as modulators, target molecules, specific ions or functional regulators^[2]. These transitions form the basis of their functional regulation that could be controlled by introducing their modifiers as drugs. The IUPs show interesting features such as modulation in their functional behavior through posttranscriptional and posttranslational modifications of their corresponding mRNAs, multiple interactions which have low affinity but high specificity and most importantly, their regulation through coupled binding and folding transitions^[6].

Here, based on all the above features of IUPs, we present the hypothesis that disease associated proteins can be targeted for structural transition by using structure based drugs that mimic the binding partner of targeted IUP and induce modulation in structure and behavior of targeted IUP. Hence it may be possible to alter the folding of target protein to regulate its activity and ultimately its function.

Disorders in protein structure: Using data of many of the direct and indirect approaches such as X-ray crystallography, multidimensional Nuclear Magnetic Resonance (NMR), Circular Dichroism (CD) spectroscopy, proteolytic sensitivity and heat stability has led to the identification of more than 200 proteins and protein domains which lack native folded 3-Dimensional structure^[7]. In addition to this, the Swiss Protein Data Bank has predicted that more than 15,000 proteins may contain disordered regions of at least

40 consecutive amino acid residues, with more than 1050 of them having high scores indicating disorder^[8,9]. This observation was accomplished by primary sequence analysis of proteins, which helped to conclude that "a large portion of gene sequences appear to code not for folded, globular proteins, but for long patches of amino acids which tend to be either unfolded in solution or adopt non-globular structures of unstructured conformation".

It is known that the sequence of amino acid encoded by gene sequence is responsible for the ultimate stable 3D structure of proteins. Apart from amino acid sequence, many things such as charge, bulkiness and hydrophathy index of amino acids, their interactions with each other etcetera are involved in deciding the stable and functional three dimensional confirmation of any protein. Similar to the encoded folded structure of any proteins, it can be hypothesized that the unstructured or unfolded characteristic of a given protein is also encoded by amino acid sequence and hence by underlying genetic sequence^[10]. Certain amino acid residues have been found to be highly "order-promoting" (namely cysteine, tryptophan, tyrosine, isoleucine, phenylalanine, valine, leucine, histidine, threonine and asparagine) while others are highly "disorder-promoting" (namely aspartic acid, methionine, lysine, arginine, serine, glutamine, proline and glutamic acid) (Fig. 1)^[11-14]. These order, disorder promoting amino acids can be shown by Fig. 1, in which the relative amino acid compositions of intrinsically disordered regions is according to their availability in the DisProt database^[7,15] in comparison with a set of structured (or ordered) proteins (Fig. 2)^[12]. In this case, these amino acid compositions were compared by means of a profiling approach^[16,17].

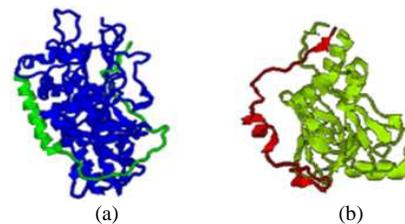


Fig. 1: Some IUPs wrap around their partner upon binding. For some IUPs, the structure in the complexed state is known from X-ray crystallography or NMR. The structures (PDB code in parenthesis) shown are: (a): SNAP-25 bound to BoNT/A (1XTG); (b): SARA SBD domain bound to Smad2 MH2 domain (1DEV), the structures have been visualized by the RasMol^[11]

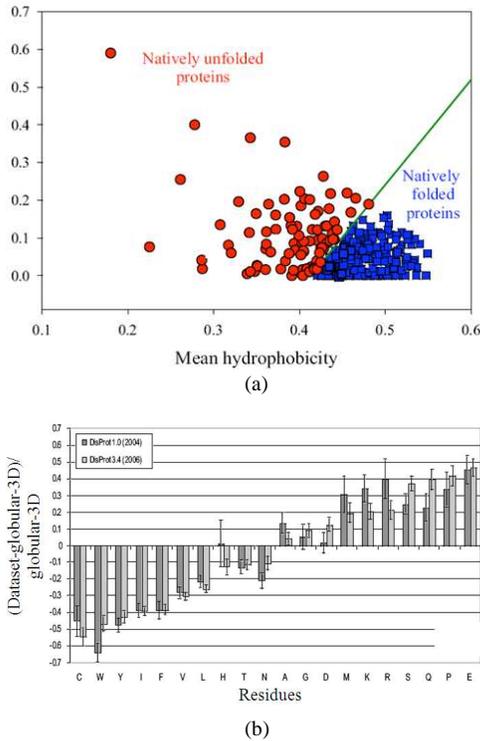


Fig. 2: Peculiarities of the amino acid sequences of intrinsically disordered proteins (a): Mean net charge versus mean hydrophobicity plot (charge-hydrophobicity plot) for the set of 275 folded (blue squares) and 91 natively unfolded proteins (red circles)^[18]; (b): Amino-acid composition, relative to the set of globular proteins Globular-3D, of intrinsically disordered regions 10 residues or longer from the DisProt database. Dark gray indicates DisProt 1.0 (152 proteins), whereas light gray indicates DisProt 3.4 (460 proteins). Amino acid compositions were calculated per disordered regions and then averaged. The arrangement of the amino acids is by peak height for the DisProt 3.4 release. Confidence intervals were estimated using per-protein bootstrapping with 10,000 iterations^[12]

Unstructured proteins: Structure and functions:

Structural transitions and functioning: The existence of native cellular proteins can be explained by the Protein Trinity model, which suggests three structural confirmations; ordered, molten globule and random coil for any given protein in its natural conditions in cell. It has been suggested that the absolute function of the protein occurs due to its transitions between these structures (Fig. 3)^[4]. Experimental results on the conformational behavior of intrinsically unstructured

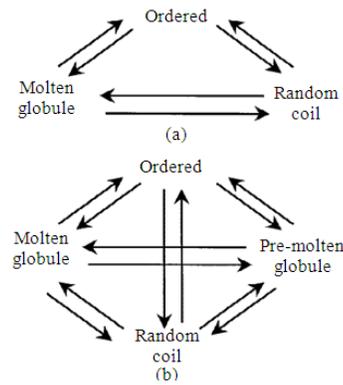


Fig. 3: Extension of The Protein Trinity (a): To the Protein Quartet model of protein functioning; (b): In accordance with this model, function arises from four specific conformations of the polypeptide chain (ordered forms, molten globules, pre-molten globules and random coils) and transitions between any of the states^[10]

proteins indicate that these proteins did not possess uniform structural properties and their states may be designated as intrinsic coils and intrinsic pre-molten globules. Evidences have been reported that the Protein Trinity Model can now be extended to Protein Quartet model because of existence of four different structural confirmations depicted as ordered forms, molten globules, pre-molten globules and random coils. The function of unstructured protein can be fulfilled by their rapid fluctuation among alternative states such as coil to pre-molten globule transitions, coil to molten globule transitions, coil to rigid structure transitions, pre-molten globule to molten globule transitions, intrinsic pre-molten globules to rigid conformation^[5].

Ligand induced folding: In many instances it has been shown that ultimate structural transitions of any IUPs arise due to its interaction with binding partners that can be a modulator, regulator or any other protein as ligand. For example, partial folding in IUPs such as thymosin $\alpha 1$ ^[19] and prothymosin α induced by Zn^{++} ^[20], partial folding of human α -synuclein in presence of several divalent and trivalent metal ions^[21], lipid induced transformation of water-soluble form of myelin basic protein into the molten globule-like conformation^[22], folding of RNase P from *B. subtilis* (unfolded in 10 mM sodium cacodylate at neutral pH) into a native α/β structure upon addition of various small molecular anions^[23]. Other example of induced folding is shown in Fig. 1 exemplifying the behavior of IUPs.

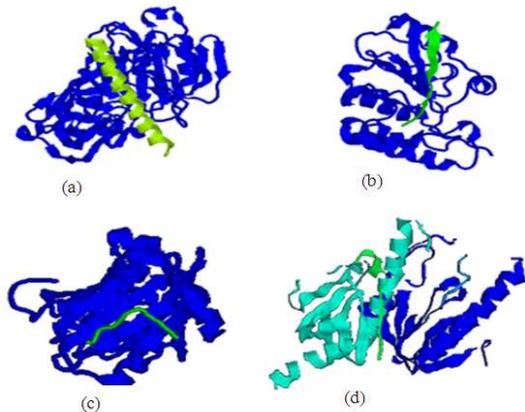


Fig. 4: Examples of structurally divergent MoRFs: (a): An α -MoRF, Proteinase Inhibitor IA3, bound to Proteinase A (1DP5); (b): A β -MoRF, viral protein pVIc, bound to Human Adenovirus 2 Proteinase (AVP); (c): An τ -MoRF, Amphiphysin, bound to α -adaptin C (1KY7); (d): A complex-MoRF, β -amyloid precursor protein (β APP), bound to the PTB domain of the neuron specific protein X11 (1X11). Partner interfaces (gray surface) are also indicated. All MoRFs are shown in green. The structures have been visualized by the RasMol and PDB code is shown in parenthesis

Molecular recognition in IUPs: Though structural data shows lack of secondary and tertiary structure, lack of understanding of atypical structural and functional attributes of IUPs would raise many questions against the premature idea of complete lack of order. There are many well defined binding sites existing in IUPs that allow them to undergo transitions upon interactions with their binding partners. A specialized subset of these interacting domains have been recognized as ‘Molecular Recognition Elements’ or ‘Molecular Recognition Features’ which are protein regions that specifically participate in protein-protein interactions. These MoRFs have the ability to undergo significant induced folding steps or disorder-to-order transition^[1,3,18] together with the change in the structure of their binding partners. Such molecular recognition mechanism, which is coupled to the folding process, has been noted to confer exceptionally high specificity and low affinity; binding diversity and binding commonality to their IUPs^[10].

According to current understanding depending on their structures in the bound state MoRFs can be divided into three subtypes: α -MoRFs form α -helices, β -MoRFs form β -strands and τ -MoRFs form structures without a regular pattern of backbone hydrogen bonds (Fig. 4), along with their observed mixtures^[24,25].

Although only a few MoRFs have been studied experimentally, it has been suggested that all MoRFs are intrinsically disordered in the absence of their binding partners in line with observations of Gunasekaran *et al.*^[26] who have demonstrated that intrinsic disorder in the unbound state is reflected in the structures of the bound state through relatively large surface and interface areas. With the discovery of MoRFs it can be supposed that IUPs function by molecular recognition which may be transient or permanent binding to a binding partner. In general, for IUPs to assume a folded confirmation, prior to binding to modulators or targets, is difficult as a consequence of many topological/structural constraints. Binding of target molecules and/or modulators to the IUP increases its flexibility in terms of topological stress to reach to the final functional state aided by MoRFs. Numbers of IUPs have been found to bind to many different types of partners and vice versa a molecule of modulator can bind with many different types of IUPs. Interestingly, both these conditions result in different confirmations of IUPs with different combination of an IUP and its binding partner. This flexibility and low affinity with high specificity of IUPs for their binding partners, which actually is an attribute of MoRFs, increases the probability of multiple interactions that allow control of many processes that have common checkpoints.

The structural flexibility that occurs due to molecular recognition has actually been demonstrated for the C-terminal domain of DNA dependent RNA polymerase II (RNAP II) bound to either RNA guanylyl transferase Cgt1 or peptidyl-proline isomerase Pin1^[27] and the HIF-1 α -interaction domain bound to either the TAZ1 domain of cAMP response element binding protein (CREB)-binding protein (CBP)^[28] or the asparagine hydroxylase FIH^[29]. Thus, the function of IUPs arises either from their transitions between alternative states as described previously or through molecular recognition and induced folding.

Unstructured proteins in diseases: The D² concept:

Proteins are essential for functioning of any cell because of their involvement in each and every step of cellular activity such as metabolism and its regulation, signaling, transport, defense and many more. Hence it is not surprising that number of diseased conditions arise due to failure of a specific peptide or protein to adopt its proper structure. Such diseases are associated with protein misfolding which include protein aggregation (and/or fibril formation), loss of normal function and gain of toxic function. Some proteins assume pathologic state due to some of the endogenous factors such as chaperones, intracellular or extra

cellular matrices, other proteins and small molecules, which can alter the confirmation of a pathogenic protein and thereby increase its chances of getting misfolded. Increase in rate of pathogenic misfolded protein accumulation may increase because of ageing, mutation, or other induced conditions leading to disease progression.

Some of the proteins shown to be reason for human diseases such as cancer, Parkinson's disease and other synucleinopathies, Alzheimer's, prion diseases, diabetes and cardiovascular disease are either completely disordered or contain long disordered regions. In fact, the analysis on Swiss-Prot suggests strong association of intrinsic disorder and diseases such as malaria, trypanosomiasis, Human Immunodeficiency Virus (HIV) and acquired immunodeficiency syndrome (AIDS), deafness, obesity, cardiovascular disease, diabetes mellitus, albinism and prion (Fig. 5)^[30]. Thus, intrinsic disorder is very common in disease-associated proteins that have given rise to a disorder in disordered proteins concept, which is now calling the "D2 concept"^[6].

Drug discovery based on protein structure: Proteins are highly favored targets for future drug discoveries for protein associated diseases since long. However, literature surveys have failed to give example of currently used structure based drugs targeted towards proteins. There has been only a little success in finding drugs which act by blocking protein-protein interaction^[10,31].

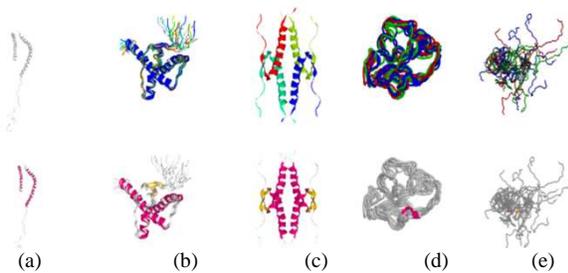


Fig. 5: Disease associated unstructured proteins: (a): α -Synuclein; (b): Major prion protein; (c): p53 tumor antigen; (d): Tat protein from HIV; (e) Apical membrane antigen 1 (AMA1) of the malarial parasite *Plasmodium falciparum* upper row shows the structures of proteins according to their chains and the lower row shows the structures according to the confirmation^[30]

Many studies have revealed several important features of protein-protein, protein-modulator and protein-binding partner interactions which can serve to determine drug targets and potential ligands. Further, all these interactions can be extended for those which comprise one of the partners having an unstructured region else, they themselves are wholly unstructured protein. Moreover there are many features of unstructured proteins that make them suitable drug targets. Firstly, their structural transitions which make them prone to induced folding which in turn regulate their functionality. Secondly, MoRFs exhibit some special features that make them stabilize the transient forms. For example, binding with partner promotes organization of disordered region into a helix or other structure with hydrophobic side chains that project away from the backbone and into the cleft^[24]. With respect to this the p53/Mdm2 interaction can be considered here, in which the p53 binding site is predicted to be an α -MoRF and this binding site contains hydrophobic side chains that project deeply into the cleft located on the surface of the Mdm2 partner upon binding. This in turn provides utilization of disorder to order transition with spending of energy which results in decrease of entropy. The interaction between p53/Mdm2 can be easily blocked and/or targeted by small structure based drugs. Some of the examples discussed in recent reviews depend on one structured partner and one disorder partner that undergo helix formation upon binding MoRF and its partner^[31]. Given this, IUPs involved in some very important diseases can be targeted for drug discoveries.

The preS1 surface antigen of Hepatitis B Virus (HBV) is known to play important role in the initial attachment of HBV to hepatocytes and is natively unstructured protein. The N-terminal 50 residues of preS1 that are populated with multiple pre-structured motifs (MoRFs), contribute critically to hepatocyte binding. There are some overlapping pre-structured motifs identified in preS1 that show folding upon binding to monoclonal antibodies and should help to determine regions that can be used to design structure-based inhibitors against HBV attachment to hepatocytes once an HBV receptor is identified^[32].

α -Synuclein is a small (14 kDa) highly conserved protein and is intrinsically unstructured, i.e., natively unfolded^[33]. Upon interaction with intracellular molecules such as phospholipids, α -Synuclein undergoes conformational change from an unstructured monomer in solution to organized structure related to β conformation^[30]. This ultimate organized structure forms the basis of aggregation and fibrillation, along with an intermediate that is found in Lewy bodies^[34].

However, the existence of such an intermediate between folded and unfolded conformations, on the pathway to fibrils, lead to population of the intermediate that shifts the equilibrium position from the natively unfolded state to a partially folded intermediate through any intracellular factors increasing the likelihood of α -synuclein fibril formation and development of Parkinson's disease. Such factors could include relatively nonpolar molecules that would preferentially bind to the intermediate. Similarly, the tau protein exists as unstructured protein with little α -helical or β -sheet structure^[35] that is modulated by heparin binding which in turn is an IUP^[30]. However, the association of tau protein with Alzheimer's disease shows the aggregation consisting of PHF (parallel helix filaments) in β amyloid fibrils mediating intracellular inclusions and neurodegeneration suggesting transition of an unfolded structure into an ordered state^[36]. In fact the abnormal phosphorylation observed in tau protein aggregates are at the unstructured sites of the protein^[37]. This finding can be linked to an assumption that hyperphosphorylation might be involved in the observed structural transition of tau protein causing abnormal microtubule association and hence bringing neurological disorder. In support of such a connection, it can be said that those molecules that can preferentially bind to α -synuclein or tau may decrease or prevent the rate of fibril formation and even prevent accumulation of partially folded intermediate or aggregates.

The central event in the pathogenesis of prion diseases is a major conformational change in the prion protein (PrP) normal cellular form containing a preponderance of helical secondary structure to a plaque forming conformer containing a greater proportion of β -sheet. Structure determination of fragments of the prion protein revealed that around 100 residues at the C terminus are folded into a largely helical domain with completely unfolded N-terminal^[38]. Partial folding of a local region is shown to be assisted by presence Cu(II) ions^[39] and may give a clue to the overall physiological function of the prion protein, which at present is unknown. If this protein functions as a copper storage or transport protein, the extreme flexibility of the N terminus is probably of functional significance that can be aimed for regulating transition of prion protein.

The presence of disorder has been directly observed in many cancer-associated proteins, a few examples of which include p53^[40], p57kip2^[41], Bcl-XL and Bcl-2^[42], c-Fos^[43], proto-oncogene securin^[44] and a breast cancer associated protein BRCA1^[45] along with recently discovered E6 and E7 oncoproteins from the high-risk types of Human Papillomaviruses (HPVs)^[46]. Amongst

these Bcl-2 has been characterized as drug targets. Additionally, small-molecule antagonists have recently been described for several new targets, including Rac1-Tiam1, beta-catenin-T cell factor (Tcf) and Sur-2-ESX that are involved in cancer progression^[31].

Other unstructured proteins that are shown to be associated with diseases and are potential targets for structure based drug design include: (A) Insulin-like Growth Factor Binding Proteins (IGFBPs) that are carriers and regulators of the insulin like growth factors having C-terminal domain with three highly disordered loops^[47]. (B) Two intracellular domains of connexin43 (Cx43), cytoplasmic loop 95-144 and C-terminal domain (amino acids 254-382) possessing short transient R-helices with numerous binding partners identified; these include tubulin, v-Src, c-Src, ZO-1, Casein Kinase 1 (CK1), Mitogen-Activated Protein Kinase (MAPK), cGMP-dependent protein kinase, cAMP-dependent protein kinase and protein kinase C^[48]. (C) Numerous proteins associate with obesity and CVD have been identified to have long regions of intrinsic disorder containing R-helices and α -MoRFs, predicted in 101 proteins from CVD data set^[49]. (D) Apical Membrane Antigen 1 (AMA1) of the malarial parasite *P. falciparum* is a merozoite antigen, has well-defined, disulfide-stabilized core region separated by a disordered loop and both the N-and C-terminal regions of the molecule are unstructured^[50]. (E) Tat (transactivator of transcription) is a small RNA-binding protein that helps in HIV-1 replication is an IUP and shows induced folding when it interacts with its binding partner-cyclin T1 along with its ability to interact with wide variety of proteins^[51].

CONCLUSION

Altogether, it can be said that the discovery of small drug molecules that can target protein-protein interaction is difficult even for structured proteins. Consequently there are some important prospects that should be considered while targeting interactions involving unstructured proteins with molecular drugs. For example, affinity of molecule to the protein (frequently subnanomolar) that can successfully compete for binding to its site^[52], the biology of the system and understanding molecular recognition of protein surfaces.

REFERENCES

1. Tompa, P., 2002. Intrinsically unstructured proteins. *Trends Biochem. Sci.*, 27: 527-533. PMID: 12368089

2. Dyson, H.J. and P.E. Wright, 2005. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.*, 6: 197-208. PMID: 15738986
3. Wright, P.E. and H.J. Dyson, 1999. Intrinsically unstructured proteins: Reassessing the protein structure-function paradigm. *J. Mol. Biol.*, 293: 321-331. PMID: 10550212
4. Dunker, A.K., J.D. Lawson, C.J. Brown, R.M. Williams and P. Romero *et al.*, 2001. Intrinsically disordered proteins. *J. Mol. Graph. Model.*, 19: 26-59. PMID: 11381529
5. Uversky, V.N., 2002. Natively unfolded proteins: A point where biology waits for physics. *Protein Sci.*, 11: 739-756. PMID: 11910019
6. Uversky, V.N., C.J. Oldfield and A.K. Dunker, 2008. Intrinsically disordered proteins in human diseases: Introducing the D2 concept. *Ann. Rev. Biophys. Mol. Biol.*, 37: 215-246. PMID: 18573080
7. Vucetic, S., Z. Obradovic, V. Vacic, P. Radivojac and J.G. Sikes *et al.*, 2005. DisProt: A database of protein disorder. *Bioinformatics*, 21: 137-140. PMID: 15310560
8. Dunker, A.K., E. Garner, S. Guillot, P. Romero and K. Albercht *et al.*, 1998. Protein disorder and the evolution of molecular recognition: Theory, predictions and observations. *Pac. Symp. Biocomput.*, 3: 473-484. PMID: 9697205
9. Romero, P., Z. Obradovic, C. Kissinger, J.E. Villafranca and E. Garner *et al.*, 1998. Thousands of proteins likely to have long disordered regions. *Pac. Symp. Biocomput.*, 3: 437-448. PMID: 9697202
10. Dunker, A.K., C.J. Oldfield, J. Meng, P. Romero and J.Y. Yang *et al.*, 2008. The unfoldomics decade: An update on intrinsically disordered proteins. *BMC. Genom.*, 9: S1. PMID: 18831774
11. Tompa, P., 2005. The interplay between structure and function in intrinsically unstructured proteins. *FEBS. Lett.*, 579: 3346-3354. PMID: 15943980
12. Radivojac, P., L.M. Iakoucheva, C.J. Oldfield and Z. Obradovic *et al.*, 2007. Intrinsic disorder and functional proteomics. *Biophys. J.*, 92: 1439-1456. PMID: 17158572
13. Williams, R.M., Z. Obradovi, V. Mathura, W. Braun and E.C. Garner *et al.*, 2001. The protein non-folding problem: Amino acid determinants of intrinsic order and disorder. *Pac. Symp. Biocomput*, 6: 89-100. PMID: 11262981
14. Campen, A., R.M. Williams, C.J. Brown, V.N. Uversky and A.K. Dunker, 2008. TOP-IDP-Scale: A new amino acid scale measuring propensity for intrinsic disorder. *Protein Peptide Lett.*, 15: 956-963. PMID: 18991772
15. Sickmeier, M., J.A. Hamilton, T. LeGall, V. Vacic and V.N. Uversky *et al.*, 2007. DisProt: The database of disordered proteins. *Nucleic Acids Res.*, 35: D786-D793. DOI: 10.1093/nar/gkl893
16. Dunker, A.K., J.D. Lawson, C.J. Brown, R.M. Williams and K.W. Hipps *et al.*, 2001. Intrinsically disordered protein. *J. Mol. Graph. Model.*, 19: 26-59. PMID: 11381529
17. Vacic, V., V.N. Uversky, A.K. Dunker and S. Lonardi, 2007. Composition profiler: A tool for discovery and visualization of amino acid composition differences. *BMC. Bioinform.*, 8: 211. PMID: 17578581
18. Uversky, V.N., J.R. Gillespie and A.L. Fink, 2000. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins*, 41: 415-427. PMID: 11025552
19. Grottesi, A., M. Sette, T. Palamara, G. Rotilio, E. Garaci and M. Paci, 1998. The conformation of peptide thymosin alpha 1 in solution and in a membrane-like environment by circular dichroism and NMR spectroscopy: A possible model for its interaction with the lymphocyte membrane. *Peptides*, 19: 1731-1738. PMID: 9880079
20. Uversky, V.N., J.R. Gillespie, I.S. Millett, A.V. Khodyakova and V.M. Abramov *et al.*, 2000. Zn(2+)-mediated structure formation and compaction of the "natively unfolded" human prothymosin. *Biochem. Biophys. Res. Commun.*, 267: 663-668. PMID: 10631119
21. Uversky, V.N., J. Li and A.L. Fink, 2001. Metal-triggered structural transformations, aggregation and fibrillation of human alpha-synuclein: A possible molecular link between Parkinson's disease and heavy metal exposure. *J. Biol. Chem.*, 276: 44284-44296. PMID: 11553618
22. Polverini, E., A. Fasano, F. Zito, P. Riccio and P. Cavatorta, 1999. Conformation of bovine myelin basic protein purified with bound lipids. *Eur. Biophys. J.*, 28: 351-355. PMID: 10394626
23. Henkels, C.H., J.C. Kurz, C.A. Fierke and T.G. Oas, 2001. Linked folding and anion binding of the *Bacillus subtilis* ribonuclease P protein. *Biochemistry*, 40: 2777-2789. PMID: 11258888
24. Mohan, A., C.J. Oldfield, P. Radivojac, V. Vacic and V.N. Uversky *et al.*, 2006. Analysis of Molecular Recognition Features (MoRFs). *J. Mol. Biol.*, 362: 1043-1059. PMID: 16935303
25. Vacic, V., C.J. Oldfield, A. Mohan, P. Radivojac and A.K. Dunker *et al.*, 2007. Characterization of molecular recognition features, MoRFs and their binding partners. *J. Proteome Res.*, 6: 2351-2366. PMID: 17488107

26. Gunasekaran, K., C.J. Tsai and R. Nussinov, 2004. Analysis of ordered and disordered protein complexes reveals structural features discriminating between stable and unstable monomers. *J. Mol. Biol.*, 341: 1327-1341. PMID: 15321724
27. Fabrega, C., V. Shen, S. Shuman and C.D. Lima, 2003. Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Mol. Cell*, 11: 1549-1561. PMID: 12820968
28. Dames, S.A., M. Martinez-Yamout, R.N. Guzman, H.J. Dyson and P.E. Wright, 2002. Structural basis for Hif-1 alpha/CBP recognition in the cellular hypoxic response. *Proc. Natl. Acad. Sci. USA.*, 99: 5271-5276. PMID: 11959977
29. Elkins, J.M., K.S. Hewitson, L.A. McNeill and J.F. Seibel *et al.*, 2003. Structure of factor-inhibiting Hypoxia-Inducible Factor (HIF) reveals mechanism of oxidative modification of HIF-1 alpha. *J. Biol. Chem.*, 278: 1802-1806. PMID: 12446723
30. Xie, H., S. Vucetic, L.M. Iakoucheva, C.J. Oldfield, A.K. Dunker, Z. Obradovic and V.N. Uversky, 2007. Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications and diseases associated with intrinsically disordered proteins. *J. Proteome Res.*, 6: 1917-1932. PMID: 17391016
31. Arkin, M., 2005. Protein-protein interactions and cancer: Small molecules going in for the kill. *Curr. Opin. Chem. Biol.*, 9: 317-324. PMID: 15939335
32. Chi, S.W., D.H. Kim, S.H. Lee, I. Chang and K.H. Han, 2007. Pre-structured motifs in the natively unstructured preS1 surface antigen of hepatitis B virus. *Protein Sci.*, 16: 2108-2117. PMID: 17766372
33. Qin, Z., D. Hu, S. Han, S.H. Reaney, D.A. Di-Monte and A.L. Fink, 2007. Effect of 4-Hydroxy-2-nonenal modification on alpha-synuclein aggregation. *J. Biol. Chem.*, 282: 5862-5870. PMID: 17189262
34. Vladimir N. Uversky, Jie Li and Anthony L. Fink, 2001. Evidence for a partially folded intermediate in alpha-synuclein fibril formation. *J. Biol. Chem.*, 276: 10737-10744. PMID: 11152691
35. Von Bergen, M., S. Barghorn, J. Biernat, E.M. Mandelkow and E. Mandelkow, 2005. Tau aggregation is driven by a transition from random coil to beta sheet structure. *Biochem. Biophys. Acta*, 1739: 158-166. PMID: 15615635
36. Schweers, O., E. Schonbrunn-Hanebeck, A. Marx and E. Mandelkow, 1994. Structural studies of tau protein and Alzheimer paired helical filaments show no evidence for beta-structure. *J. Biol. Chem.*, 269: 24290-24297. PMID: 7929085
37. Orosz, F., G.G. Kovács, A. Lehotzky, J. Oláh, O. Vincze and J. Ovádi, 2004. TPPP/p25: From unfolded protein to misfolding disease: Prediction and experiments. *Biol. Cell*, 96: 701-711. PMID: 15567525
38. Donne, D.G., J.H. Viles, D. Groth, I. Mehlhorn, T.L. James, F.E. Cohen, S.B. Prusiner, P.E. Wright and H. J. Dyson, 1997. Structure of the recombinant full-length hamster prion protein PrP (29-328): The N terminus is highly flexible. *Proc. Natl. Acad. Sci. USA.*, 94: 13452-13457. PMID: 9391046
39. Viles, J.H., F.E. Cohen, S.B. Prusiner, D.B. Goodin, P.E. Wright and H.J. Dyson, 1999. Copper binding to the prion protein: Structural implications of four identical cooperative binding sites. *Proc. Natl. Acad. Sci. USA.*, 96: 2042-2047. PMID: 10051591
40. Lee, H., K.H. Mok, R. Muhandiram, K.H. Park and J.E. Suk *et al.*, 2000. Local structural elements in the mostly unstructured transcriptional activation domain of human p53. *J. Biol. Chem.*, 275: 29426-29432. PMID: 10884388
41. Adkins, J.N. and K.J. Lumb, 2002. Intrinsic structural disorder and sequence features of the cell cycle inhibitor p57Kip2. *Proteins*, 4: 1-7. PMID: 11746698
42. Chang, B.S., A.J. Minn, S.W. Muchmore, S.W. Fesik and C.B. Thompson, 1997. Identification of a novel regulatory domain in Bcl-X(L) and Bcl-2. *EMBO. J.*, 16: 968-977. PMID: 9118958
43. Campbell, K.M., A.R. Terrell, P.J. Laybourn and K.J. Lumb, 2000. Intrinsic structural disorder of the C-terminal activation domain from the bZIP transcription factor Fos. *Biochemistry*, 39: 2708-2713. PMID: 10704222
44. Sanchez-Puig, N., D.B. Veprintsev and A.R. Fersht, 2005. Human full length securin is a natively unfolded protein. *Protein Sci.*, 14: 1410-1418. PMID: 15929994
45. Mark, W.Y., J.C. Liao, Y. Lu, A. Ayed and R. Laister *et al.*, 2005. Characterization of segments from the central region of BRCA1: An intrinsically disordered scaffold for multiple protein-protein and protein-DNA interactions. *J. Mol. Biol.*, 345: 275-287. PMID: 15571721
46. Uversky, V.N., A. Roman, C.J. Oldfield and A.K. Dunker, 2006. Protein intrinsic disorder and human papillomaviruses: Increased amount of disorder in E6 and E7 oncoproteins from high risk HPVs. *J. Proteome Res.*, 5: 1829-1842. PMID: 16889404

47. Headey, S.J., D.W. Keizer, S. Yao, G. Brasier and P. Kantharidis *et al.*, 2004. C-terminal domain of Insulin-like Growth Factor (IGF) binding protein-6: structure and interaction with IGF-II. *Mol. Endocrinol.*, 18: 2740-2750. PMID: 15308688
48. Giepmans, B.N., 2004 Gap junctions and connexin-interacting proteins. *Cardiovasc. Res.*, 62: 233-245. PMID: 15094344
49. Cheng, Y., T. Le Gall, C.J. Oldfield, A.K. Dunker and V.N. Uversky, 2006. Abundance of intrinsic disorder in proteins associated with cardiovascular disease. *Biochemistry*, 45: 10448-10460. PMID: 16939197
50. Nair, M., M.G. Hinds, A.M. Coley, A.N. Hodder and M. Foley *et al.*, 2002. Structure of domain III of the blood-stage malaria vaccine candidate, *Plasmodium falciparum* Apical Membrane Antigen 1 (AMA1). *J. Mol. Biol.*, 322: 741-753. PMID: 12270711
51. Shojania, S. and J.D. O'Neil, 2006. HIV-1 Tat is a natively unfolded protein: The solution conformation and dynamics of reduced HIV-1 tat-(1-72) by NMR spectroscopy. *J. Biol. Chem.*, 281: 8347-8356. PMID: 16423825
52. Cochran, A.G., 2000. Antagonists of protein-protein interactions. *Chem. Biol.*, 7: 85-94. PMID: 10779412