Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Structural Biology

Unstructural biology coming of age

Peter Tompa[1,2]

It is now generally accepted that many proteins or protein domains (intrinsically disordered proteins, IDPs) lack a well-defined tertiary structure under functional conditions. Due to recent concerted activity, a critical transition in this field is gaining momentum, in which qualitative observations are turned into quantitative structural models of IDPs. Here, it is suggested that the transition is set up by the synergy of: (i) more advanced bioinformatic tools for the prediction of disorder and function of IDPs, (ii) ensemble description of their structure and dynamics in both free and bound states, down to the single molecule level, (iii) advent of in-cell approaches for characterizing their structure and function *in vivo*, and (iv) generation of small-molecule inhibitors both against their binding partners and IDPs themselves. In all, we suggest that due to steady advance in these areas, the field of 'unstructural' biology is rapidly maturing to a state where it can provide quantitative models of proteins functioning without well-defined three-dimensional structures.

Addresses
[1] Institute of Enzymology, Biological Research Center, Hungarian Academy of Sciences, Budapest, Hungary
[2] 2VIB Department of Structural Biology, Brussels, Belgium

Corresponding author: Tompa, Peter (tompa@enzim.hu)

Current Opinion in Structural Biology 2011, **21**:1–7

This review comes from a themed issue on
Sequences and topology
Edited by Julian Gough and Keith Dunker

0959-440X/$ – see front matter
© 2011 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.sbi.2011.03.012

Introduction

The second half of the last century witnessed the continuous advance of structural biology, glorifying the notion that structure is the prerequisite of function. Arguably, the recent most exciting development in structural biology, however, is not the solution of yet another giant protein or complex, but the recognition that many proteins or regions of proteins exist and function without a well-defined structure. These IDPs demand a radical change in concept for describing biological events at the molecular level. Whereas this field has raised eyebrows for a decade, it is now becoming evident that a critical transition is taking place by 'unstructural' biology getting into the mainstream of molecular biology. Rapid growth of the field has been marked by several excellent recent reviews [1–5], including a textbook of comprehensive coverage [6]. Here, the most important recent developments of the field are surveyed.

Computational studies: prediction of disorder and functional sites

Bioinformatic predictions still play a decisive role in studies of structural disorder. Whereas the latest release (2010 November) of the DisProt database [7] contains 1342 disordered regions in 627 IDPs (www.disprot.org), there is still a very wide gap between experimentally demonstrated and expected structural disorder, which leaves much room for bioinformatics in large-scale functional association studies. Based on the compositional bias of IDPs, several dozen predictors of different principles have been developed [8]. To handle limitations inherent in prediction accuracy due to distinct flavors of disorder, however, different predictors are recently combined into metapredictors, such as metaPrDOS [9] or PONDR-FIT [10]. These combined predictors do show improved performance over their composite ones.

Predicting function and/or functional sites of IDPs is a task even more difficult. Recently, significant advance has been made in this direction, based on the observation that interactions of IDPs are often mediated by short linear motifs [11]. Because linear motifs are 3–15 residues in length, they contain very little sequence information and their prediction from sequence alone is fraught with very high false positive rates. A critical advance in this direction has been made by applying context-filters, which significantly increase prediction accuracy by taking into consideration motif enrichment in proteins that share the same binding partner or evolutionary history (e.g. SLiMFinder [12]). Completely different logic forms the basis of ANCHOR, which predicts disordered binding sites by estimating their interaction energy with a general partner [13], and of the molecular recognition feature predictor (α-MoRF-PredII), which uses patterns of order/disorder prediction [14].

Toward describing the structural ensemble of IDPs

It is becoming evident that IDPs are not fully disordered, but they have all sort of function-related transient short- and long-range structural organization. The major techniques toward describing the ensuing structural ensemble apply structural calculations restrained by NMR and small-angle X-ray scattering (SAXS) data (Table 1). Residue-level parameters carrying information mostly on the local structure of the IDP, such as chemical shifts,

Please cite this article in press as: Tompa P. Unstructural biology coming of age, Curr Opin Struct Biol (2011), doi:10.1016/j.sbi.2011.03.012

Table 1

Methods developed to describe the structural ensemble of IDPs. The methods usually rely on a set of structural constraints determined by NMR residual dipolar coupling (RDC), various chemical shift values, amide proton relaxation rate ($^{15}N\ R_2$), distance restraints from paramagnetic relaxation enhancement (PRE), small-angle X-ray scattering (SAXS, usually Kratky plot), dynamic light scattering (DLS) or analytical ultracentrifugation (AUC). In each method an ensemble of conformations is generated and iterated to match the experimental restraints as closely as possible

Parameters	Ensemble method	Protein	Reference
RDC, SAXS	Flexible-Meccano, Accelerated Molecular Dynamics (AMD)	p53	[58]
RDC	Flexible-Meccano	Sendai virus PX	[59]
Chemical shifts, R_h (DLS), $^{15}N\ R_2$, SAXS	ENSEMBLE	Inhibitor-2, spinophilin, DARPP-32	[34]
PRE, RDC	ASTEROIDS	α-Synuclein	[60]
Chemical shifts, $^{15}N\ R_2$, RDC, PRE, SAXS	ENSEMBLE	Sic1	[16•]
SAXS, $^{15}N\ R_2$,	Ensemble optimization method	Ribosomal L12	[61]
Chemical shifts, R_h (AUC) SAXS	Molecular dynamics	$p27^{Kip1}$	[62]
PRE	Molecular dynamics	α-Synuclein	[15••]

hetNOE values, relaxation parameters and dipolar couplings, are determined in NMR. To obtain long-range structural constraints, spin probes are applied in paramagnetic relaxation enhancement (PRE) measurements. SAXS, on the other hand, mostly contributes information on the hydrodynamic behavior and topology of the polypeptide chain. To interpret the values by either technique, a large number of random conformers are generated, the parameters for each conformer are calculated and an optimization procedure is carried out to select a limited number (around 50) of conformers which together satisfy the constraints (for details and references, see Table 1).

If such an ensemble description of structure can be related to the function of IDP, we may state that 'unstructural' biology approaches the descriptive power we ascribe to traditional structural biology at present. For example, in a recent study on α-synuclein PRE distances were incorporated into molecular dynamics (MD) simulations to map the free energy landscape of the structural ensemble of the protein [15••]. In another key study, it was shown for pSic1 binding to Cdc4 [16•] that the ensemble can even be solved for the bound state (Figure 1), which provides insight into intricate details of binding and ubiquitination of this protein. These types of studies may represent the first steps toward quantitative structure–function models of IDPs.

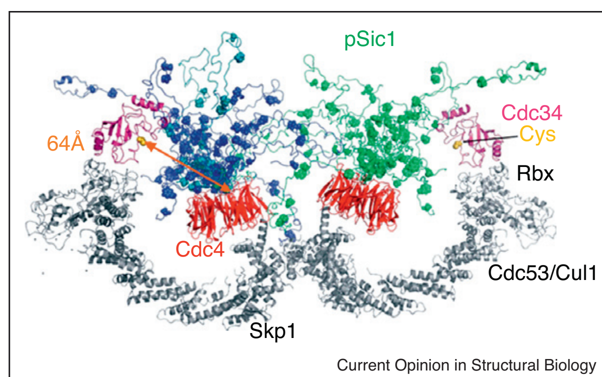
Single-molecule studies

Single-molecule studies of IDP structure, such as atomic-force microscopy (AFM) pulling studies of unfolding transitions [17], fast tapping AFM visualization of structural changes [18•,19], and single-molecule fluorescence resonance energy transfer (smFRET) measurements [20,21,22••] of the range and dynamics of global conformational changes may even surpass the descriptive power of ensemble methods. These approaches allow the observation of transient intermediates and both static and dynamic heterogeneity of structure.

Single-molecule mechanical unfolding by AFM was used to study the conformational heterogeneity of wild-type and mutant α-synuclein [17]. The molecule was found to have three main conformations, disordered, some soluble oligomeric state and 'β-like', which might be important in the transition of the protein toward its pathological amyloid state. AFM can also be used in an ultrafast scanning mode, when it can provide 'movies' of the topological details and conformational transitions of an IDP, such as that of myosin V motor molecules moving along actin tracks enabled by disordered linker regions [18•].

SmFRET is also unparalleled in its spatial and dynamic resolution, as demonstrated by studying of α-synuclein membrane association, structural distributions and dynamics [20]. In a similar study of p53 [21], it was found that its N-terminal domain has multiple preferred

Figure 1



Structural model of pSic1 bound to the SCF^{Cdc4} dimer/Cdc34 complex. The dynamic ensemble of pSic1–Cdc4 complex is calculated by ENSEMBLE using structure constraints obtained by NMR and SAXS. The ensemble is superimposed on a structural model of the SCF^{Cdc4} ubiquitin ligase dimer (the E3) bound to Cdc34 (the E2). Cdc4 is shown in red, Cdc34 in magenta, and the other subunits Skp1, Cdc53/Cul1 and Rbx1 in grey. One pSic1 ensemble is in blue, with the pSic1 ensemble binding to the other Cdc4 subunit shown in green.Adapted from [16•].

Please cite this article in press as: Tompa P. Unstructural biology coming of age, Curr Opin Struct Biol (2011), doi:10.1016/j.sbi.2011.03.012

conformations, some but not all interacting with the DNA-binding domain. The technique even enabled structural studies in live cells, as recently demonstrated for individual SNARE proteins [22••], which became incorporated into folded complexes at the cell membrane.

Structure of IDPs in the bound state

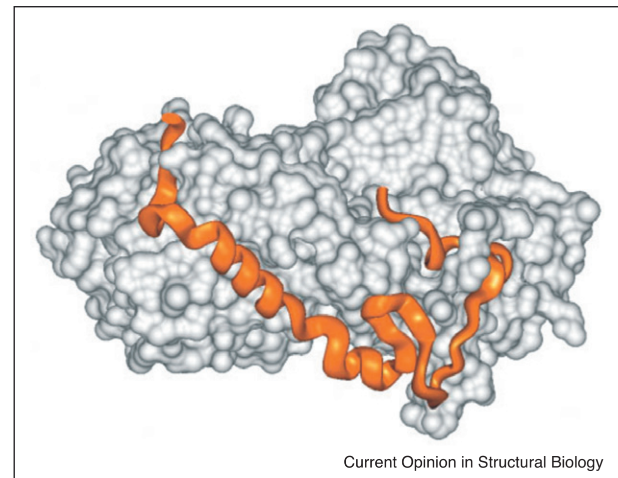
Binding of IDPs to their partners via linear motifs is often weak and is of limited specificity [11]. When stronger, more specific binding is required, IDPs use two distinct strategies. For one, they may use disordered domains for recognition [23], which are longer than 20 residues and conform to all three domain definitions, that is they are autonomous structural, functional, and evolutionary units. These long disordered regions (Figure 2) should be recognized as novel structural–functional elements of IDPs [23]. The other strategy is the combined action of several motifs, as observed in the case of the tripartite binding of calpastatin to calpain [24] or inhibitor 2 to protein phosphatase 1 [25].

This mode of binding (Figure 3a) may add binding strength and specificity, but may also mediate remote initial interaction, in accord with the 'fly casting' model [26]. Such an interaction has been observed in nonsense-mediated decay (NMD) that degrades mRNAs carrying a premature stop codon. NMD is triggered by the assembly of a multiprotein complex that includes three up-frameshift factor (UPF) proteins. Assembly of the NMD complex is initiated by the long disordered C-terminal domain of UPF2 initially binding UPF1 by separate α-helical and β-hairpin elements (Figure 3b) from a distance, thereby 'catching' UPF1 in an encounter complex and bringing various parts of the complex in proximity via fly casting [27].

Interestingly, the mechanism of fly-casting is somewhat confused with 'induced folding' in the literature. For example, when KID domain of transcription factor CREB binds to the KIX domain of CBP [28] or the regulatory domain of p53 binds to S100B(ββ) [29], folding does occur after binding, but an initial weak and long-range interaction may not occur. In fact, kinetic data on the interactions of ordered and disordered proteins [30] show very little difference, probably because their slower diffusion compensates for the larger capture radius of IDPs.

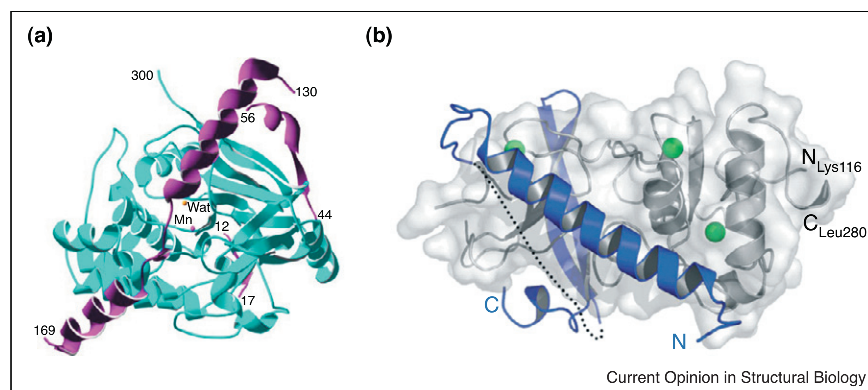
Structural, functional and kinetic dissection of the assembly of multicomponent complexes provides important functional insight into complex regulatory phenomena

Figure 2



Structure of an intrinsically disordered domain bound to its partner. Longer binding regions of IDPs conform to the structural, functional and evolutionary definition of domains. Here the kinase-inhibitory domain (KID) of Cdk inhibitor p27^{Kip1} (orange) bound to the CycA/Cdk2 complex (grey) is shown.
Adapted from [23].

Figure 3



Multipartite binding of IDPs. IDPs often bind their partners via short motifs used in combination to enhance specificity and/or binding strength. **(a)** Inhibitor 2 (purple) wraps around protein phosphatase 1 (green) and binds via three discontinuous binding motifs by nanomolar affinity. Adapted from [25]. **(b)** In assembling the nonsense-mediated decay (NMD) complex, UPF2 (blue ribbon) binds on two opposite surfaces of the CH domain of UPF1 (grey). Adapted from [27]. Please note that parts of IDPs remain disordered in the bound state, which represent a case of fuzziness [33••].

Please cite this article in press as: Tompa P. Unstructural biology coming of age, Curr Opin Struct Biol (2011), doi:10.1016/j.sbi.2011.03.012

[31,32•]. In the case of bacterial toxin/antitoxin (T/A) pairs, the binding of the disordered antitoxin (CcdA or Phd) to the toxin (CcdB or Doc) results in inhibition of the toxin, allosteric release (rejuvenation) of inhibition, and the transition of the T/A operon from a repressed to derepressed state depending on T:A stoichiometry (conditional cooperativity). The subtle interplay between these events demonstrate the inherent complexity of regulation by IDPs [31,32•].

Fuzziness: structural ambiguity in the bound state

Many observations suggest that the dominant mode of IDP function is binding to a partner and concomitant folding [4]. This notion, however, contains a significant element of simplification, because IDPs hardly ever become fully ordered in the bound state [33••], and often their region(s) that remain disordered are important for function. This phenomenon termed 'fuzziness' represents the extension of structural disorder to the bound state. Fuzziness may turn out to be a general structural–functional phenomenon, as suggested by many important cases, such as Sic1 binding to Cdc4 (Figure 1 [16•]), inhibitor 2 binding to PP1c (Figure 3a [34]), and UPF2 binding to UPF1 (Figure 3b [16•]).

What about *in vivo*?

Understanding how IDPs exist and function in cells is complicated by crowding elicited by extreme macromolecular concentrations [35] and binding partners [4], both of which may strongly favor folded states. Several recent studies addressed this question. By applying extremely high concentrations of macromolecular crowding agents, disordered dehydrins of *Arabidopsis thaliana* were found to maintain their disordered character *in vitro* [36]. In addition, functional studies have corroborated that their chaperone function associated with structural disorder *in vitro* is also witnessed *in vivo*, which underlines their structural disorder in a living cell [37•]. In-cell NMR studies addressed the structural state of tau protein in Xenopus oocytes [38], where its microtubule-binding region became ordered, whereas its long projection domain remained largely disordered. On the contrary, α-synuclein, which becomes compacted by crowding conditions *in vitro* [39], remains largely disordered when overexpressed in *Escherichia coli* [40].

Indirect approaches also provide important information on how IDPs behave in a living cell. Recently, it was shown that a simple 'operational' definition of structural disorder can be provided by ubiquitin-independent degradation of IDPs by the 20S proteasome [41]. This relation enables the identification of IDPs *in vivo*, as shown through the regulation of p53 degradation by NAP(P)H quinine oxidoreductase 1 (NQO1) [42]. Similar studies confirm that other IDPs are also disordered and susceptible to 20S-proteasomal degradation *in vivo*, probably regulated by specialized accessory proteins termed 'nannies' [42]. The regulation of IDPs *in vivo* has also been addressed in a bioinformatic study of high-throughput datasets of transcripts and proteins [43•]. It was found that proteins of a high level of disorder are more tightly regulated than proteins of a low level of disorder at all levels of transcription, mRNA clearance, protein synthesis and degradation.

Structural disorder in disease-associated and 'less-evolved' proteins

Structural disorder is enriched in proteins involved in diseases, such as cancer, diabetes, cardiovascular disease and neurodegenerative diseases [44,45]. Disease state caused by IDPs may result not only from protein misfolding [46], but also misidentification, missignaling, and unnatural or nonnative folding, as summarized in the novel D2 (disorder in disorders) concept [45]. Several recent studies provided further details of this correlation.

One observation pertains to chromosomal translocations, which fuse segments of distinct genes and generate oncogenic protein chimeras in cancer. In a comprehensive bioinformatic analysis of 406 translocation-related human proteins, such as BCR-ABL and CBP-MLL [47], these proteins, and their translocation breakpoints in particular, were shown to be significantly enriched in disorder. Apparently, structural disorder enables these chimeras to evade cellular surveillance mechanisms and exert their deleterious functions. Another recent paper addressed the related phenomenon of dosage sensitivity [48•], that is asked what renders gene products harmful when they are overexpressed. It was found that predicted intrinsic protein disorder is the strongest determinant of this effect [48•], suggesting that the likely cause of dosage sensitivity is binding promiscuity of IDPs. Dosage-sensitive genes were also found to be tightly regulated at the transcriptional, RNA and protein levels, as reported for IDPs in general [43•]. Structural disorder is also apparent in viruses. Every step of the viral cell cycle is orchestrated through interactions with cellular proteins for the epigenetic reprogramming of the cell. In most cases viruses use motif-mimicry for this purpose, that is short motifs in disordered regions that compete off similar interactions of the host [49••]. Due to the pressure on the viral genome for compaction, such motifs also represent a very economic solution for high functional density, as demonstrated in the case of the adenoviral E1A oncoprotein [50].

A corollary of the foregoing observations is that structural disorder seems to enable the rapid appearance of novel, 'less-evolved' proteins that have not undergone a long evolutionary selection. For example, alternative splicing (AS) may shift translation reading frame resulting in dual coding, which is only conceivable if the protein product is

Please cite this article in press as: Tompa P. Unstructural biology coming of age, Curr Opin Struct Biol (2011), doi:10.1016/j.sbi.2011.03.012

disordered in at least one of the frames. Comparison of genomic sequences and transcripts has led to the identification of 67 human genes with dual-coding regions at least 75 nucleotides in length [51•]. Predictions did show either a high disorder in both frames, or a significant tendency to become more disordered upon shifting the frame.

Drug development: the new frontier

As seen, IDPs are often involved in disease [44–46], and it is of no doubt that proteins such as p53, BRCA1, CFTR or α-synuclein are preferred targets in drug development. Because the binding pockets of IDPs resemble the actives sites of enzymes, the binding partners of IDPs have been suggested as targetable proteins [44]. The potency of this approach has been demonstrated by nutlins, which can inhibit p53-MDM2 interaction and reactivate p53 pathway in cancer cells [52].

The recent buzz, however, is aroused by the observation that IDPs themselves can be targeted by small molecules [53••,54], as demonstrated in the case of the oncoprotein c-Myc, which can form a heterodimeric complex with Max. In a systematic search several small-molecule inhibitors were found that bind to distinct disordered regions of c-Myc, promote its disordered state and prevent its interaction with Max [54]. This concept probably can be generalized, because small molecules have also been found against other important IDP targets, such as Aβ, EWS-Fli1 and various peptides [53••]. Given the frequent involvement of IDPs in disease [45], the ability to interfere with their action represents tremendous potential in drug discovery, as also suggested in a recent excellent review [55]. An independent corroboration of this generalization has come from analyzing small-molecule γ-secretase modulators (GSM) aimed at selectively lowering Aβ42 levels in Alzheimer's disease [56]. GSMs were found to cross-link to the substrates amyloid precursor protein (APP) and Aβ, rather than γ-secretase, which suggests 'substrate targeting' by small-molecule effectors. Because substrate sites of enzymatic modifications correlate strongly with local disorder [57], this observation is highly relevant to the IDP field.

Conclusions

Although only about a decade old, the field of IDPs has already brought many surprises. The very idea of structural disorder rocked the building of structural biology, and the prevalence of IDPs in normal cell function and importance in pathology has brought the field into the limelight. With a steady advance in our ability to describe their structure and function in detail, now the next transition in the field is gaining momentum in which the generation of quantitative structural models of IDP function becomes possible. We reckon this transition will bring 'unstructural' biology the full recognition and appreciation it deserves, to be surely witnessed in the coming years.

Acknowledgements

This research was supported by grant NK71582 from the Hungarian Scientific Research Fund (OTKA), a Korean-Hungarian Joint Laboratory grant from Korea Research Council of Fundamental Science and Technology (KRCF), and both an FP7 Marie Curie Initial Training Network grant (no. 264257, IDPbyNMR) and an FP7 Infrastructures grant (no. 261863, BioNMR) from the European Commission.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Dyson HJ, Wright PE: **Intrinsically unstructured proteins and their functions**. *Nat Rev Mol Cell Biol* 2005, **6**:197-208.
2. Dunker AK, Silman I, Uversky VN, Sussman JL: **Function and structure of inherently disordered proteins**. *Curr Opin Struct Biol* 2008, **18**:756-764.
3. Gsponer J, Babu MM: **The rules of disorder or why disorder rules**. *Prog Biophys Mol Biol* 2009, **99**:94-103.
4. Wright PE, Dyson HJ: **Linking folding and binding**. *Curr Opin Struct Biol* 2009, **19**:1-8.
5. Uversky VN, Dunker AK: **Understanding protein non-folding**. *Biochim Biophys Acta* 2010, **1804**:1231-1264.
6. Tompa P: *Structure and Function of Intrinsically Disordered Proteins*. Boca Raton, FL: CRC Press (Taylor and Francis Group); 2009.
7. Sickmeier M, Hamilton JA, LeGall T, Vacic V, Cortese MS, Tantos A, Szabo B, Tompa P, Chen J, Uversky VN *et al.*: **DisProt: the database of disordered proteins**. *Nucleic Acids Res* 2007, **35**:D786-793.
8. He B, Wang K, Liu Y, Xue B, Uversky VN, Dunker AK: **Predicting intrinsic disorder in proteins: an overview**. *Cell Res* 2009, **19**:929-949.
9. Ishida T, Kinoshita K: **Prediction of disordered regions in proteins based on the meta approach**. *Bioinformatics* 2008, **24**:1344-1348.
10. Xue B, Dunbrack RL, Williams RW, Dunker AK, Uversky VN: **PONDR-FIT: a meta-predictor of intrinsically disordered amino acids**. *Biochim Biophys Acta* 2010, **1804**:996-1010.
11. Gould CM, Diella F, Via A, Puntervoll P, Gemund C, Chabanis-Davidson S, Michael S, Sayadi A, Bryne JC, Chica C *et al.*: **ELM: the status of the 2010 eukaryotic linear motif resource**. *Nucleic Acids Res* 2009, **38**:D167-180.
12. Davey NE, Haslam NJ, Shields DC, Edwards RJ: **SLiMFinder: a web server to find novel, significantly over-represented, short protein motifs**. *Nucleic Acids Res* 2010, **38**:W534-539.
13. Meszaros B, Simon I, Dosztanyi Z: **Prediction of protein binding regions in disordered proteins**. *PLoS Comput Biol* 2009, **5**:e1000376.
14. Cheng Y, Oldfield CJ, Meng J, Romero P, Uversky VN, Dunker AK: **Mining alpha-helix-forming molecular recognition features with cross species sequence alignments**. *Biochemistry* 2007, **46**:13468-13477.
15. •• Allison JR, Varnai P, Dobson CM, Vendruscolo M: **Determination of the free energy landscape of alpha-synuclein using spin label nuclear magnetic resonance measurements**. *J Am Chem Soc* 2009, **131**:18314-18326.

The complete structural description of the ensemble of IDP structures is an important development of the field. This paper is a nice illustration of even one step further, because here a combination of PRE NMR distance restraints are used in molecular dynamics simulations to obtain a mapping of the relative weight of conformations in the ensemble providing the

Please cite this article in press as: Tompa P. Unstructural biology coming of age, Curr Opin Struct Biol (2011), doi:10.1016/j.sbi.2011.03.012

free energy landscape of a natively unfolded protein, α-synuclein. The importance of this method is that it may lead to developing quantitative structural-functional models of IDPs.

16. • Mittag T, Marsh J, Grishaev A, Orlicky S, Lin H, Sicheri F, Tyers M, Forman-Kay JD: **Structure/function implications in a dynamic complex of the intrinsically disordered Sic1 with the Cdc4 subunit of an SCF ubiquitin ligase**. *Structure* 2010, **18**:494-506.
Detailed ensemble characterization of the structure of the yeast Cdk-inhibitor Sic1 in complex with its partner ubiquitin-ligase subunit Cdc4. This is a prime example of a highly dynamic 'fuzzy' complex between the two proteins, which is promoted by multiple phosphorylation of several sub-optimal binding elements of Sic1. To characterize the ensuing poly-electrostatic interaction, here small-angle X-ray scattering (SAXS) and NMR data are used for calculations of the ensemble of structures of Sic1, pSic1 and the pSic1-Cdc4 complexes. These results provide one of the first examples of the real structural model of an IDP that is predominantly disordered in both its free and bound states, which enables its structure/function relationship to be elucidated.

17. Sandal M, Valle F, Tessari I, Mammi S, Bergantino E, Musiani F, Brucale M, Bubacco L, Samori B: **Conformational equilibria in monomeric alpha-synuclein at the single-molecule level**. *PLoS Biol* 2008, **6**:e6.

18. • Kodera N, Yamamoto D, Ishikawa R, Ando T: **Video imaging of walking myosin V by high-speed atomic force microscopy**. *Nature* 2010, **468**:72-76.
A key recent transition in the description of IDP structures is the application of single-molecule techniques, which may allow the observation of both static and dynamic heterogeneity in IDP structure without ensemble averaging. This paper describes the invention of one such technique, high-speed atomic force microscopy, which allows direct visualization of conformational transitions in structural disorder. The dynamic behavior of myosin V molecules translocating along actin filaments enabled by disordered linker regions has been directly visualized. The 'movie' provides direct evidence of dynamic molecular behavior, leading to a comprehensive understanding of the motor mechanism. This technique may become one of the most powerful approaches to studying the structure and dynamics of IDPs in action.

19. Yamamoto D, Uchihashi T, Kodera N, Yamashita H, Nishikori S, Ogura T, Shibata M, Ando T: **High-speed atomic force microscopy techniques for observing dynamic biomolecular processes**. *Methods Enzymol* 2010, **475**:541-564.

20. Ferreon AC, Gambin Y, Lemke EA, Deniz AA: **Interplay of alpha-synuclein binding and conformational switching probed by single-molecule fluorescence**. *Proc Natl Acad Sci U S A* 2009, **106**:5645-5650.

21. Huang F, Rajagopalan S, Settanni G, Marsh RJ, Armoogum DA, Nicolaou N, Bain AJ, Lerner E, Haas E, Ying L *et al.*: **Multiple conformations of full-length p53 detected with single-molecule fluorescence resonance energy transfer**. *Proc Natl Acad Sci U S A* 2009, **106**:20758-20763.

22. •• Sakon JJ, Weninger KR: **Detecting the conformation of individual proteins in live cells**. *Nat Methods* 2010, **7**:203-205.
As suggested, single-molecule characterization provides unprecedented detail in the description of static and dynamic heterogeneity of IDP structures. Here the use of single-molecule fluorescence resonance energy transfer (smFRET) in combination with single-particle tracking is demonstrated to allow the detection of *in vivo* conformation of individual SNARE proteins. It was found the proteins got rapidly incorporated into folded complexes at the cell membrane, which demonstrates the potential of this technique to reveal dynamic conformational changes and interactions of IDPs in cells.

23. 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* 2009, **31**:328-335.

24. Kiss R, Bozoky Z, Kovacs D, Rona G, Friedrich P, Dvortsak P, Weisemann R, Tompa P, Perczel A: **Calcium-induced tripartite binding of intrinsically disordered calpastatin to its cognate enzyme, calpain**. *FEBS Lett* 2008, **582**:2149-2154.

25. Hurley TD, Yang J, Zhang L, Goodwin KD, Zou Q, Cortese M, Dunker AK, DePaoli-Roach AA: **Structural basis for regulation of protein phosphatase 1 by inhibitor-2**. *J Biol Chem* 2007, **282**:28874-28883.

26. Shoemaker BA, Portman JJ, Wolynes PG: **Speeding molecular recognition by using the folding funnel: the fly-casting mechanism**. *Proc Natl Acad Sci U S A* 2000, **97**:8868-8873.

27. Clerici M, Mourao A, Gutsche I, Gehring NH, Hentze MW, Kulozik A, Kadlec J, Sattler M, Cusack S: **Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2**. *EMBO J* 2009, **28**:2293-2306.

28. Turjanski AG, Gutkind JS, Best RB, Hummer G: **Binding-induced folding of a natively unstructured transcription factor**. *PLoS Comput Biol* 2008, **4**:e1000060.

29. Chen J: **Intrinsically disordered p53 extreme C-terminus binds to S100B(betabeta) through "fly-casting"**. *J Am Chem Soc* 2009, **131**:2088-2089.

30. Huang Y, Liu Z: **Kinetic advantage of intrinsically disordered proteins in coupled folding-binding process: a critical assessment of the "fly-casting" mechanism**. *J Mol Biol* 2009, **393**:1143-1159.

31. De Jonge N, Garcia-Pino A, Buts L, Haesaerts S, Charlier D, Zangger K, Wyns L, De Greve H, Loris R: **Rejuvenation of CcdB-poisoned gyrase by an intrinsically disordered protein domain**. *Mol Cell* 2009, **35**:154-163.

32. • Garcia-Pino A, Balasubramanian S, Wyns L, Gazit E, De Greve H, Magnuson RD, Charlier D, van Nuland NA, Loris R: **Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity**. *Cell* 2010, **142**:101-111.
Detailed study of the Doc/Phd toxin–antitoxin system, which shows complex structure–function phenomena leading to the concentration-dependent autoregulatory repression or de-repression of the doc/phd operon. Intrinsic disorder of the antitoxin, its induced folding, allostery, and complex stoichiometries in binding constitute elements of conditionally cooperative regulation of transcription typical not only of this, but possibly other toxin–antitoxin modules. A nice example of the possible complexities of IDP function outlining an experimental strategy to characterize it in detail.

33. •• Tompa P, Fuxreiter M: **Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions**. *Trends Biochem Sci* 2008, **33**:2-8.
It is generally thought that IDPs undergo induced folding to their final, function-related structured state when they function by molecular recognition. In such a case, their function can be interpreted in terms of (bound) structure, which is very similar to the classical structure–function paradigm. On the contrary, it is demonstrated in this review that IDPs are never fully ordered even in the bound state, but show a significant amount of structural disorder or polymorphism in protein complexes. This function-related phenomenon termed 'fuzziness' represents the extension of the paradigm of structural disorder to the functional-bound-state. Due to the central role of disorder in protein–protein interactions and in regulatory processes, fuzziness may turn out to be a mechanism of fundamental importance in the interactome.

34. Marsh JA, Dancheck B, Ragusa MJ, Allaire M, Forman-Kay JD, Peti W: **Structural diversity in free and bound states of intrinsically disordered protein phosphatase 1 regulators**. *Structure* 2010, **18**:1094-1103.

35. Ellis RJ: **Macromolecular crowding: obvious but underappreciated**. *Trends Biochem Sci* 2001, **26**:597-604.

36. Mouillon JM, Eriksson SK, Harryson P: **Mimicking the plant cell interior under water stress by macromolecular crowding: disordered dehydrin proteins are highly resistant to structural collapse**. *Plant Physiol* 2008, **148**:1925-1937.

37. • Chakrabortee S, Meersman F, Kaminski Schierle GS, Bertoncini CW, McGee B, Kaminski CF, Tunnacliffe A: **Catalytic and chaperone-like functions in an intrinsically disordered protein associated with desiccation tolerance**. *Proc Natl Acad Sci U S A* 2010, **107**:16084-16089.
A critical issue of the IDP field is how much *in vitro* structural and functional observations can be extrapolated to live cells. Here in-cell functional characterization of an anhydrin thought to provide desiccation tolerance in an anhydrobiotic nematode, *Aphelenchus avenae*, is carried out. When the protein is expressed in cells, it can reduce protein aggregation, in which a loose association with its client protein could be shown. This function is consistent with a physiological role of this protein as a molecular shield.

Please cite this article in press as: Tompa P. Unstructural biology coming of age, Curr Opin Struct Biol (2011), doi:10.1016/j.sbi.2011.03.012

38. Bodart JF, Wieruszeski JM, Amniai L, Leroy A, Landrieu I, Rousseau-Lescuyer A, Vilain JP, Lippens G: **NMR observation of Tau in Xenopus oocytes**. *J Magn Reson* 2008, **192**:252-257.

39. Uversky VN, Li J, Fink AL: **Trimethylamine-N-oxide-induced folding of alpha-synuclein**. *FEBS Lett* 2001, **509**:31-35.

40. McNulty BC, Young GB, Pielak GJ: **Macromolecular crowding in the *Escherichia coli* periplasm maintains alpha-synuclein disorder**. *J Mol Biol* 2006, **355**:893-897.

41. Tsvetkov P, Asher G, Paz A, Reuven N, Sussman JL, Silman I, Shaul Y: **Operational definition of intrinsically unstructured protein sequences based on susceptibility to the 20S proteasome**. *Proteins* 2008, **70**:1357-1366.

42. Tsvetkov P, Reuven N, Shaul Y: **The nanny model for IDPs**. *Nat Chem Biol* 2009, **5**:778-781.

43. • Gsponer J, Futschik ME, Teichmann SA, Babu MM: **Tight regulation of unstructured proteins: from transcript synthesis to protein degradation**. *Science* 2008, **322**:1365-1368.

An ambitious computational study of multiple high-throughput datasets aimed at elucidating the regulation of the level and availability of IDPs in cells. It was found that regulation of transcript clearance, proteolytic degradation and translational rate contribute to controlling the abundance of IDPs, some of which are only present in low amounts and for short periods of time. Fidelity in signaling of IDPs indicated by low stochasticity in transcription and translation rates indicate fine-tuning of the availability of IDPs, which may ensure that most IDPs are only available in appropriate amounts and only for the time they are needed.

44. Cheng Y, Legall T, Oldfield CJ, Mueller JP, Van YY, Romero P, Cortese MS, Uversky VN, Dunker AK: **Rational drug design via intrinsically disordered protein**. *Trends Biotechnol* 2006, **24**:435-442.

45. Uversky VN, Oldfield CJ, Dunker AK: **Intrinsically disordered proteins in human diseases: introducing the D2 concept**. *Annu Rev Biophys* 2008, **37**:215-246.

46. Chiti F, Dobson CM: **Protein misfolding, functional amyloid, and human disease**. *Annu Rev Biochem* 2006, **75**:333-366.

47. Hegyi H, Buday L, Tompa P: **Intrinsic structural disorder confers cellular viability on oncogenic fusion proteins**. *PLoS Comput Biol* 2009, **5**:e1000552.

48. • Vavouri T, Semple JI, Garcia-Verdugo R, Lehner B: **Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity**. *Cell* 2009, **138**:198-208.

Structural disorder is known to correlate with disease, where mutations are thought to impair function or expression of the IDP, and disease results from loss of function. In this paper it is asked why certain genes are harmful when they are overexpressed. By analyzing overexpression phenotypes in yeast, intrinsic protein disorder is identified as an important determinant of dosage sensitivity. This inference is also validated in other species, fruit fly, worm, mouse and humans. It is suggested that disordered regions are prone to make promiscuous molecular interactions at elevated concentrations, which is the likely cause of pathology when genes are overexpressed.

49. •• Davey NE, Trave G, Gibson TJ: **How viruses hijack cell regulation**. *Trends Biochem Sci* 2010, **36**:159-169.

A central theme of the IDP field is that disordered proteins/regions function by molecular recognition, in which their short binding elements (linear motifs) bind their partner and undergo induced folding. In this important paper it is reviewed that viruses also use this mechanism very frequently for interactions with host proteins and epigenetic reprogramming of the host cell for their own needs. As shown by adequate examples, viruses use motif mimicking for this purpose, which has many advantages for them, such as rapid evolution, stronger binding their host competitor and high functional density.

50. Ferreon JC, Martinez-Yamout MA, Dyson HJ, Wright PE: **Structural basis for subversion of cellular control mechanisms by the adenoviral E1A oncoprotein**. *Proc Natl Acad Sci U S A* 2009, **106**:13260-13265.

51. • Kovacs E, Tompa P, Liliom K, Kalmar L: **Dual coding in alternative reading frames correlates with intrinsic protein disorder**. *Proc Natl Acad Sci U S A* 2010, **107**:5429-5434.

It is shown in this paper that there are many human genes in which alternative splicing shifts the reading frame, that is generates mature transcripts in which a segment is translated in two distinct frames. It is shown that the protein folding problem caused by dual coding is resolved by an elevated level of structural disorder in at least one of the frames. It is also shown that dual coding is under adaptive evolution and it may be an effective mechanism for the evolutionary appearance of novel functions.

52. Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, Filipovic Z, Kong N, Kammlott U, Lukacs C, Klein C *et al.*: **In vivo activation of the p53 pathway by small-molecule antagonists of MDM2**. *Science* 2004, **303**:844-848.

53. •• Metallo SJ: **Intrinsically disordered proteins are potential drug targets**. *Curr Opin Chem Biol* 2010, **14**:481-488.

Despite increasing investments, the pace of discovery of new drugs is leveling off, which demands novel approaches of drug development. Binding partners of IDPs have recently been suggested as attractive drug targets. In this review it is shown through the example of c-Myc and other IDPs that this idea can be taken one step further because IDPs themselves can be targeted by small-molecule interactors. Because structural disorder is very frequent in disease-associated proteins, this approach may be one of the fastest-developing areas in the IDP field.

54. 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, **131**:7390-7401.

55. Dunker AK, Uversky VN: **Drugs for 'protein clouds': targeting intrinsically disordered transcription factors**. *Curr Opin Pharmacol* 2010, **10**:782-788.

56. Kukar TL, Ladd TB, Bann MA, Fraering PC, Narlawar R, Maharvi GM, Healy B, Chapman R, Welzel AT, Price RW *et al.*: **Substrate-targeting gamma-secretase modulators**. *Nature* 2008, **453**:925-929.

57. Fuxreiter M, Tompa P, Simon I: **Local structural disorder imparts plasticity on linear motifs**. *Bioinformatics* 2007, **23**:950-956.

58. Wells M, Tidow H, Rutherford TJ, Markwick P, Jensen MR, Mylonas E, Svergun DI, Blackledge M, Fersht AR: **Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain**. *Proc Natl Acad Sci U S A* 2008, **105**:5762-5767.

59. Jensen MR, Markwick PR, Meier S, Griesinger C, Zweckstetter M, Grzesiek S, Bernado P, Blackledge M: **Quantitative determination of the conformational properties of partially folded and intrinsically disordered proteins using NMR dipolar couplings**. *Structure* 2009, **17**:1169-1185.

60. Salmon L, Nodet G, Ozenne V, Yin G, Jensen MR, Zweckstetter M, Blackledge M: **NMR characterization of long-range order in intrinsically disordered proteins**. *J Am Chem Soc* 2010, **132**:8407-8418.

61. Bernado P, Modig K, Grela P, Svergun DI, Tchorzewski M, Pons M, Akke M: **Structure and dynamics of ribosomal protein L12: an ensemble model based on SAXS and NMR relaxation**. *Biophys J* 2010, **98**:2374-2382.

62. Galea CA, Nourse A, Wang Y, Sivakolundu SG, Heller WT, Kriwacki RW: **Role of intrinsic flexibility in signal transduction mediated by the cell cycle regulator, p27 Kip1**. *J Mol Biol* 2008, **376**:827-838.

Please cite this article in press as: Tompa P. Unstructural biology coming of age, Curr Opin Struct Biol (2011), doi:10.1016/j.sbi.2011.03.012