Published in final edited form as:

J Biomol NMR. 2019 December ; 73(12): 651–659. doi:10.1007/s10858-019-00280-2.

Perspective: the essential role of NMR in the discovery and characterization of intrinsically disordered proteins

H. Jane Dyson¹, Peter E. Wright¹

¹Department of Integrative Structural and Computational Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

Abstract

The 2019 ISMAR Prize recognized NMR studies of disordered proteins. Here we provide a highly personal perspective on the discovery of intrinsically disordered proteins and the development and application of NMR methods to characterize their conformational ensembles, dynamics, and interactions.

Keywords

Intrinsically disordered proteins; Immunogenic peptides; Protein folding; Relaxation dispersion

Today, the central role that intrinsically disordered proteins (IDPs) play in cellular processes is widely recognized. This was not always the case. To understand where the IDP concept came from, and why it had not surfaced previously, we need to take ourselves back to the scientific mindset before 1990. The prevailing dogma at that time was that protein function was directly encoded in three-dimensional structure. The human genome project, proposed in 1988, was not on anybody's radar—very few if any genes had been sequenced. Cloning and expression of isotopically labeled proteins was in its infancy. Computing power was pitiful compared to today—the capacity of our institute-wide mainframe was less than that of a smart phone today. For protein NMR, the dominant technique was homonuclear ¹H, and 2D was only introduced in the late 1970s (Wokaun and Ernst 1977; Jeener et al. 1979). Heteronuclear spectroscopy of uniformly-labeled proteins was introduced in 1989 (Gronenborn et al. 1989).

In 1984 we moved from Sydney, Australia to Scripps Research in La Jolla, California, where a big incentive was our very own state-of-the-art 500 MHz NMR spectrometer! In our new environment, we set about working with our new neighbors. Richard Lerner and his colleagues were very interested at that time in explaining the phenomenon of anti-peptide antibodies, that is, how antibodies raised against a small peptide immunogen could recognize the intact natively folded protein (Niman et al. 1983). Reasoning that there must be structural correspondence of some sort between the same sequence when present in the peptide and in the protein, we undertook to examine one of these peptides, a 9-residue

H. Jane Dyson, dyson@scripps.edu, Peter E. Wright, wright@scripps.edu.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

sequence from influenza virus hemagglutinin, in solution by NMR. Of course, we could not approach NMR studies of the entire trimeric, glycosylated protein, which would remain challenging (though not impossible) today, but we were also fortunate in that another colleague, Ian Wilson, had solved the crystal structure of the protein (Wilson et al. 1981).

For the NMR studies of the immunogenic hemagglutinin peptide, we initially employed both 1D and 2D homonuclear ^1H NMR spectroscopy to make assignments and probe the peptide conformation (Dyson et al. 1985). Remarkably, the NMR data showed that the hemagglutinin peptide formed a detectable population of a type II β -turn in aqueous solution, stabilized by an intramolecular hydrogen bond (Dyson et al. 1985). This discovery was met with great skepticism at that time since it was generally believed that short linear peptides, with the sole exception of the C-peptide fragment of ribonuclease A (Brown and Klee 1971), are unstructured in water solution and, in particular, that competition with solvent water would preclude formation of intrapeptide hydrogen bonds. However, our NMR evidence for a reverse turn in the hemagglutinin peptide was unequivocal, and subsequent experiments, where we used 2D rotating-frame NOESY (ROESY) spectra to analyze the conformations of a whole series of hemagglutinin peptide variants in which the amino acid sequence was changed in systematic ways (Dyson et al. 1988a), showed us that local sequence directly determines the propensity of a peptide to take up a preferred structure. The data also showed clearly that the reverse turn structure is in rapid dynamic exchange with a multitude of unfolded states. NMR studies of another immunogenic peptide, derived from a helical region of myohemerythrin, revealed formation of nascent helical structure that rapidly interconverts with fully unfolded states (Dyson et al. 1988b). These studies, and those of peptide fragments comprising the entire sequence of several proteins (Dyson et al. 1992a, b; Reymond et al. 1997), taught us that all peptides, even those with the highest populations of structured forms, are ensembles of interchanging structures in water solution, and that the NMR spectrum is a population-weighted average over the entire conformational ensemble. From these studies, we were able to make quantitative interpretations of conformational propensities of peptides (Wright et al. 1988; Dyson and Wright 1991), comparing the relative intensities of sequential HN–HN and HN–H α NOE cross peaks to probe the distribution of backbone dihedral angles on the Ramachandran map, and using medium-range NOEs and coupling constants to identify folded conformers within the ensemble, together with amide proton temperature coefficients to identify intramolecular hydrogen bonds.

That the NMR spectrum of a peptide is a population-weighted average of the spectra of all of the component conformer structures seems intuitively obvious to anyone who has looked at the NMR spectra of peptides in aqueous solution, but bear in mind what the general scientific consensus was before the mid-1990s: proteins must be folded to be functional. We became very interested in looking at protein folding, another story that has been told elsewhere (Dyson and Wright 2017). Again NMR was central to our approach to the problem, for example in the identification of protein folding initiation sites (Dyson et al. 1992a, b) and characterization of unfolded and partly folded protein states (Yao et al. 1997; Eliezer et al. 1998). Although these studies were aimed at the elucidation of protein folding, the methods that we developed to characterize the molten globule and acid unfolded states of apomyoglobin laid the foundation for our NMR studies of IDPs. Throughout this early work,

the theme that conformational ensembles are always present in water solutions of peptides (and unfolded forms of globular proteins) was constant. But were functional proteins qualitatively different? They consist, like peptides, of polypeptide chains. Why should every protein have to be folded into a well-defined 3D structure before it can function? (In fact, we already had clues that they do not. One of us (PEW) had shown earlier using 1D ^1H NMR that the protein chromogranin A, which functions in storage of ATP and adrenaline in the chromaffin granules of the adrenal gland, is a highly dynamic “random coil” both as a purified protein and in intact tissue (Daniels et al. 1976, 1978).)

That question was not at the forefront of our minds in the 1990s. We were occupied like many other labs in using the new cloning and expression methods, together with the exciting new heteronuclear NMR experiments that could now be used efficiently with isotopically-labeled protein samples, to determine three-dimensional solution structures of proteins (Moore et al. 1988; Lee et al. 1989; Dyson et al. 1990). Yet it was clear that even proteins that were well-structured contained regions where the family of structures did not cluster closely but was rather disordered. This disorder could be due to a scarcity of NOE data in the region, or to actual dynamics, which could be measured by NMR relaxation. Genes were beginning to be sequenced routinely in the early 1990s, but it was still long before the completion of the human genome project! Collaborators Marc Montminy at the Salk Institute and Rudi Grosschedl at UCSF worked out the gene and protein sequences corresponding to a number of important transcription-related proteins, and pinpointed which regions of these proteins were involved in interactions in the cell. We could look at these interactions by NMR, using different labeling combinations and engineering of NMR-amenable fragments, later brought to a fine art by our in-house design genius Maria Yamout. The same paradoxical observation was made in three studies at that time, of the HMG domain of the DNA-bending protein LEF-1 (Love et al. 1995, 2004), the p21 cyclin-dependent kinase inhibitor (Kriwacki et al. 1996), and the complex between the phosphorylated kinase-inducible domain (KID) of CREB and the KIX domain of CREB-binding protein (CBP) (Radhakrishnan et al. 1997). In each case, the key interaction sites all showed the same NMR signature: limited ^1H dispersion, more typical of peptides than proteins, in the free state, and increased dispersion upon binding to the target protein or DNA, indicating that these proteins become more structured in complexes with their binding partners (Fig. 1). This folding-upon-binding phenomenon has been a recurring theme in the systems that we have studied, though we and other groups find that bound structures can also be disordered (Borg et al. 2007), referred to as fuzzy interactions (Tompa and Fuxreiter 2008).

In addition, we characterized the prion protein by NMR, including the lengthy disordered N-terminus that contributes to disease-related amyloid formation (Viles et al. 2001). The presence of both a structured domain and a disordered N-terminal region resulted in the observation of NMR relaxation behavior that could not be analyzed through the standard model-free methods since tumbling of the globular core is influenced by the disordered tail (Bae et al. 2009).

The situation at the end of the 1990s was this: several projects in our lab involved complexes in which at least one of the partners is completely unstructured in the absence of the other

[and examples where both partners are intrinsically disordered (Demarest et al. 2002)]. This led us in 1999 to write a review article “Intrinsically Unstructured Proteins: Re-assessing the Protein Structure–Function Paradigm” (Wright and Dyson 1999). “Intrinsically unstructured proteins” was the term we had come up with, but a similar idea had arisen from observations of missing density in local regions of X-ray structures, and in bioinformatic analyses. Keith Dunker had noticed that an astonishingly high proportion of the genome sequences that were by now beginning to appear would code for proteins that would not be expected to adopt stable globular structures with an ordered hydrophobic core (Romero et al. 1997; Dunker et al. 2001). His term was “intrinsically disordered proteins”, and this was the term eventually adopted as a standard by the field (Dunker et al. 2013; van der Lee et al. 2014a).

Similarly to the reaction to our earlier discovery of secondary structure in small peptide fragments in water solution, the idea that fully-functional proteins could be intrinsically disordered (“unstructured”) was met with extreme skepticism. It was argued that unstructured proteins were an artifact of sample preparation and could not exist in the cell because they would be degraded by proteolytic enzymes. It was further argued that IDPs are merely “proteins waiting for a partner” (PWPs) (Janin and Sternberg 2013)—that disorder is observed because IDPs are studied in the absence of their natural binding targets. However, the vast body of data that has accumulated over the past 25 years shows that such criticisms are unfounded. It is now clear that IDPs are highly abundant in eukaryotes, that they are essential to an enormous array of biological processes, including organization of phase separated sub-cellular structures, that they do not always fold on interaction with binding partners, that they have comparable cellular lifetimes to globular proteins, and that they can remain unstructured in the cell (van der Lee et al. 2014b; Wright and Dyson 2015; Theillet et al. 2014, 2016).

It is impossible to imagine the advent and growth of the IDP concept without the contribution of solution NMR, which gives quantitative information about ensembles that is not available from any other technique. ^{15}N and ^{13}C labeling is mandatory for assignment of the spectrum and for conformational analysis for any but the smallest disordered peptides; the intrinsically large chemical shift dispersion of ^{15}N and ^{13}CO is essential for assignment of most IDP spectra (Yao et al. 1997; Neri et al. 1992). Once assignments have been made, deviations of ^{13}Ca and ^{13}CO chemical shifts from sequence corrected “random coil” shifts (Schwarzinger et al. 2000, 2001; Tamiola et al. 2010; Nielsen and Mulder 2018) allow quantitative assessment of helical and other secondary structure populations (Eliezer et al. 1998; Radhakrishnan et al. 1998; Yao et al. 2001). Other NMR parameters provide detailed insights into the conformational ensemble; ^{15}N relaxation can be used to probe dynamics and identify local collapse of the polypeptide chain (Schwarzinger et al. 2002), residual dipolar couplings identify local compaction and secondary structural preferences (Mohana-Borges et al. 2004), and paramagnetic relaxation enhancement (PRE) from coupled spin label probes reveals transient long-range interactions (Gillespie and Shortle 1997; Lietzow et al. 2002). These NMR parameters are now used routinely for characterization of IDPs and form the basis for quantitative modeling of the conformational ensemble and energy landscape (Marsh and Forman-Kay 2012; Jensen et al. 2014).

A major advantage of NMR is that it can be used to characterize the conformation and dynamics of the free and bound states of the IDP as well as the binding process itself. ^1H – ^{15}N HSQC titrations allow facile identification of the number and location of binding sites and, if exchange between free and bound states is in the fast exchange regime, a quantitative measure of the binding affinity (Lee et al. 2009). Remarkably, in cases where an IDP binds to two sites on its target protein, primary dissociation constants of 10 nM or less can be determined by chemical shift titrations, provided the second binding process is weak and coupled to the primary binding event (Arai et al. 2012). Relaxation-compensated Carr-Purcell-Meiboom-Gill R_2 relaxation dispersion experiments provide novel insights into the kinetics and mechanism of IDP binding processes. ^{15}N R_2 relaxation dispersion was used to characterize the kinetics and mechanism of coupled folding and binding of two disordered transcription factors to the structured KIX domain of CBP. In one case, the phosphorylated KID domain of CREB, the NMR data showed that the binding mechanism was primarily “induced-fit”, where the disordered partner (pKID) folds after binding to the structured target protein KIX (Sugase et al. 2007). In contrast, the transcriptional activation domain of c-Myb, which has a strong propensity to populate helical structure, binds to KIX by a “conformational selection” mechanism, where the initial contact is made by a highly populated helical region of c-Myb (Arai et al. 2015).

NMR studies of disordered regions within larger proteins that contain both intrinsically disordered and folded globular domains, which are common in eukaryotic proteomes, are even more challenging. Not only are the relaxation properties of these proteins problematic, giving rise to spectra that contain both intense and broad, weak resonances, but their large size and the presence of disorder causes resonance overlap that frequently cannot be resolved even with multiply-labeled proteins and multi-dimensional spectroscopy. We have recently employed segmental labeling of proteins such as the 180 kD p53 tetramer using intein methods, which resolves much of the overlap problem (Krois et al. 2016, 2018). It is also frequently the case that relaxation properties are unfavorable for complexes of IDPs due to exchange broadening associated with the kinetics of binding or due to conformational fluctuations in the bound state. Solutions to this problem include the use of direct ^{13}C or ^{15}N detection methods (Takeuchi et al. 2015) or fusion of the disordered and target molecules to reduce exchange broadening (Krois et al. 2016). These examples (plus many from labs other than our own) illustrate the crucial importance of NMR in the understanding of both the free disordered domains and their structured or fuzzy complexes.

Our peptide and protein-folding studies laid the ground-work for acceptance of the idea of ensembles as the norm for flexible polypeptides. Once we accepted that an unstructured peptide in solution could fold upon binding to a partner, the idea began to make a lot of sense. The interaction sites of many eukaryotic signaling proteins are disordered—this means that they can interact with multiple partners, mediating cross-talk between pathways, and allowing an efficient use of cellular resources for multiple processes. Disordered proteins have many different, energetically-equivalent structures available within the ensemble, and can fold into different structures upon binding to different partners. An example of this is the transactivation domain of the hypoxia-inducible factor alpha (HIF-1 α), which folds into a largely-helical structure on binding to the TAZ1 domain of CBP/p300 (Dames et al. 2002; Freedman et al. 2002), but as an extended structure on the

enzyme that hydroxylates it (Elkins et al. 2003) (Fig. 2). Disordered regions of larger proteins frequently also contain multiple sites for post-translational modification, allowing not only the utilization of these molecules under many different cellular conditions, but also providing a mechanism for signals to be turned on/off and modulated. Disordered proteins too can readily be degraded in a ubiquitin-independent manner, by the 20S proteasome, providing a further means for fine control of cellular processes (Tsvetkov et al. 2009).

Since the realization in the late 1990s that disordered proteins are abundant and perform essential biological functions, there has been an explosion of interest in IDPs, with hundreds of papers published each year. We now recognize that IDPs participate in an enormous range of cellular processes and new and quite unexpected functions and properties are constantly being revealed. For example, recent NMR and single molecule FRET experiments showed that two entirely disordered proteins can associate with high affinity to form a disordered complex (Borgia et al. 2018). Our own recent work led to the unexpected discovery that intrinsic disorder forms the mechanistic basis of unique unidirectional switches that regulate critical signaling pathways (Berlow et al. 2017).

The cellular response to hypoxia is essential for survival at low or fluctuating oxygen levels and is tightly regulated. Under all cellular conditions, the hypoxia-inducible factor HIF-1 α is continuously synthesized, but in the presence of normal levels of oxygen, it is rapidly degraded following hydroxylation of two proline residues (Pro402 and Pro564) in an intrinsically disordered region of the protein (Semenza 2004). In addition, the hydroxylation of asparagine 803 (Land et al. 2002) in the intrinsically disordered C-terminal activation domain (shown in Fig. 2) destabilizes the TAZ1:HIF-1 α complex. Under conditions of hypoxia, hydroxylation no longer occurs and the HIF-1 α is stabilized to mediate the transcription of hypoxia-responsive genes (Semenza 2004; Ivan et al. 2001; Jaakkola et al. 2001). The hypoxic response upregulates transcription of oxygen stress genes that must be rapidly turned off when normal oxygen levels are restored; a prolonged hypoxic response can lead to apoptosis mediated by p53 (Schmid et al. 2004). Robust feedback mechanisms are in place to halt the hypoxic response when it is no longer needed. One of these mechanisms employs the negative regulator CITED2, which is encoded by one of the genes controlled by HIF-1 α . CITED2, another IDP, competes with HIF-1 α for binding to the TAZ1 domain of CBP. The disordered activation domains of HIF-1 α and CITED2 bind to TAZ1 in sites that partially overlap. NMR-derived structures of the binary complexes are shown in Fig. 3.

The HIF-1 α and CITED2 activation domains bind to TAZ1 with equal and very high affinity; the K_d for both binary complexes is 10 nM. To probe the mechanism by which CITED2 displaces the tightly bound HIF-1 α from TAZ1 we turned to NMR. The addition of HIF-1 α or CITED2 to ^{15}N -labeled TAZ1 causes shifts in the cross peaks of the TAZ1 ^1H - ^{15}N HSQC spectrum (Fig. 4a). Paradoxically, mixing the three proteins in a 1:1:1 ratio leads not to the expected 50% ratio of TAZ1:HIF-1 α and TAZ1:CITED2 complexes (expected because their K_d s are identical), but to a 100% population of the CITED2 complex (Fig. 4b).

Extensive further experiments, including fluorescence anisotropy measurements, stopped flow kinetics, and NMR relaxation measurements (Berlow et al. 2019) showed that the

system functions as a hypersensitive, unidirectional allosteric switch. CITED2 rapidly and efficiently displaces HIF-1 α from the TAZ1 complex (apparent $K_d = 0.2$ nM) but, conversely, HIF-1 α is highly ineffective at displacing TAZ1-bound CITED2 (apparent $K_d = 900$ nM). NMR was absolutely essential for reaching a molecular mechanistic understanding of this remarkable unidirectional process (Fig. 5). The NMR and kinetic data show that displacement of HIF-1 α from TAZ1 occurs by an associative process involving formation of a transient ternary complex. The inherent dynamics of HIF-1 α , CITED2, and TAZ1, in their free and bound states, are essential to this process. ^{15}N relaxation measurements show that binding of HIF-1 α dampens ps–ns time scale fluctuations of the TAZ1 backbone, increasing its order and prepaying part of the entropic cost for binding of CITED2 (Berlow et al. 2019). The N-terminal region of bound HIF-1 α is highly dynamic and makes fuzzy interactions with TAZ1, allowing the CITED2 α_A region to compete for binding and form an intermediate ternary complex. Docking of the CITED2 α_A region drives a conformational change in TAZ1 that disfavors the interaction with the C-terminal α_C helix of HIF-1 α , resulting in displacement of the HIF-1 α (Fig. 5). Unidirectionality arises because of the altered conformation of the TAZ1:CITED2 complex, which disfavors binding of the α_C helix of HIF-1 α , and because thermodynamic coupling between the LPQL- α_B and α_C interaction motifs of HIF-1 α , which are separated by a highly flexible linker, is weak (Berlow et al. 2017). This unidirectional allosteric process is entirely enabled by the unique dynamic and thermodynamic properties of the HIF-1 α and CITED2 IDPs and by the conformational dynamics and plasticity of the TAZ1 target domain.

The HIF-1 α -CITED2-TAZ1 system provides an excellent example of the essential role of NMR experiments in the elucidation of the structural, dynamic and mechanistic characteristics of important signaling systems in the cell. It is now recognized that IDPs function across almost the entire spectrum of cellular processes, on a molecular scale that ranges from localized interactions that control critical regulatory and signaling networks and extends to higher order assembly of macromolecular machines and even the organization of the cell itself, through formation of membraneless compartments. Despite the explosion of interest in IDPs in recent years, much remains to be learned about their biophysical properties and their biological functions. There can be little doubt that future research on IDPs will continue to yield many surprises.

In this article we have attempted to illustrate, from a personal perspective, the critical role that NMR played in the discovery and understanding of intrinsically disordered proteins. While single molecule FRET and other biophysical techniques are beginning to yield new and unexpected insights into the behavior of IDPs, NMR remains a central technology for characterization of their conformational ensembles, their dynamics, and their complexes with other biomolecules. We refer the reader to recent reviews for a detailed account of NMR methods for characterization of IDPs (Konrat 2014; Gibbs et al. 2017; Milles et al. 2018; Sciolino et al. 2019). NMR studies continue to reveal surprising new aspects of the roles of disorder in cellular processes. Disorder confers important advantages to the efficient operation of cells, but recognition of the existence and importance of IDPs would not have been possible without the unique insights provided by NMR.

Acknowledgements

We thank the many members of the Wright and Dyson laboratories, past and present, who have made essential contributions to the IDP work. Sadly, only a few examples of their contributions could be highlighted here. Our work on disordered proteins has been supported by the National Institutes of Health for over 30 years, through grants AI19499 and CA27489 (on anti-peptide antibodies), GM38794 (on peptides), NS14069 and AG21601 (on prions), DK34909 (on protein folding), GM57374 (on peptide NMR), CA96865 (on CBP), CA229652 (on HIF-1 α /CITED2 competition), GM113251 (on chaperone interactions), GM56879 and GM075995 (on protein dynamics and p53), and CA214054 (on CREB). We also thank the Skaggs Institute for Chemical Biology for long-term support of our work.

References

- Arai M, Ferreon JC, Wright PE (2012) Quantitative analysis of multisite protein-ligand interactions by NMR: binding of intrinsically disordered p53 transactivation subdomains with the TAZ2 domain of CBP. *J Am Chem Soc* 134:3792–3803 [PubMed: 22280219]
- Arai M, Sugase K, Dyson HJ, Wright PE (2015) Conformational propensities of intrinsically disordered proteins influence the mechanism of binding and folding. *Proc Natl Acad Sci USA* 112:9614–9619 [PubMed: 26195786]
- Bae SH, Dyson HJ, Wright PE (2009) Prediction of the rotational tumbling time for proteins with disordered segments. *J Am Chem Soc* 131:6814–6821 [PubMed: 19391622]
- Berlow RB, Dyson HJ, Wright PE (2017) Hypersensitive termination of the hypoxic response by a disordered protein switch. *Nature* 543:447–451 [PubMed: 28273070]
- Berlow RB, Martinez-Yamout MA, Dyson HJ, Wright PE (2019) role of backbone dynamics in modulating the interactions of disordered ligands with the TAZ1 domain of the CREB-binding protein. *Biochemistry* 58:1354–1362 [PubMed: 30775911]
- Borg M, Mittag T, Pawson T, Tyers M, Forman-Kay JD, Chan HS (2007) Polyelectrostatic interactions of disordered ligands suggest a physical basis for ultrasensitivity. *Proc Natl Acad Sci USA* 104:9650–9655 [PubMed: 17522259]
- Borgia A, Borgia MB, Bugge K, Kissling VM, Heidarsson PO, Fernandes CB, Sottini A, Soranno A, Buholzer KJ, Nettels D, Kragelund BB, Best RB, Schuler B (2018) Extreme disorder in an ultrahigh-affinity protein complex. *Nature* 555:61–66 [PubMed: 29466338]
- Brown JE, Klee WA (1971) Helix-coil transition of the isolated amino terminus of ribonuclease. *Biochemistry* 10:470–476 [PubMed: 5543977]
- Dames SA, Martinez-Yamout M, De Guzman RN, Dyson HJ, Wright PE (2002) Structural basis for Hif-1 α /CBP recognition in the cellular hypoxic response. *Proc Natl Acad Sci USA* 99:5271–5276 [PubMed: 11959977]
- Daniels A, Williams RJP, Wright PE (1976) Nuclear magnetic resonance studies of the adrenal gland and some other organs. *Nature* 261:321–323 [PubMed: 1272407]
- Daniels AJ, Williams RJP, Wright PE (1978) The character of the stored molecules in chromaffin granules of the adrenal medulla: a nuclear magnetic resonance study. *Neuroscience* 3:573–585 [PubMed: 692872]
- De Guzman RN, Martinez-Yamout M, Dyson HJ, Wright PE (2004) Interaction of the TAZ1 domain of CREB-binding protein with the activation domain of CITED2: regulation by competition between intrinsically unstructured ligands for non-identical binding sites. *J Biol Chem* 279:3042–3049 [PubMed: 14594809]
- Demarest SJ, Martinez-Yamout M, Chung J, Chen H, Xu W, Dyson HJ, Evans RM, Wright PE (2002) Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* 415:549–553 [PubMed: 11823864]
- Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z (2001) Intrinsically disordered protein. *J Mol Graph Model* 19:26–59 [PubMed: 11381529]
- Dunker AK, Babu MM, Barbar E, Blackledge M, Bondos SE, Dosztanyi Z, Dyson HJ, Forman-Kay J, Fuxreiter M, Gsponer J, Han KH, Jones DT, Longhi S, Metallo SJ, Nishikawa K, Nussinov R,

Obradovic Z, Pappu RV, Rost B, Selenko P, Subramaniam V, Sussman JL, Tompa P, Uversky VN (2013) What's in a name? Why these proteins are intrinsically disordered. *Intrins Disord Prot* 1:e24157

Dyson HJ, Wright PE (1991) Defining solution conformations of small linear peptides. *Annu Rev Biophys Chem* 20:519–538 [PubMed: 1867725]

Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6:197–208 [PubMed: 15738986]

Dyson HJ, Wright PE (2017) How does your protein fold? Elucidating the apomyoglobin folding pathway. *Acc Chem Res* 50:105–111 [PubMed: 28032989]

Dyson HJ, Cross KJ, Houghten RA, Wilson IA, Wright PE, Lerner RA (1985) The immunodominant site of a synthetic immunogen has a conformational preference in water for a Type-II reverse turn. *Nature* 318:480–483 [PubMed: 4069219]

Dyson HJ, Rance M, Houghten RA, Lerner RA, Wright PE (1988a) Folding of immunogenic peptide fragments of proteins in water solution. I. Sequence requirements for the formation of a reverse turn. *J Mol Biol* 201:161–200 [PubMed: 2843644]

Dyson HJ, Rance M, Houghten RA, Wright PE, Lerner RA (1988b) Folding of immunogenic peptide fragments of proteins in water solution. II. The nascent helix. *J Mol Biol* 201:201–217 [PubMed: 3418697]

Dyson HJ, Gippert GP, Case DA, Holmgren A, Wright PE (1990) Three-dimensional solution structure of the reduced form of *Escherichia coli* thioredoxin determined by nuclear magnetic resonance spectroscopy. *Biochemistry* 29:4129–4136 [PubMed: 2193685]

Dyson HJ, Merutka G, Waltho JP, Lerner RA, Wright PE (1992a) Folding of peptide fragments comprising the complete sequence of proteins. Models for initiation of protein folding. I. Myohemerythrin. *J Mol Biol* 226:795–817 [PubMed: 1507227]

Dyson HJ, Sayre JR, Merutka G, Shin H-C, Lerner RA, Wright PE (1992b) Folding of peptide fragments comprising the complete sequence of proteins: models for the initiation of protein folding. II. Plastocyanin. *J Mol Biol* 226:819–835 [PubMed: 1507228]

Eliezer D, Yao J, Dyson HJ, Wright PE (1998) Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat Struct Biol* 5:148–155 [PubMed: 9461081]

Elkins JM, Hewitson KS, McNeill LA, Seibel JF, Schlemminger I, Pugh CW, Ratcliffe PJ, Schofield CJ (2003) Structure of factorinhibiting hypoxia-inducible factor (HIF) reveals mechanism of oxidative modification of HIF-1 α . *J Biol Chem* 278(3):1802–1806 [PubMed: 12446723]

Freedman SJ, Sun ZY, Poy F, Kung AL, Livingston DM, Wagner G, Eck MJ (2002) Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1 α . *Proc Natl Acad Sci USA* 99:5367–5372 [PubMed: 11959990]

Gibbs EB, Cook EC, Showalter SA (2017) Application of NMR to studies of intrinsically disordered proteins. *Arch Biochem Biophys* 628:57–70 [PubMed: 28502465]

Gillespie JR, Shortle D (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *J Mol Biol* 268:158–169 [PubMed: 9149149]

Gronenborn AM, Bax A, Wingfield PT, Clore GM (1989) A powerful method of sequential proton resonance assignment in proteins using relayed ^{15}N - ^1H multiple quantum coherence spectroscopy. *FEBS Lett* 243:93–98 [PubMed: 2537752]

Ivan M, Kondo K, Yang H, Kim W, Valiando J, Ohh M, Salic A, Asara JM, Lane WS, Kaelin WG (2001) HIF α targeted for VHL-mediated destruction by proline hydroxylation: implications for O_2 sensing. *Science* 292(5516):464–468 [PubMed: 11292862]

Jaakkola P, Mole DR, Tian Y-M, Wilson MI, Gielbert J, Gaskell SJ, Kriegsheim A, Hebestreit HF, Mukherji M, Schofield CJ, Max-well PH, Pugh CW, Ratcliffe PJ (2001) Targeting of HIF- α to the von Hippel-Lindau ubiquitylation complex by O_2 -regulated prolyl hydroxylation. *Science* 292(5516):468–472 [PubMed: 11292861]

Janin J, Sternberg MJ (2013) Protein flexibility, not disorder, is intrinsic to molecular recognition. *F1000 Biol Rep* 5:2 [PubMed: 23361309]

- Jeener J, Meier BH, Bachman P, Ernst RR (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy. *J Chem Phys* 71:4546–4553
- Jensen MR, Zweckstetter M, Huang J-R, Blackledge M (2014) Exploring free-energy landscapes of intrinsically disordered proteins at atomic resolution using nmr spectroscopy. *Chem Rev* 114:6632–6660 [PubMed: 24725176]
- Konrat R (2014) NMR contributions to structural dynamics studies of intrinsically disordered proteins. *J Magn Reson* 241:74–85 [PubMed: 24656082]
- Kriwacki RW, Hengst L, Tennant L, Reed SI, Wright PE (1996) Structural studies of p21^{Waf1/Cip1/Sdi1} in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proc Natl Acad Sci USA* 93:11504–11509 [PubMed: 8876165]
- Krois AS, Ferreon JC, Martinez-Yamout MA, Dyson HJ, Wright PE (2016) Recognition of the disordered p53 transactivation domain by the transcriptional adapter zinc finger domains of CREB-binding protein. *Proc Natl Acad Sci USA* 113:E1853–E1862 [PubMed: 26976603]
- Krois AS, Dyson HJ, Wright PE (2018) Long-range regulation of p53 DNA binding by its intrinsically disordered N-terminal transactivation domain. *Proc Natl Acad Sci USA* 115:E11302–E11310 [PubMed: 30420502]
- Land D, Peet DJ, Whelan DA, Gorman JJ, Whitelaw ML (2002) Asparagine hydroxylation on the HIF transactivation domain: a hypoxic switch. *Science* 295:858–861 [PubMed: 11823643]
- Lee MS, Gippert G, Soman KY, Case DA, Wright PE (1989) Three-dimensional solution structure of a single zinc finger binding domain. *Science* 245:635–637 [PubMed: 2503871]
- Lee CW, Arai M, Martinez-Yamout MA, Dyson HJ, Wright PE (2009) Mapping the interactions of the p53 transactivation domain with the KIX domain of CBP. *Biochemistry* 48:2115–2124 [PubMed: 19220000]
- Lietzow MA, Jamin M, Dyson HJ, Wright PE (2002) Mapping long-range contacts in a highly unfolded protein. *J Mol Biol* 322:655–662 [PubMed: 12270702]
- Love JJ, Li XA, Case DA, Giese K, Grosschedl R, Wright PE (1995) Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature* 376:791–795 [PubMed: 7651541]
- Love JJ, Li X, Chung J, Dyson HJ, Wright PE (2004) The LEF-1 HMG domain undergoes a disorder-to-order transition upon complex formation with cognate DNA. *Biochemistry* 43:8725–8734 [PubMed: 15236581]
- Marsh JA, Forman-Kay JD (2012) Ensemble modeling of protein disordered states: experimental restraint contributions and validation. *Prot Struct Func Bioinform* 80:556–572
- Milles S, Salvi N, Blackledge M, Jensen MR (2018) Characterization of intrinsically disordered proteins and their dynamic complexes: from in vitro to cell-like environments. *Prog NMR Spect* 109:79–100
- Mohana-Borges R, Goto NK, Kroon GJA, Dyson HJ, Wright PE (2004) Structural characterization of unfolded states of apomyoglobin using residual dipolar couplings. *J Mol Biol* 340:1131–1142 [PubMed: 15236972]
- Moore JM, Case DA, Chazin WJ, Gippert GP, Havel TF, Powls R, Wright PE (1988) Three-dimensional solution structure of plastocyanin from the green alga *Scenedesmus obliquus*. *Science* 240:314–317 [PubMed: 3353725]
- Neri D, Billeter M, Wider G, Wüthrich K (1992) NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science* 257:1559–1563 [PubMed: 1523410]
- Nielsen JT, Mulder FAA (2018) POTENCI: prediction of temperature, neighbor and pH-corrected chemical shifts for intrinsically disordered proteins. *J Biomol NMR* 70:141–165 [PubMed: 29399725]
- Niman HL, Houghten RA, Walker LE, Reisfeld RA, Wilson IA, Hogle JM, Lerner RA (1983) Generation of protein-reactive antibodies by short peptides is an event of high frequency: implications for the structural basis of immune recognition. *Proc Natl Acad Sci USA* 80:4949–4953 [PubMed: 6192445]
- Radhakrishnan I, Pérez-Alvarado GC, Parker D, Dyson HJ, Montminy MR, Wright PE (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. *Cell* 91:741–752 [PubMed: 9413984]

- Radhakrishnan I, Pérez-Alvarado GC, Dyson HJ, Wright PE (1998) Conformational preferences in the Ser¹³³-phosphorylated and non-phosphorylated forms of the kinase inducible transactivation domain of CREB. *FEBS Lett* 430:317–322 [PubMed: 9688563]
- Reymond MT, Merutka G, Dyson HJ, Wright PE (1997) Folding propensities of peptide fragments of myoglobin. *Protein Sci* 6:706–716 [PubMed: 9070453]
- Romero P, Obradovic Z, Kissinger CR, Villafranca JE, Dunker AK (1997) Identifying disordered regions in proteins from amino acid sequences. In: *Proceedings of IEEE International Conference on Neural Networks* pp 90–95
- Schmid T, Zhou J, Brune B (2004) HIF-1 and p53: communication of transcription factors under hypoxia. *J Cell Mol Med* 8:423–431 [PubMed: 15601571]
- Schwarzinger S, Kroon GJA, Foss TR, Wright PE, Dyson HJ (2000) Random coil chemical shifts in acidic 8 M urea: implementation of random coil chemical shift data in NMRView. *J Biomol NMR* 18:43–48 [PubMed: 11061227]
- Schwarzinger S, Kroon GJA, Foss TR, Chung J, Wright PE, Dyson HJ (2001) Sequence dependent correction of random coil NMR chemical shifts. *J Am Chem Soc* 123:2970–2978 [PubMed: 11457007]
- Schwarzinger S, Wright PE, Dyson HJ (2002) Molecular hinges in protein folding: the urea-denatured state of apomyoglobin. *Biochemistry* 41:12681–12686 [PubMed: 12379110]
- Sciolino N, Burz DS, Shekhtman A (2019) In-cell NMR spectroscopy of intrinsically disordered proteins. *Proteomics* 19:1800055
- Semenza GL (2004) Hydroxylation of HIF-1: oxygen sensing at the molecular level. *Physiology* 19:176–182 [PubMed: 15304631]
- Sugase K, Dyson HJ, Wright PE (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447:1021–1025 [PubMed: 17522630]
- Takeuchi K, Arthanari H, Shimada I, Wagner G (2015) Nitrogen detected TROSY at high field yields high resolution and sensitivity for protein NMR. *J Biomol NMR* 63:323–331 [PubMed: 26497830]
- Tamiola K, Acar B, Mulder FAA (2010) Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *J Am Chem Soc* 132:18000–18003 [PubMed: 21128621]
- Theillet FX, Binolfi A, Frembgen-Kesner T, Hingorani K, Sarkar M, Kyne C, Li C, Crowley PB, Gierasch L, Pielak GJ, Elcock AH, Gershenson A, Selenko P (2014) Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs). *Chem Rev* 114:6661–6714 [PubMed: 24901537]
- Theillet F-X, Binolfi A, Bekei B, Martorana A, Rose HM, Stuijver M, Verzini S, Lorenz D, van Rossum M, Goldfarb D, Selenko P (2016) Structural disorder of monomeric α -synuclein persists in mammalian cells. *Nature* 530:45–50 [PubMed: 26808899]
- Tomba P, Fuxreiter M (2008) Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem Sci* 33:2–8 [PubMed: 18054235]
- Tsvetkov P, Reuven N, Prives C, Shaul Y (2009) Susceptibility of p53 unstructured N terminus to 20 S proteasomal degradation programs the stress response. *J Biol Chem* 284:26234–26242 [PubMed: 19617345]
- van der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, Fuxreiter M, Gough J, Gsponer J, Jones DT, Kim PM, Kriwacki RW, Oldfield CJ, Pappu RV, Tompa P, Uversky VN, Wright PE, Babu MM (2014a) Classification of intrinsically disordered regions and proteins. *Chem Rev* 114:6589–6631 [PubMed: 24773235]
- van der Lee R, Lang B, Kruse K, Gsponer J, Sánchez de Groot N, Huynen Martijn A, Matouschek A, Fuxreiter M, Babu MM (2014b) Intrinsically disordered segments affect protein half-life in the cell and during evolution. *Cell Rep* 8:1832–1844 [PubMed: 25220455]
- Viles JH, Donne DG, Kroon GJA, Prusiner SB, Cohen FE, Dyson HJ, Wright PE (2001) Local structural plasticity of the prion protein. Analysis of NMR relaxation dynamics. *Biochemistry* 40:2743–2753 [PubMed: 11258885]
- Wilson IA, Wiley DC, Skehel JJ (1981) Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å. *Nature* 289:366–373 [PubMed: 7464906]
- Wokaun A, Ernst RR (1977) Selective detection of multiple quantum transitions in NMR by two-dimensional spectroscopy. *Chem Phys Lett* 52:407–412

- Wright PE, Dyson HJ (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol* 293:321–331 [PubMed: 10550212]
- Wright PE, Dyson HJ (2015) Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* 16:18–29 [PubMed: 25531225]
- Wright PE, Dyson HJ, Lerner RA (1988) Conformation of peptide fragments of proteins in aqueous solution: implications for initiation of protein folding. *Biochemistry* 27:7167–7175 [PubMed: 3061450]
- Yao J, Dyson HJ, Wright PE (1997) Chemical shift dispersion and secondary structure prediction in unfolded and partly folded proteins. *FEBS Lett* 419:285–289 [PubMed: 9428652]
- Yao J, Chung J, Eliezer D, Wright PE, Dyson HJ (2001) NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding. *Biochemistry* 40:3561–3571 [PubMed: 11297422]

Fig. 1. ^1H - ^{15}N HSQC spectra of disordered partners for **a** the DNA-binding domain of LEF-1 (Love et al. 1995, 2004), black—free LEF-1, red, LEF-1 bound to a 15 base-pair cognate DNA sequence. Selected resonances shifted between the free and bound spectra are connected and labeled. Resonances that are absent from the spectrum of the free protein but appear in the bound spectrum are labeled in green. **b** p21 residues 9–85 bound to the cyclin-dependent kinase Cdk2 (Kriwacki et al. 1996), green—resonances found at the same chemical shift in the free protein and the Cdk2 complex; red, new cross peaks that appear in the complex. **c** free pKID (top) and pKID bound to the KIX domain (bottom) (Radhakrishnan et al. 1997). Significantly shifted resonances are labeled. The intense, narrow resonances that appear between 7.8 and 8.5 ppm arise from residues in the unstructured terminal regions of pKID in the complex. The box at lower right shows (aliased) arginine side chain resonances. **b**, **c** reproduced with permission from references (Kriwacki et al. 1996; Radhakrishnan et al. 1997)

Fig. 2. Structure of a HIF-1 α interaction domain fragment (residues 798–805) bound to two targets, **a** TAZ1 domain of CBP (Dames et al. 2002), and **b** asparagine hydroxylase FIH (Elkins et al. 2003). The backbone of each partner protein is shown as a gray ribbon, the backbone of the HIF-1 α fragment in green and side chains in yellow. The asparagine residue that functions as a redox switch is highlighted. Figure adapted from reference (Dyson and Wright 2005) with permission

Fig. 3. Single NMR structures of TAZ1 in complex with HIF-1 α and CITED2. **a** TAZ1:HIF-1 α complex (Dames et al. 2002). **b** TAZ1:CITED2 complex (De Guzman et al. 2004). The HIF-1 α and CITED2 sequences are color-coded from N (green) to C (red). **c** Superposition of the two structures on the TAZ1 structure (Berlow et al. 2017). The gray surface shows TAZ1, with HIF-1 α in gold and CITED2 in blue

Fig. 4. ^1H - ^{15}N HSQC spectra of TAZ1 and complexes (Berlow et al. 2017). **a** Superimposed spectra of ^{15}N -labeled TAZ1 alone (black), and in complex with HIF-1 α (gold) and CITED2 (blue). Assignments are shown for selected cross peaks. The tryptophan indole resonance is shown as an inset in the bottom left corner. **b** Superimposed spectra of ^{15}N -labeled TAZ1 in complex with HIF-1 α (gold) and with CITED2 (blue). The red spectrum (with fewer contours displayed for visibility) was obtained from a solution containing equimolar amounts of ^{15}N -labeled TAZ1, HIF-1 α and CITED2. Adapted from reference (Berlow et al. 2017), with permission

Fig. 5.

Schematic mechanism for displacement of HIF-1α (gold) by CITED2 (blue) from TAZ1 (gray) (Berlow et al. 2017). g^C and g^H represent thermodynamic coupling between the α_A and LPEL motifs of CITED2 and between the LPQL- α_B and α_C motifs of HIF-1α, respectively. Reproduced from reference (Berlow et al. 2017), with permission