Published in final edited form as:

Curr Phys Chem. 2012 January ; 2(1): 107–114.

Folding of Intrinsically Disordered Protein Phosphatase 1 Regulatory Proteins

Wolfgang Peti^{1,*}, Angus C. Nairn², and Rebecca Page³

¹Department of Molecular Pharmacology, Physiology and Biotechnology & Department of Chemistry, Brown University, Providence, RI 02912, USA

²Department of Psychiatry, Yale University School of Medicine, New Haven, CT 06508, USA

³Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, RI 02912, USA

Abstract

Intrinsically disordered but biologically active proteins, commonly referred to as IDPs, are readily identified in many biological systems and play critical roles in multiple protein regulatory processes. While disordered in their unbound states, IDPs often, but not always, fold upon binding with their protein interaction partners. Here, we discuss how a class of IDPs directs the targeting, specificity and activity of Protein Phosphatase 1 (PP1). PP1 is major ser/thr phosphatase that plays a critical role in a broad range of biological processes, from muscle contraction to memory formation. In the cell, PP1 is regulated through its interaction with more than 200 regulatory proteins, the majority of which are IDPs. Critically, these PP1:regulatory protein holoenzyme complexes confer specificity to PP1 and are thus the functional forms of the PP1 enzyme *in vivo*. Furthermore, we discuss the distinct modes of interaction utilized by IDPs to complex with their protein binding partners. We subsequently show, by integrating multiple biophysical tools, that the majority of IDPs that regulate PP1, prefer a conformational selection model.

Keywords

Intrinsically disordered proteins; proteins; protein phosphatase 1; folding

INTRODUCTION

Biology has long been a source of inspiration for the physical sciences, and the desire to understand biological mechanisms have led to a more clear understanding of numerous fundamental physical principles. Until recently, the central dogma of protein science was that a well-defined 3-dimensional structure is required for knowledge of protein function [1–4]. Thus, the initial interest in the protein folding problem was focused on understanding the physical principles that drive proteins into a single, folded conformation from an ensemble of denatured/unstructured states. However, the results of these studies are applicable to only a subset of proteins, i.e., those that have well-defined 3-dimensional structures in their native state. Although structurally dynamic but functional peptides and hormones, such as glucagon and insulin [5, 6], have been known since the mid 1970's, it has only recently been recognized that *intrinsically disordered but biologically active proteins* are not only

prevalent in biological systems, but also play critical roles in protein:protein interactions and protein signaling [3]. These proteins are disordered in their unbound states, but are able to fold into at least one and perhaps multiple conformations upon interacting with their protein binding partners [1, 7].

Intrinsically disordered proteins (IDPs) lack the common hydrophobic residues, such as Val, Leu, Ile, Met, Phe, Trp and Tyr, which form the central hydrophobic core of folded proteins and, as a consequence, IDPs lack a single, stable fold. Instead, IDPs are enriched in charged and often short side chain amino acids, such as Gln, Ser, Pro, Glu, Lys and Gly. Nevertheless, while IDPs are highly flexible, they exhibit a variety of dynamics that play critical roles in their ability to bind targeting proteins. Indeed, for IDPs, flexibility is the major characteristic that allows them to form novel unexpected complexes that are essential for a multitude of vital biological processes.

Many eukaryotic proteins contain such intrinsically disordered segments, with lengths ranging from a few tens to hundreds of residues, or, in some cases, the entire protein can be completely unstructured. Computational studies have shown that ~25% of eukaryotic proteins contain a disordered fragment longer than 30 residues [8, 9]. The recent focus on such unfolded and partially folded proteins comes from post-genomic biochemical methodology: proteins are now investigated from the gene and no longer from biochemical isolation, a process that had tended to bias protein discovery and structure determination towards well-structured and stable proteins.

IDPs also play essential roles in critical regulatory processes because of their ability to form a multitude of transient multi-partner complexes. Evidently, the potential role of residual secondary and tertiary structure and the dynamics of IDPs require new, innovative methods of investigation that combine both established and novel techniques for data measurement and evaluation. An outstanding question in the field is to understand, at atomic resolution, the parameters that drive the coupling between folding and binding of IDPs to their target proteins.

ATOMIC RESOLUTION ANALYSIS OF IDPs

A large number of experimental techniques can be used to derive structural information for the folded state of proteins. This is vastly different for the unstructured state, for which fewer methods exist to obtain atomic resolution information [10]. Dynamic unfolded regions cannot be studied by X-ray crystallography, as unstructured proteins do not crystallize (even if they did, the ensemble of structures in the crystal would misrepresent the full conformational ensemble). Among all biophysical tools, NMR spectroscopy is the only tool that provides *atomic resolution data* of a complete ensemble in the absence of a well-defined structure [11, 12]. ^1H , ^{15}N , ^{13}C resonance assignments have been reported for a large number of IDPs [7, 10, 12–15]. The availability of NMR multidimensional methods to obtain site-specific resonance assignments for IDPs forms the basis for a detailed analysis of NMR parameters such as chemical shifts and heteronuclear relaxation rates.

Traditional NMR tools, such as chemical shift analysis, paramagnetic relaxation enhancements (PRE), nuclear Overhauser effects (nOe), residual dipolar couplings (RDC), nitrogen-15 relaxation, among others, have been employed to characterize the conformational space of IDPs (Fig. 1). The analysis of the NMR parameters of IDPs is challenging, as they reflect an average taken over the ensemble of conformers populated in the unstructured state of a protein. Therefore, NMR data are currently interpreted with new, adapted approaches for the dynamic nature of these proteins. Much effort has been devoted to the development of mathematical techniques for the description of accurate structural ensembles [16–22].

Nevertheless, it must also be mentioned that comprehensive statistics of the data based on the NMR analysis of IDPs are currently less well defined than for well-structured proteins. Consequently, the definition of “preferred conformation” is subjective and often differs between different laboratories. For example, we consider a propensity of 20% to be significant; however, other laboratories consider much lower propensities (< 10%) of a preferred conformation to be significant. One other interesting experimental observation must also be mentioned in regards to preformed α -helical secondary structure. In folded proteins, (i,i+1) H^N - H^N NOEs (distance is ~ 2.8 Å) are typically key for identifying α -helical structures [23]. We have looked in all of our published and a large number of unpublished IDPs for (i,i+1) H^N - H^N NOEs in pre-populated α -helical regions. However, so far, we were able to experimentally detect these (i,i+1) H^N - H^N NOEs using a 3D ^{15}N -resolved [1H , 1H] NOESY spectrum, in only a single IDP (Inhibitor-2, see below), and only in a single region that was $\sim 80\%$ populated. This clearly shows that in folded proteins, α -helices are commonly nearly 100% populated. Lastly, many IDPs show slow timescale cis/trans proline isomerization, where the minor conformation is usually 5–15% populated. While this is of interest, no systematic, detailed analysis has so far been reported. During the next years it will be important to further analyze the rapidly growing pool of NMR data for IDPs to understand their common structural hallmarks.

Lastly, additional biophysical techniques, including single-molecule fluorescence spectroscopy [24] as well as small angle X-ray scattering (SAXS) [10, 25, 26] have also emerged as additional complimentary and powerful tools for studying IDPs.

IDPs – BINDING TO TARGETS: FOLDING?

A large number of published studies of IDPs have now shown that IDPs can adopt a variety of structures. While IDPs can exist as true random coil polymers, more often they exhibit preferences for elements of transient secondary or tertiary structure. Clearly, this pre-organization, or “preferred conformation”, and spatial restriction in IDPs can expose primary interaction sites to enable faster and more effective binding as well as folding to target molecules [27–29].

There are two potential pathways that lead to the selection of a single folded conformation when IDPs bind to a folded scaffold. First, the IDP has an *intrinsic preference* for its binding conformer and therefore the interacting protein is only a scaffold as it does not direct the formation of the preferred conformation. In this “conformational selection” model of binding, the IDP adopts the bound-state conformation in its free state in a limited population [30]. Only these pre-populated binding-competent conformations will interact with the target protein and thus shift the equilibrium toward the bound state.

Second, the interacting protein *guides the folding process* of the IDP and is needed to identify the folded, stable conformer. IDPs can also follow the second proposed model, often referred to as the “induced fit” model. In this model, the bound conformation becomes energetically accessible only in the presence of the binding partner [31] and is either not detected in the unbound state or is only populated in very low amounts, undetectable using current techniques.

During the last 5 years, it has become evident that both mechanisms, conformational selection and induced fit, are important for IDP function [32–36]. Indeed, the interaction of an IDP that is undergoing a significant folding-upon-binding conformational change likely involves an induced-fit component, as there is only a minute probability that the necessary bound-state conformation with the correct 3D geometry occurs simultaneously in an rapidly interchanging ensemble of structures, typical for an unbound IDP. Nevertheless, conformational selection will be equally important, as individual elements of secondary or

tertiary structure (or both) are fractionally populated and selected for during the binding process.

An additional recently introduced mechanism that some IDPs use is the formation of dynamic complexes between IDPs and their binding partners. In these dynamic complexes, IDPs continue to behave as an ensemble of rapidly inter-converting conformations, mostly unstructured, even when bound to their targeting protein [37].

In this review we specifically focus on the analysis of functionally related IDPs, *e.g.* IDPs that bind to a single target protein, Protein Phosphatase 1 (PP1). This analysis reveals a conservation of structure-function motifs in IDPs that are common in folded proteins, as well as additional novel mechanisms that are employed for binding.

PROTEIN PHOSPHATASE 1 (PP1)

Protein Phosphatase 1 (PP1, ~38.5 kDa) is the best-characterized serine/threonine phosphatase [38]. Mammalian genomes contain three different genes that encode four distinct catalytic subunits of PP1: PP1 α , PP1 β / δ and the splice variants PP1 γ_1 and PP1 γ_2 , which differ in the amino acid compositions of the N- and C-terminal extremities. Remarkably, free PP1 has only little intrinsic specificity for substrates. Rather, it is tightly regulated by its interaction with >200 known targeting proteins [38], proteins that: 1) localize PP1 to distinct regions of the cell and 2) modulate its substrate specificity. Furthermore, PP1 specific inhibitor proteins, such as inhibitor-2 and DARPP-32, bind and block the PP1 active site, which contains two metal ions (Mn²⁺, when heterologously expressed in bacteria [25, 39]) that are required for PP1 activity. The catalytic site of PP1 is at the intersection of three potential substrate binding grooves: the hydrophobic, acidic and C-terminal grooves (Fig. 2) [38]. The majority of PP1 regulators and also some substrates interact with PP1 *via* a primary PP1-binding motif, the RVxF motif, which generally conforms to the consensus sequence [K/R][K/R][V/I][x][F/W], where x is any residue other than Phe, Ile, Met, Tyr, Asp, or Pro [40–42]. While the RVxF interaction is necessary for regulatory protein binding, it does not influence PP1 enzymatic activity as it is 20 Å away from the active site (compare magenta versus pink surface patches in Fig. 2). Recently identified additional docking sites [26, 43–45], such as the SILK and MyPhoNE motifs [40], also have essential roles in the regulation of PP1 substrate specificity (Fig. 2). However, a detailed understanding and, more importantly, the possibility of predicting how the >200 PP1 targeting proteins direct PP1 specificity from sequence alone, is still entirely missing.

Interestingly, a large number (>70%) of PP1 regulators are predicted to be IDPs. Thus, the PP1 regulatory system provides the unique possibility to study how a large number of IDPs bind and regulate a single protein, PP1, and in turn a myriad of critical biological functions. Here we describe the interaction of two targeting proteins (MYPT1, spinophilin) and one inhibitor protein (Inhibitor-2), as well as the simultaneous interaction of a targeting and inhibitor protein (spinophilin and Inhibitor-2) to PP1.

PP1:SPINOPHILIN – THE HIGHLY DISORDERED IDP

A high percentage (62%) of the spinophilin PP1-binding domain, spinophilin residues 417–494 (hereafter referred to as spinophilin), consists of disorder inducing amino acids (16 Glu, 7 Ser, 8 Pro, 5 Arg, 7 Asp, 2 Lys and 3 Gly), while only 23% have a hydrophobic character (4 Val, 3 Leu, 3 Ile, 1 Met, 2 Phe and 4 Tyr). Using local and long range probes for secondary and tertiary structure it was identified that unbound spinophilin has no preferred conformation in solution; rather it behaves much like a highly disordered peptide (Fig. 3A). This is remarkable as spinophilin interacts tightly and specifically with PP1, with a K_d of ~8 nM [44].

As stated NMR spectroscopy has been the premier tool for the structural analysis of IDPs. As IUPs adopt rapid interchangeable conformations in solution, only a very small number of ^1H - ^1H NOEs (and these are mainly sequential) can be detected. This is a problem, as the most common method for NMR structure determination is based on the collection of a large number of ^1H - ^1H NOE distances, which are subsequently used to fold the protein. However, numerous exciting approaches for the calculation of 3-dimensional IDP ensemble structures have been reported during the last few years, including those by Martin Blackledge [17, 18], Julie Forman-Kay [46] and Michele Vendruscolo [47]. In our studies, the program ENSEMBLE (Julie Forman-Kay) [19, 46] was used to investigate the unbound ensemble structure(s) of a number of IDPs, including spinophilin. ENSEMBLE assigns populations to a large pool of pre-generated input structures that satisfy experimental restraints. As a complete sampling of the conformational space is impossible, ENSEMBLE uses an iterative approach, in which an experimental data based conformer selection method is used to increase the pool of available ensemble structures for subsequent ENSEMBLE runs.

For spinophilin, ENSEMBLE detected two major populated clusters (Fig. 3B). In both clusters, a ~30% populated helix was detected at the C-terminus. In the lower 18% populated cluster, tertiary interactions that clearly resembled β -strand contacts were detected. Most of the ensemble structures (68%) showed a large number of tertiary contacts, particularly near the N- and C-termini and thus adopted a fairly compact conformation. IDPs often have at least one cluster with tertiary structures that involve the N- and C-termini [19, 26]. This goes hand-and-hand with a postulation by Dill [48], which states that contacts involving termini should be more likely due to the excluded volume of the polymer chain.

Recently, the PP1:spinophilin holoenzyme crystal structure was reported [44, 49]. This allows for a direct comparison between the bound and unbound IDP. The structures of the unbound spinophilin ensemble reveal that many of the intramolecular structural features observed in PP1-bound spinophilin are preformed in unbound spinophilin (Figs. 3B, 3C, black boxes). Thus, for spinophilin, it has become evident that conformational selection is likely key for PP1 binding; i.e. PP1 stabilizes the preformed lowly populated fold of spinophilin during binding to PP1.

PP1:INHIBITOR-2 – RESIDUAL FLEXIBILITY DEFINES THE COMPLEX, BUT NOT THE BIOLOGICAL FUNCTION

The PP1:Inhibitor-2 (hereafter referred to as I-2) crystal structure was the second determined PP1 holoenzyme complex structure [43] (Fig. 4). However, electron density for I-2 in the PP1:I-2 complex was visible for only ~25% of the I-2 protein and, thus, I-2 remains largely disordered, even upon binding to PP1. This intrinsic disorder of I-2 makes it challenging to evaluate the overall structure of the complex. Nevertheless, this study identified that I-2 interacts with PP1 at three different sites. The first site is the anticipated RVxF binding site. A second site (residues $^{12}\text{SILK}^{16}$), which had never previously been observed, conforms to the general sequence [GS]IL[RK] and was subsequently termed SILK motif [40]. The third site is comprised of I-2 residues 130–169 which form an α -helix that lies across the PP1 active site and block access for substrates.

In follow-up studies, the residual structure of the PP1:I-2 complex was assessed by NMR spectroscopy and SAXS [26, 50]. 2D [^1H , ^{15}N] HSQC/TROSY spectra of unbound I-2 and NMR-active-labeled I-2 in complex with unlabeled PP1 were directly compared. In complete agreement with the crystal structure, chemical shift changes in I-2 were only detected for the I-2 residues directly interacting with PP1 in the PP1:I-2 complex, i.e. those that were not flexible and became rigid upon binding to PP1. In contrast, NMR signals of I-2 residues in regions for which no electron density was observed retained the characteristic

traits of highly dynamic proteins, including very narrow line-widths typical for fast reorientation in solution. Thus, as was done for unbound spinophilin, experimentally determined NMR parameters were used as the input for the program ENSEMBLE to determine a model of the full PP1:I-2 complex. The most remarkable feature of this model is a long loop connecting I-2 residues 55–127. Although this loop is somewhat heterogeneous, with varying conformations between the different ensemble conformers, a significant fraction of structures adopt a distinct loop structure. This loop contains residue Thr72 which is essential for the biological function of the PP1:I-2 complex, as phosphorylation of Thr72 by GSK-3b reactivates the PP1:I-2 complex [51]. Furthermore, the ENSEMBLE derived model for the PP1:I-2 holoenzyme was experimentally validated using SAXS measurements. The calculated PP1:I-2 envelope nicely represented the largest population of the ensemble models with the distinct loop structure of I-2 extended away from the center of PP1.

So how is the PP1:I-2 complex formed? 53% of I-2 is comprised of disorder inducing amino acids, while 24% have a hydrophobic character. Indeed, I-2 even has a tryptophan residue (part of the RVxF sequence), which is rarely found in IDPs. Based on extensive NMR analysis, three preformed secondary structural elements can be detected in unbound I-2 including residues 36–42, 96–106, and 127–154, with these residues populating the α -helical conformation space ~30, 48, and 67% of the time, respectively. I-2 residues 132–138, which belong to the long helix spanning residues 127–154, have >98% α -helical character, as measured using chemical shifts. Indeed, even (i,i+1) H^N - H^N NOEs can be experimentally detected for these residues. Thus clearly, the most notable pre-populated secondary structure element is the 67% populated α -helix in I-2. This helix corresponds directly to the long α -helix present in the PP1:I-2 complex, which folds over the PP1 active site and is critical for PP1 inhibition. This is a prime example for a conformational selection folding event. In striking contrast the other two transient pre-populated helices are in regions that stay flexible even after PP1 binding. This clearly shows that not all pre-populated secondary structures elements automatically guide a direct binding event.

PP1:SPINOPHILIN:INHIBITOR-2 – THE BATTLE OVER THE RVxF PP1-BINDING SITE LEADS TO EVEN MORE FLEXIBILITY

It was recently shown that pairs of targeting and inhibitor proteins can bind PP1 simultaneously [52–54]. This leads to an additional layer of complexity in PP1 regulation.

One heterotrimeric PP1 complex of interest is PP1:spinophilin:I-2 (PSI). A detailed biochemical, NMR spectroscopy and SAXS analysis showed that in the PSI complex, I-2 releases from the RVxF binding site while staying bound *via* the SILK motif and the long I-2 α -helix, which binds the PP1 active site [25]. This results in an increase in the size of the flexible I-2 loop in the PSI complex as compared to the I-2 complex (Fig. 4). Conversely, while the RVxF motif of spinophilin remains bound to PP1 in the PSI complex, the shorter α -helix that connects the spinophilin PP1 binding and PDZ domains dissociates from PP1 in order to accommodate bound I-2.

As this is the first PP1 triple complex that has been investigated in molecular detail, it is too early to draw detailed conclusions about whether the observations observed for the PSI complex will hold for other potential PP1 triple complexes. However, it is interesting to note that the I-2 secondary structure elements that were most highly populated in the unbound form stay bound in the PSI complex, indicating that these pre-formed secondary structural elements might also influence the binding strength of the interaction.

PP1:MYPT1 – THE PSEUDO IDP

MYPT1 (myosin phosphatase targeting subunit 1) is responsible for both targeting the PP1 holoenzyme to subcellular compartments in the muscle and directing PP1 specificity toward myosin. While 3-dimensional structures of PP1 in complex with small molecular toxins, as well as an RVxF peptide were reported in 1995–1998 [39, 55, 56], the PP1:MYPT1 complex was the first reported protein regulatory complex [45]. How is the complex formed?

Using NMR spectroscopy methods and biophysical tools similar to those previously used to understand the unbound form of spinophilin, I-2 and other PP1 regulatory proteins, the structure of unbound MYPT1 in solution was investigated [57]. The structure can be readily split into two structurally distinct regions: the N-terminal 40 residues are intrinsically disordered, while the C-terminal residues form a folded ankyrin-repeat domain. However, auto-correlated backbone relaxation measurements showed that fast-timescale motions in MYPT1 are much more restricted than in spinophilin or I-2.

The N-terminal intrinsically disordered region of MYPT1 contains a transiently populated α -helix (~25%; Fig. 5). Interestingly, this α -helix has been identified to contain a PP1 consensus binding sequence (R-X-X-Q-[V/I/L]-[K/R]-X-[Y/W], where X is any residue; residues 10–17 in MYPT1) and which is referred to as the myosin phosphatase N-terminal element or MyPhoNE (Figs. 2, 5) [40]. Based on this data, ensemble structures of unbound MYPT1 were calculated in order to understand the interaction of MYPT1 with PP1. The MYPT1 RVxF site-PP1 interaction is strong and driven by extensive hydrophobic contacts. In contrast, the MyPhoNE motif interaction contains both hydrophobic and electrostatic interactions. In particular, the electrostatic interactions are often more long range in nature. Thus, it is possible that electrostatic interactions are initially formed by the MyPhoNE helix, whereas the strong hydrophobic RVxF motif interaction provides the solid lock. Moreover, the MyPhoNE motif is more distant from the folded ankyrin-repeat region, possibly leading to an enlarged capture radius of MYPT1 for PP1. Thus, these two interaction sites provide an effective, faster “fly-casting” model for reeling in PP1.

CONCLUSION

IDPs play vital roles in a multitude of essential regulatory processes [4, 10, 13, 26, 38, 58]. PP1 interacting proteins provide a unique system for investigating the similarities and differences in a family of IDPs (the PP1 binding proteins) that bind a single macromolecule (the ser/thr phosphatase PP1). Our studies during the last years have clearly shown that a conformational selection mechanism is preferentially used by PP1 regulators, such as spinophilin and I-2, for PP1 binding. So far in all analyzed PP1 binding proteins/IDPs we were able to identify preformed/pre-populated secondary structural elements that were also detected in the correlated crystal structures. This enabled us to identify the bound form and thus allowed for direct comparison between the bound and unbound states. Interestingly, the preformed secondary structural elements varied in their pre-populated percentages, from ~20–80%. Equally interesting, not all identified pre-populated secondary structural elements were necessary for PP1 binding, and often were identified in regions that stayed flexible even after PP1 interaction, e.g. for I-2. These sites likely mediate interactions with additional regulatory proteins. Taken together, conformational selection, which might go hand-in-hand with folding-upon-binding to PP1, plays a major role in the formation of PP1 holoenzymes.

Acknowledgments

The authors thank all current and previous members of the Page and Peti laboratory for their contribution to this ongoing research during the last 7 years. The authors thank Dr. Mathieu Bollen for stimulating input into the extensive biology of PP1 and Dr. Julie Forman-Kay for support with the ENSEMBLE method. The work was

supported by grant R01NS056128 from the National Institute of Neurological Disorders and Stroke to W.P., grant R01GM098482 from the National Institute of General Medicine and an American Cancer Society research scholar grant (RSG-08-067-01-LIB) to R.P.

ABBREVIATIONS

IDP	Intrinsically disordered protein
NMR	Nuclear magnetic resonance
SAXS	Small angle X-Ray scattering
PP1	Protein phosphatase 1
I-2	Inhibitor-2
PI	PP1:inhibitor-2 complex
PS	PP1:spinophilin complex
PSI	PP1:spinophilin:I-2 complex
MYPT1	Myosin Phosphatase Targeting Subunit 1
HSQC	Heteronuclear single quantum coherence
TROSY	Transverse relaxation optimized spectroscopy

References

1. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol.* 2005; 6(3):197–208. [PubMed: 15738986]
2. Dyson HJ, Wright PE. According to current textbooks, a well-defined three-dimensional structure is a prerequisite for the function of a protein. Is this correct? *IUBMB Life.* 2006; 58(2):107–109. [PubMed: 16608823]
3. Wright PE, Dyson HJ. Intrinsically unstructured proteins: reassessing the protein structure-function paradigm. *J Mol Biol.* 1999; 293(2):321–331. [PubMed: 10550212]
4. 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. Intrinsically disordered protein. *J Mol Graph Model.* 2001; 19(1):26–59. [PubMed: 11381529]
5. Sasaki K, Dockerill S, Adamiak DA, Tickle IJ, Blundell T. X-ray analysis of glucagon and its relationship to receptor binding. *Nature.* 1975; 257(5529):751–757. [PubMed: 171582]
6. Boesch C, Bundi A, Oppliger M, Wuthrich K. 1H nuclear-magnetic-resonance studies of the molecular conformation of monomeric glucagon in aqueous solution. *Eur J Biochem.* 1978; 91(1):209–214. [PubMed: 720338]
7. Wright PE, Dyson HJ. Linking folding and binding. *Curr Opin Struct Biol.* 2009; 19(1):31–38. [PubMed: 19157855]
8. Ward JJ, McGuffin LJ, Bryson K, Buxton BF, Jones DT. The DISOPRED server for the prediction of protein disorder. *Bioinformatics.* 2004; 20(13):2138–2139. [PubMed: 15044227]
9. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol.* 2004; 337(3):635–645. [PubMed: 15019783]
10. Eliezer D. Biophysical characterization of intrinsically disordered proteins. *Curr Opin Struct Biol.* 2009; 19(1):23–30. [PubMed: 19162471]
11. Shortle DR. Structural analysis of non-native states of proteins by NMR methods. *Curr Opin Struct Biol.* 1996; 6(1):24–30. [PubMed: 8696969]
12. Dyson HJ, Wright PE. Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance. *Adv Protein Chem.* 2002; 62:311–340. [PubMed: 12418108]

13. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z. Intrinsic disorder and protein function. *Biochemistry*. 2002; 41(21):6573–6582. [PubMed: 12022860]
14. Dyson HJ, Wright PE. Unfolded proteins and protein folding studied by NMR. *Chem Rev*. 2004; 104(8):3607–3622. [PubMed: 15303830]
15. Eliezer D. Characterizing residual structure in disordered protein states using nuclear magnetic resonance. *Methods Mol Biol*. 2007; 350:49–67. [PubMed: 16957317]
16. Huang JR, Grzesiek S. Ensemble calculations of unstructured proteins constrained by RDC and PRE data: a case study of urea-denatured ubiquitin. *J Am Chem Soc*. 2010; 132(2):694–705. [PubMed: 20000836]
17. 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(9):1169–1185. [PubMed: 19748338]
18. Jensen MR, Salmon L, Nodet G, Blackledge M. Defining conformational ensembles of intrinsically disordered and partially folded proteins directly from chemical shifts. *J Am Chem Soc*. 2010; 132(4):1270–1272. [PubMed: 20063887]
19. Marsh JA, Forman-Kay JD. Structure and disorder in an unfolded state under nondenaturing conditions from ensemble models consistent with a large number of experimental restraints. *J Mol Biol*. 2009; 391(2):359–374. [PubMed: 19501099]
20. Meier S, Grzesiek S, Blackledge M. Mapping the conformational landscape of urea-denatured ubiquitin using residual dipolar couplings. *J Am Chem Soc*. 2007; 129(31):9799–9807. [PubMed: 17636913]
21. Mittag T, Forman-Kay JD. Atomic-level characterization of disordered protein ensembles. *Curr Opin Struc Biol*. 2007; 17(1):3–14.
22. Vendruscolo M. Determination of conformationally heterogeneous states of proteins. *Curr Opin Struc Bio*. 2007; 17(1):15–20.
23. Wuthrich, K. *NMR of Proteins and Nucleic Acids*. John Wiley & Sons; 1986.
24. Muller-Spath S, Soranno A, Hirschfeld V, Hofmann H, Ruegger S, Reymond L, Nettels D, Schuler B. From the Cover: Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc Natl Acad Sci U S A*. 2010; 107(33):14609–14614. [PubMed: 20639465]
25. Dancheck B, Ragusa MJ, Allaire M, Nairn AC, Page R, Peti W. Molecular investigations of the structure and function of the protein phosphatase 1-spinophilin-inhibitor 2 heterotrimeric complex. *Biochemistry*. 2011; 50(7):1238–1246. [PubMed: 21218781]
26. 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(9):1094–1103. [PubMed: 20826336]
27. Csizmok V, Bokor M, Banki P, Klement E, Medzihradszky KF, Friedrich P, Tompa K, Tompa P. Primary contact sites in intrinsically unstructured proteins: the case of calpastatin and microtubule-associated protein 2. *Biochemistry*. 2005; 44(10):3955–3964. [PubMed: 15751971]
28. Tompa P. The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett*. 2005; 579(15):3346–3354. [PubMed: 15943980]
29. Hoang TX, Marsella L, Trovato A, Seno F, Banavar JR, Maritan A. Common attributes of native-state structures of proteins, disordered proteins, and amyloid. *Proc Natl Acad Sci U S A*. 2006; 103(18):6883–6888. [PubMed: 16624879]
30. Weber G. Ligand binding and internal equilibria in proteins. *Biochemistry*. 1972; 11(5):864–78. [PubMed: 5059892]
31. Koshland DE. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc Natl Acad Sci U S A*. 1958; 44(2):98–104. [PubMed: 16590179]
32. Sugase K, Dyson HJ, Wright PE. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature*. 2007; 447(7147):1021–1025. [PubMed: 17522630]
33. Turjanski AG, Gutkind JS, Best RB, Hummer G. Binding-induced folding of a natively unstructured transcription factor. *PLoS Comput Biol*. 2008; 4(4):e1000060. [PubMed: 18404207]
34. Chen HF. Molecular dynamics simulation of phosphorylated KID post-translational modification. *PLoS One*. 2009; 4(8):e6516. [PubMed: 19654879]

35. Boehr DD, Wright PE. Biochemistry. How do proteins interact? *Science*. 2008; 320(5882):1429–1430. [PubMed: 18556537]
36. Lange OF, Lakomek NA, Fares C, Schroder GF, Walter KF, Becker S, Meiler J, Grubmuller H, Griesinger C, de Groot BL. Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science*. 2008; 320(5882):1471–1475. [PubMed: 18556554]
37. Uversky VN. Multitude of binding modes attainable by intrinsically disordered proteins: a portrait gallery of disorder-based complexes. *Chem Soc Rev*. 2011; 40(3):1623–1634. [PubMed: 21049125]
38. Bollen M, Peti W, Ragusa MJ, Beullens M. The extended PP1 toolkit: designed to create specificity. *Trends Biochem Sci*. 2010; 35(8):450–458. [PubMed: 20399103]
39. Kelker MS, Page R, Peti W. Crystal structures of protein phosphatase-1 bound to nodularin-R and tautomycin: a novel scaffold for structure-based drug design of serine/threonine phosphatase inhibitors. *J Mol Biol*. 2009; 385(1):11–21. [PubMed: 18992256]
40. Hendrickx A, Beullens M, Ceulemans H, Den Abt T, Van Eynde A, Nicolaescu E, Lesage B, Bollen M. Docking motif-guided mapping of the interactome of protein phosphatase-1. *Chem Biol*. 2009; 16(4):365–371. [PubMed: 19389623]
41. Meiselbach H, Sticht H, Enz R. Structural analysis of the protein phosphatase 1 docking motif: molecular description of binding specificities identifies interacting proteins. *Chem Biol*. 2006; 13(1):49–59. [PubMed: 16426971]
42. Wakula P, Beullens M, Ceulemans H, Stalmans W, Bollen M. Degeneracy and function of the ubiquitous RVXF motif that mediates binding to protein phosphatase-1. *J Biol Chem*. 2003; 278(21):18817–18823. [PubMed: 12657641]
43. 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(39):28874–28883. [PubMed: 17636256]
44. Ragusa MJ, Dancheck B, Critton DA, Nairn AC, Page R, Peti W. Spinophilin directs protein phosphatase 1 specificity by blocking substrate binding sites. *Nat Struct Mol Biol*. 2010; 17(4):459–464. [PubMed: 20305656]
45. Terrak M, Kerff F, Langsetmo K, Tao T, Dominguez R. Structural basis of protein phosphatase 1 regulation. *Nature*. 2004; 429(6993):780–784. [PubMed: 15164081]
46. Choy WY, Forman-Kay JD. Calculation of ensembles of structures representing the unfolded state of an SH3 domain. *J Mol Biol*. 2001; 308(5):1011–1032. [PubMed: 11352588]
47. Lindorff-Larsen K, Best RB, DePristo MA, Dobson CM, Vendruscolo M. Simultaneous determination of protein structure and dynamics. *Nature*. 2005; 433(7022):128–132. [PubMed: 15650731]
48. Chan HS, Dill KA. Interchain loops in polymers: effects of excluded volume. *J Chem Phys*. 1989; 90:492–509.
49. Ragusa MJ, Allaire M, Nairn AC, Page R, Peti W. Flexibility in the PP1:spinophilin holoenzyme. *FEBS Lett*. 2011; 585(1):36–40. [PubMed: 21094159]
50. Dancheck B, Nairn AC, Peti W. Detailed structural characterization of unbound protein phosphatase 1 inhibitors. *Biochemistry*. 2008; 47(47):12346–12356. [PubMed: 18954090]
51. Jurgensen S, Shacter E, Huang CY, Chock PB, Yang SD, Vandenheede JR, Merlevede W. On the mechanism of activation of the ATP X Mg(II)-dependent phosphoprotein phosphatase by kinase FA. *J Biol Chem*. 1984; 259(9):5864–570. [PubMed: 6325452]
52. Connor JH, Weiser DC, Li S, Hallenbeck JM, Shenolikar S. Growth arrest and DNA damage-inducible protein GADD34 assembles a novel signaling complex containing protein phosphatase 1 and inhibitor 1. *Mol Cell Biol*. 2001; 21(20):6841–6850. [PubMed: 11564868]
53. Lesage B, Beullens M, Pedelini L, Garcia-Gimeno MA, Waelkens E, Sanz P, Bollen M. A complex of catalytically inactive protein phosphatase-1 sandwiched between Sds22 and inhibitor-3. *Biochemistry*. 2007; 46(31):8909–8919. [PubMed: 17630778]
54. Terry-Lorenzo RT, Elliot E, Weiser DC, Prickett TD, Brautigan DL, Shenolikar S. Neurabins recruit protein phosphatase-1 and inhibitor-2 to the actin cytoskeleton. *J Biol Chem*. 2002; 277(48):46535–46543. [PubMed: 12270929]

55. Egloff MP, Johnson DF, Moorhead G, Cohen PT, Cohen P, Barford D. Structural basis for the recognition of regulatory subunits by the catalytic subunit of protein phosphatase 1. *EMBO J.* 1997; 16(8):1876–1887. [PubMed: 9155014]
56. Goldberg J, Huang HB, Kwon YG, Greengard P, Nairn AC, Kuriyan J. Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1. *Nature.* 1995; 376(6543):745–753. [PubMed: 7651533]
57. Pinheiro AS, Marsh JA, Forman-Kay JD, Peti W. Structural signature of the MYPT1-PP1 interaction. *J Am Chem Soc.* 2011; 133(1):73–80. [PubMed: 21142030]
58. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol.* 2005; 6(3):197–208. [PubMed: 15738986]

Fig. 1. Investigating unbound IDPs

NMR parameters, including chemical shifts (translated into chemical shift index CSI or secondary structure preference SSP scores), residual dipolar couplings (RDCs) and paramagnetic relaxation enhancement measurements (PREs), provide an ensemble averaged snapshot of the unstructured state of IDPs (left and center). These parameters can be used to calculate ensemble structures of the IDP consistent with experimental data (right). The ensemble structures can be further analyzed by cluster analysis to determine structural motifs important for IDP function and provide insights into the mechanism that leads to the selection of a single folded conformation when IDPs bind to a folded scaffold.

Fig. 2. PP1 targeting protein binding surfaces and substrate binding grooves

The PP1 surface (grey) is highly grooved, allowing it to interact with multiple proteins. Illustrated here is the PP1 active site (pink), which contains two metals and the acidic (blue), the C-terminal (green), the hydrophobic (beige), the RVxF (magenta), the MyPhoNE (orange) and the SILK (cyan) binding grooves.

Fig. 3. The spinophilin PP1 binding domain exploits conformational selection to bind PP1
(A) Multiple NMR-based experiments were used to demonstrate that spinophilin adopts a highly disordered structure in solution. **(B)** Upper left triangle: fractional contact plot of the calculated ensemble structure of spinophilin (18% populated cluster) shows that the secondary structural elements observed in the PP1:spinophilin crystal structure (lower right triangle – contour plot – spinophilin:PP1 crystal structure PDBid: 3EGG) **(C)** are transiently populated in the free form (secondary structural elements are boxed in B and C); the spinophilin:PP1 crystal structure also includes the C-terminally adjacent spinophilin PDZ domain (spinophilin residues 498–583).

Fig. 4. The PP1:I-2 and PP1:spinophilin:I-2 complex

Because I-2 and spinophilin occupy overlapping sites on PP1 (i.e., the RVxF motif binding pocket), the interactions between these PP1 regulatory proteins and PP1 change upon formation of the PP1:spinophilin:I-2 triple complex. Namely, the RVxF motif of I-2 dissociates from the PP1 RVxF binding pocket, while the C-terminal helix of spinophilin dissociates from PP1. This leads to a triple complex (right panel) which is significantly more dynamic than either holoenzyme complex alone (left panels).

Fig. 5. MYPT1 is only “slightly disordered”

SSP scores of the unbound PP1 binding domain of MYPT1 indicate that the N-terminal residues form a transient α -helix (25% populated, left panel), which includes the MYPT1 residues that anchor to PP1 using the MyPhoNE binding site, while the ankyrin repeat α -helices are highly populated (80–85% populated, left panel). The primary sequence of the MyPhoNE and RVxF binding sites are annotated. Calculated ensemble structures provide further evidence that the N-terminal residues have a tendency to adopt an α -helical character. The structure of the PP1 binding domain of MYPT1 when bound to PP1 (left, PP1 is not included in this figure) confirms both the NMR and ensemble calculation predictions.