BMC Genomics

BioMed Central

Research

Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners

Christopher J Oldfield[1], Jingwei Meng[1], Jack Y Yang[1], Mary Qu Yang[1], Vladimir N Uversky[1,2] and A Keith Dunker*[1]

Address: [1]Center for Computational Biology and Bioinformatics, Indiana University Schools of Medicine and Informatics, 410 W. 10th Street, Indianapolis, IN 46202, USA and [2]Institute for Biological Instrumentation, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia

Email: Christopher J Oldfield - cjoldfie@iupui.edu; Jingwei Meng - mengj@iupui.edu; Jack Y Yang - purduejk@ecn.purdue.edu; Mary Qu Yang - purduexy@ecn.purdue.edu; Vladimir N Uversky - vuversky@iupui.edu; A Keith Dunker* - kedunker@iupui.edu

* Corresponding author

from The 2007 International Conference on Bioinformatics & Computational Biology (BIOCOMP'07)
Las Vegas, NV, USA. 25-28 June 2007

Published: 20 March 2008

BMC Genomics 2008, **9**(Suppl 1):S1 doi:10.1186/1471-2164-9-S1-S1

This article is available from: http://www.biomedcentral.com/1471-2164/9/S1/S1

Abstract

Background: Proteins are involved in many interactions with other proteins leading to networks that regulate and control a wide variety of physiological processes. Some of these proteins, called hub proteins or hubs, bind to many different protein partners. Protein intrinsic disorder, via diversity arising from structural plasticity or flexibility, provide a means for hubs to associate with many partners (Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN: Flexible Nets: The roles of intrinsic disorder in protein interaction networks. *FEBS J* 2005, 272:5129-5148).

Results: Here we present a detailed examination of two divergent examples: 1) p53, which uses different disordered regions to bind to different partners and which also has several individual disordered regions that each bind to multiple partners, and 2) 14-3-3, which is a structured protein that associates with many different intrinsically disordered partners. For both examples, three-dimensional structures of multiple complexes reveal that the flexibility and plasticity of intrinsically disordered protein regions as well as induced-fit changes in the structured regions are both important for binding diversity.

Conclusions: These data support the conjecture that hub proteins often utilize intrinsic disorder to bind to multiple partners and provide detailed information about induced fit in structured regions.

Background

Protein-protein interaction (PPI) networks integrate various biological signals including those used for energy generation, cell division and growth to give a few notable examples. The architectures of the PPI networks indicate that they are nearly scale free [1-8]. That is, a log-log plot of the number of nodes versus the number of links (or interactions) at each node gives a straight line with a negative slope. The negative slope means that these sets of interactions contain a few proteins (hubs) with many links and many proteins (non-hubs) with only a few links. The term 'hub protein' is relative to the other proteins in a given PPI network, with no agreed upon number of links separating hubs and non-hubs.

Several networks such as the internet, cellular phone systems, social interactions, author citations, and so on, exhibit scale-free architecture. With regard to PPIs, scale-free network architecture is suggested to provide several biological advantages. For example, given the small fraction of hub proteins, random deleterious mutations will more likely occur in non-hub proteins. The elimination of the functions of such non-hub proteins typically have small effects and so, generally, are not serious. In contrast, a deleterious mutation of a hub protein is more likely to be lethal [4-9]. Another advantage is that signals can traverse these networks in a small number of steps, so signal transduction efficiency is improved compared to that expected for random networks [7].

Understanding PPI network evolution across different species is an important problem [10-13]. From this body of work, hub proteins appear to evolve more slowly than non-hub proteins, an observation that is consistent with Fisher's classic proposal that pleiotropy constrains evolution [14,15]. Some proteins have multiple, simultaneous interactions ("party hubs") [16] while others have multiple, sequential interactions ("date hubs") [16]. Date hubs appear to connect biological modules to each other [17] while party hubs evidently form scaffolds that assemble functional modules [16].

The idea that PPI networks use scale-free network topology is receiving considerable attention, but some caution is in order. Currently constructed networks are noisy, with both false positive and false negative interactions [8,18-20]. Also, network coverage to date [21,21-24] is not sufficient to prove scale-free architecture [25]. Whether PPI networks are truly scale-free or only approximately so, it nevertheless appears to be true that a relatively small number of proteins interact with many partners, either as date hubs or party hubs, while many proteins interact with just a few partners.

The ability of a protein to bind to multiple partners was suggested to involve new principles [26]. Indeed, neither the lock-and-key [27] nor the original induced-fit [28] readily explains how one protein can bind to multiple partners. Note that the original induced fit mechanism was defined as changes in a structured binding site upon binding to the partner [28], changes that are analogous to a glove altering its shape to fit a hand. On the other hand, both theoretical and experimental studies over many years suggested that natively unstructured or intrinsically disordered proteins form multi-structure ensembles that present different structures for binding to different partners [29-35]. Based on these prior studies, we proposed that molecular recognition via disorder-to-order transitions provides a mechanism for hub proteins to specifically recognize multiple partners [36]. We pointed out earlier that intrinsic disorder could enable one protein to associate with multiple partners (one-to-many signaling) and could also enable multiple partners to associate with one protein (many-to-one signaling) [35].

Recent bioinformatics studies support the importance of protein disorder for hubs [37-41]. While disorder appears to be more clearly associated with date hubs [39,41] than with party hubs, some protein complexes clearly use long regions of disorder as a scaffold for assembling an interacting group of proteins [42,42-50]. Thus, the importance of disorder for party hubs needs to be examined further. Additional evidence for the importance of disorder for highly connected hub proteins comes from a structure-based study of the yeast protein interaction network [51]. The authors considered only interactions that could be mediated by domains with known structures and found that the degree distribution of the resulting network contained no proteins with more than 14 interactions, which is more than an order of magnitude less than is observed in one unfiltered, high confidence dataset (Jake Chen, personal communication). This result indicates that a structure-based view of hub proteins is insufficient to explain the multitude of partners that interact with hub proteins.

To improve understanding of the use of disorder for binding diversity, we studied two prototypical examples: p53 and 14-3-3. Both are hubs that are clearly involved in crucial biological functions. For example, p53 is a key player in a large signaling network involving the expression of genes carrying out such processes as cell cycle progression, apoptosis induction, DNA repair, response to cellular stress, etc. [52]. Loss of p53 function, either directly through mutation or indirectly through several other mechanisms, is often accompanied by cancerous transformation [53]. Cancers with mutations in p53 occur in colon, lung, esophagus, breast, liver, brain, reticuloendothelial tissues and hemopoietic tissues [53]. The p53

protein induces or inhibits over 150 genes, including *p21*, *GADD45*, *MDM2*, *IGFBP3*, and *BAX*[54].

The four regions or (not necessarily structured) domains in p53 are the N-terminal transcription activation domain, the central DNA binding domain, the C-terminal tetramerization domain, and the C-terminal regulatory domain. The last two could be considered to be a single C-terminal domain with two subregions. The transactivation region interacts with TFIID, TFIIH, Mdm2, RPA, CBP/p300 and CSN5/Jab1 among many other proteins [52]. The C-terminal domain interacts with GSK3β, PARP-1, TAF1, TRRAP, hGcn5, TAF, 14-3-3, S100B(ββ) and many other proteins [55].

As for 14-3-3 proteins, they contribute to a wide range of crucial regulatory processes including signal transduction, apoptosis, cell cycle progression, DNA replication, and cell malignant transformation [56]. These activities involve 14-3-3 interactions with various proteins in a phosphorylation-dependent manner. More than 200 proteins have been shown to interact with members of 14-3-3 family [57-59], with these 14-3-3-interacting proteins amounting to approximately 0.6% of the human proteome [59]. One proposed functional model is that 14-3-3 binds to the specific target as a molecular anvil causing conformational changes in the partner. In their turn, these changes can affect enzymatic (biological) activity of a target protein, or mask or reveal specific motifs that regulate its localization, activity, phosphorylation state, and/or stability [60].

The 14-3-3 protein has at least nine sequence isomers, called α, β, γ, δ, ε, η, σ, τ, and ζ [61]. All isomers are structured dimers with grooves that bind to more than 200 different partners, and the different partners have different sequences for their binding regions. Screening experiments have identified individual peptides that bind to all the different isomers, suggesting that the binding grooves in the different isomers have some common features [62]. A recent bioinformatics study suggests that the partners of 14-3-3 utilize intrinsic disorder for binding [63].

The interactions of p53 and 14-3-3 with their partners as reported previously [61,64-79] are examined herein but from an order-disorder point of view. In the case of p53, different regions in the disordered tails enable this protein to bind to multiple partners at the same time. In addition, one single region of disorder adopts clearly different secondary structures and uses the same amino acids to different extents in different binding interactions. For this case the plasticity of the disordered region clearly enables the binding to multiple partners. In the case of 14-3-3, the different partners have distinct sequences. Their interactions with 14-3-3 show characteristics, such as hydrogen bonds between side chains of 14-3-3 and the backbone of the partners and such as hydrogen bonds between the backbone of the partners and water, indicating that the two partners were very likely unfolded in water just prior to association with 14-3-3. The distinct sequences of the partners do not adopt identical backbone structures, and the various side chain interactions between 14-3-3 and the two different partners involve induced-fit adjustments of the 14-3-3 structure. Overall, these studies show how the plasticity of disordered proteins is used to enable the binding diversity of hub proteins, both for a single disordered region binding to multiple partners and for multiple disordered regions binding to the same partner. An earlier, less complete version of this work was reported at the Biocomp'07 meeting [80].

Results

Intrinsic disorder and the molecular interactions of p53

The p53 molecule interacts with many other proteins in order to carry out its signal transduction function. A number of these are downstream targets, such as transcription factors, and others are activators or inhibitors of p53's transactivation function. Many of these interactions have been mapped to regions of the p53 sequence (Figure 1, gray boxes): the N-terminal domain (i.e., the transactivation domain), the C-terminal domain (i.e., the regulatory domain), and the DNA binding domain (DBD). These domains have also been characterized in terms of their structure or lack thereof (Figure 1, red (disordered) and blue (structured) segments), where the DNA binding domain is intrinsically structured and the terminal domains are intrinsically disordered [81,82]. While the tetramerization domain is structured, the structure is acquired upon the formation of the complex. Additionally, multiple different posttranslational modifications have been identified in p53 (Figure 1, vertical ticks). These modifications are relevant here because they are a common method for altering protein interactions.

Comparing the regions of order and disorder reveals a strong bias towards the localization of the interactions within the intrinsically disordered regions. Overall, 60/84 = 71% of the interactions are mediated by intrinsically disordered regions in p53. A bias toward intrinsically disordered regions is even more pronounced in the sites of posttranslational modifications, with 86%, 90%, and 100% of observed acetylation, phosphorylation, and protein conjugation sites, respectfully, found in the disordered regions. This is consistent with previous observation of a strong bias for post translational modifications toward intrinsically disordered regions [83]. This concentration of functional elements within intrinsically disordered regions compares to just 29% of the residues being disordered [36]. Clearly, p53 exhibits a highly

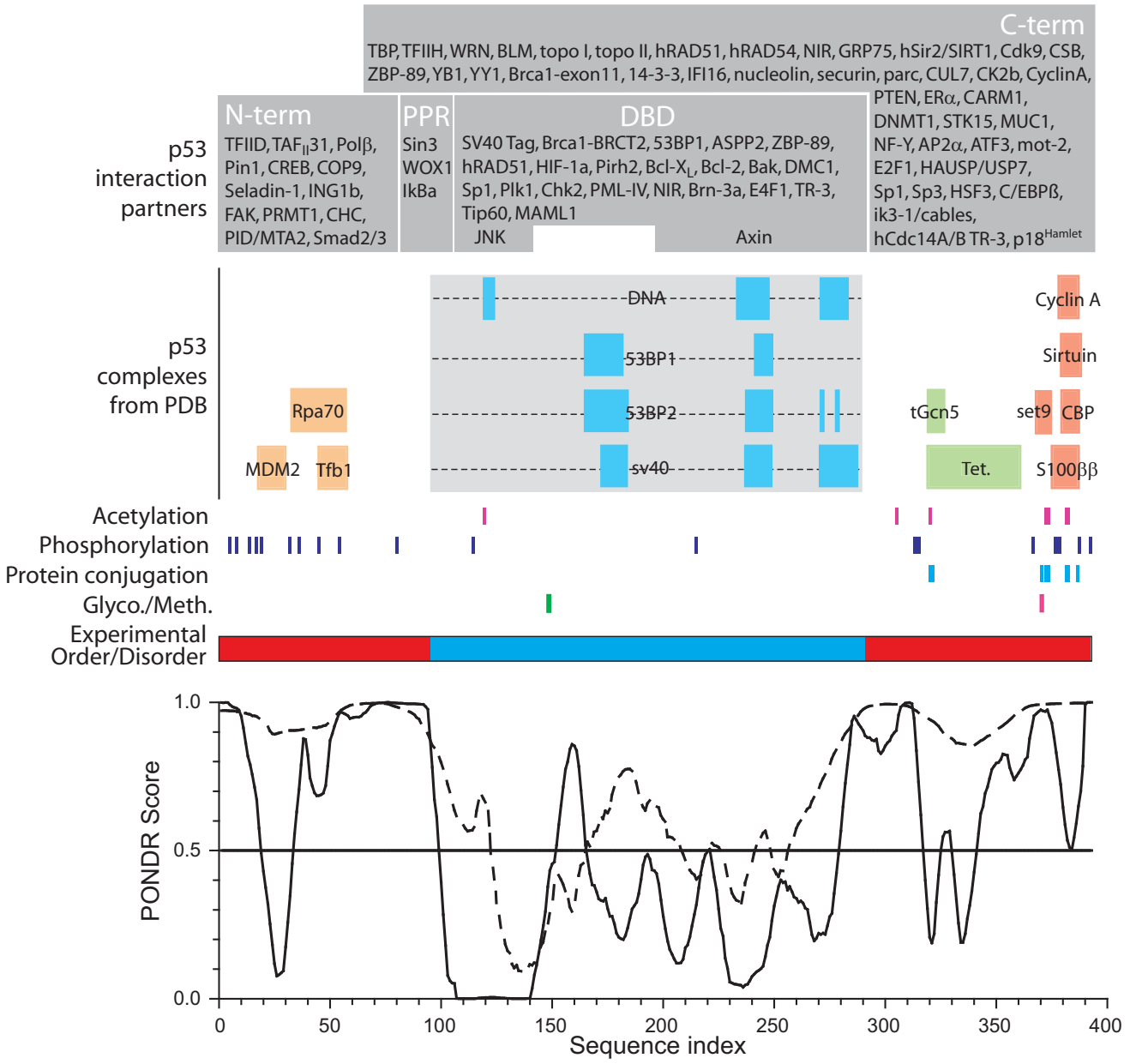

Figure 1

Summary of p53 interactions and structure. Dark gray boxes indicate the approximate binding regions of p53's known binding partners. The regions of p53 represented in structure complexes in PDB are represented by horizontal bars, labeled with the name of the binding partner. For the DBD, the extent of the globular domain is indicated by the light grey box, where the internal horizontal bars indicate regions involved in binding to a particular partner. Post translational modifications sites are represented by vertical ticks. Experimentally characterized regions of disorder (red) and order (blue) are indicated by the horizontal bar. Finally, predictions of disorder (scores > 0.5) and order (scores < 0.5) are shown for two PONDR predictors: VLXT (solid line) and VSL2P (dashed line). All, features are presented to scale, as indicated by the horizontal axis. The p53 interaction partners and post translational modification sites have been adapted from Anderson & Appella [55].

biased use of disordered regions for mediating and modulating interactions with other proteins.

In addition to experimentally characterized disorder, predictions of intrinsic disorder for p53 using both PONDR VL-XT [84] and VSL2 predictions [85] were carried out

(Figure 1, graph). The latter is one of the highest accuracy prediction algorithms available [86], whereas the former has been observed to be especially useful in identifying binding regions within longer regions of disorder [87-89] and to be much better at identifying such sites as compared to a number of different disorder predictors [90]. Both predictors give good agreement with the experimental determination of intrinsic disorder [83,91-109], and in the case of p53 both of their predictions agree well with experimental characterization.

Analysis of associations involving p53 using 3D structures

The structures of 14 complexes between various regions of p53 and unique binding partners have been determined (Figure 1, horizontal bars). For 10 of these partners, the interactions are mediated by regions experimentally characterized as intrinsically disordered, where PONDR VL-XT detects the majority of these binding regions as short predictions of order within a longer prediction of disorder. These structures are complexes between p53 and endogenous partners: cyclin A [64], sirtuin [65], CBP [66], S100ββ [67], set9 [68], tGcn5 [69], Rpa70 [70], MDM2 [71], Tfb1 [72], and itself [73]. The remaining 4 interactions are mediated by the structured DBD, namely between p53 and 3 endogenous partners – DNA [74], 53BP1 [75], and 53BP2 [76] – and one exogenous partner – the large-T antigen (LTag) from simian virus 40 [79].

Protein complexes can be formed from the association of structured proteins, by the folding of one disordered protein onto the surface of a structured partner, or by the coupled folding and binding of intrinsically disordered proteins [110-117]. Nussinov and collaborators [117] showed that a plot of normalized monomer area (NMA) versus normalized interface area (NIA) nicely separates complexes formed from structured proteins as compared to complexes formed from unfolded proteins by coupled binding and folding. That is, associations of structured proteins exhibit small NMAs and NIAs and so lie near the origin of the NMA-NIA plot. Conversely, complexes formed by coupled binding and folding have much larger NMAs and NIAs, and so are spread out and lie far from the origin of the NMA versus NIA plot. Indeed, a linear boundary separates the two groups [117]. IT should be emphasized that the NMA-NIA plot approach is a global measure of a proteins order-disorder monomeric state, and has not been characterized on local order-disorder transitions (e.g. disordered binding loops in an otherwise well ordered protein).

As described in more detail in the implementation, by developing two separate NMA-NIA plots, one for each partner of a complex (Figure 2A), and then by determining the distance to the linear boundary in each plot, a double NMA-NIA plot (Figure 2B) can be produced. Interacting pairs can be divided into the 3 groups given above, namely: (1) both partners are structured (region (i)

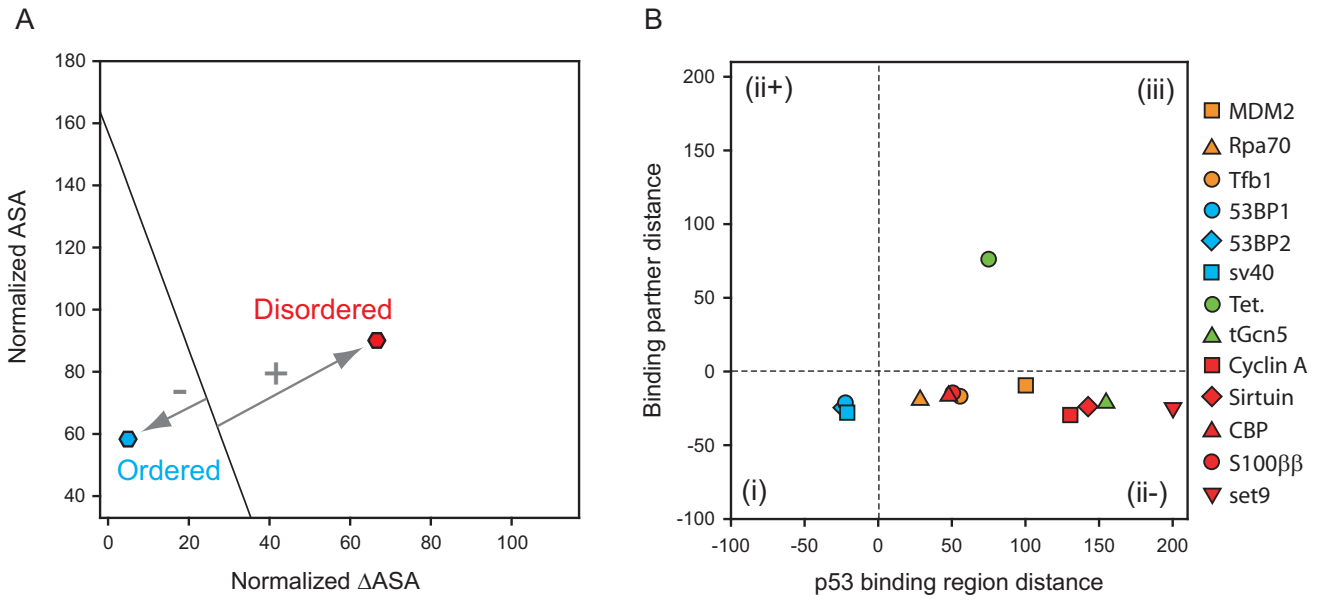

Figure 2

Double NMA-NIA plot for p53 complexes. (A) The definition of boundary distance used in the double NMA-NIA plot, where ordered structures have a negative boundary distance and disordered structures have a positive boundary distance. (B) The double NMA-NIA plot for the p53 structures shown in Figure 1, with the exception of DNA-bound p53.

of 2B), i.e. both distances are negative; (2) one partner is structured and the second partner is disordered, i.e. the ordered partner has a negative distance and the disordered partner has a positive distance (regions (ii+ and ii-) of 2B); and (3) both partners are intrinsically disordered, i.e. both distances are positive (region (iii) of 2B).

A double NMA-NIA plot was calculated for 13 of the p53 complex structures (Figure 2B). The p53-DNA complex was excluded since the NMA-NIA analysis is not relevant for nucleic acids. In the general case, the distinction between the distances of the two partners is arbitrary, so that the double NMA-NIA plot is symmetric about the diagonal. However, here we restrict the p53 distance to one axis, so that group (2) is split into two sub groups (regions (ii+ and ii-)): the p53 segment is disordered and the partner is ordered (region (ii-)) and the p53 region is ordered and the partner is disordered (region (ii+)). One interaction, the formation of the p53 tetramer, is in the third group (region (iii)) and so therefore likely involves an association between two disordered partners. This is consistent with experimental data [81]. At the opposite side of the spectrum, the three protein-protein complexes involving the p53 DBD domain are in group 1 (region (iii)), indicating that all three are ordered prior to binding, which is consistent with the solution of structures for identical or homologous monomeric domains (e.g. p53 DBD [118], 53BP1 BRCT domain [119], 53BP2 SH3 domain [120], and LTag [79]). The other nine p53 complexes found so far in the PDB are all in the group 2 quadrant (that is, in region (ii-), and so all likely involve a disordered region of p53 associating with a structured partner. These results are likewise consistent with experimental data. That is, these p53 regions are disordered in the unbound state [81,82], and the isolated partners appear to be structured: MDM2 [121], Rpa70 [70], Tfb1 [72], tGCN5 [122], Cyclin A/CDK2 [123], sirtuin [124], CBP [125], S100ββ [126] and set9 [127]).

In summary, these data point out the importance of disorder-to-order transitions for many of the structurally characterized interactions involving the p53 hub protein. While many previous studies discuss these same interactions, to our knowledge the importance of disorder has not been emphasized in those previous studies.

Analysis of multiple specificities in the p53 C-terminus

So far, complexes involving one region of the p53 sequence bound to four different partners have been determined and deposited in the PDB. This region is from residue 374 to 388 in the p53 sequence bound to one of the following: cyclin A [64], sirtuin [65], CBP [66], or S100ββ [67]. The regions that mediate these interactions and their respective secondary structures were mapped precisely to the p53 sequence (Figure 3A). Although slightly different residues of the p53 sequence are used in

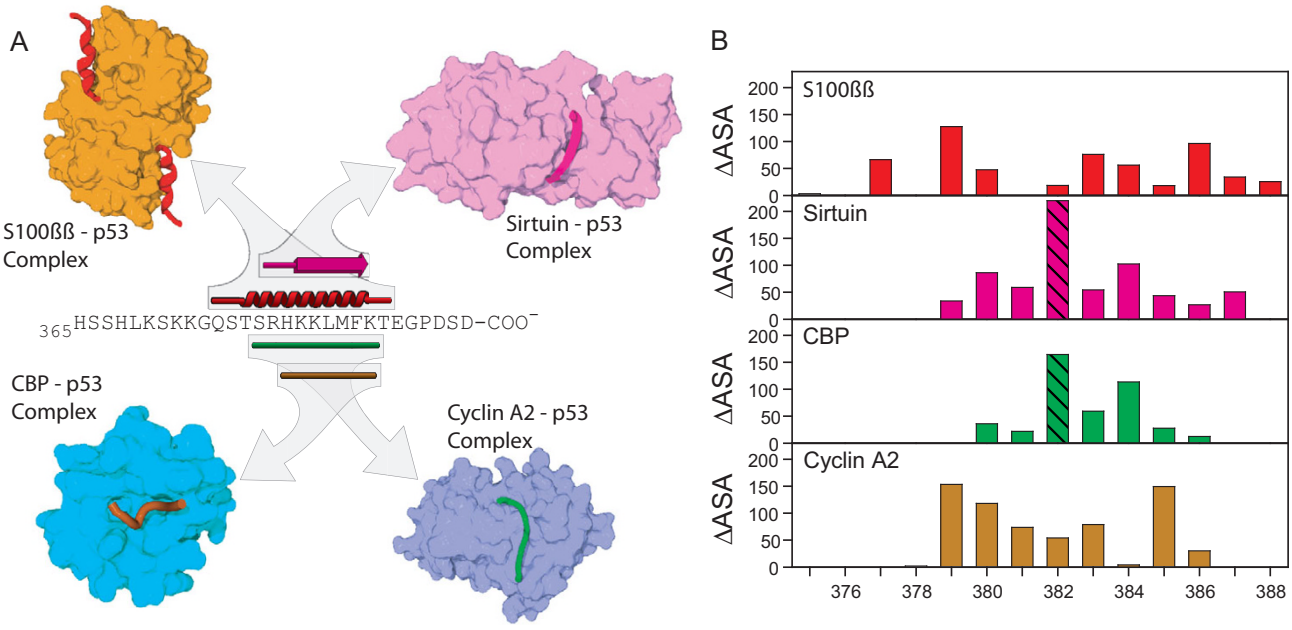

Figure 3

Sequence and structure comparison for the four overlapping complexes in the C-terminus of p53. (A) Primary, secondary, and quaternary structure of p53 complexes. (B) The ΔASA for rigid association between the components of complexes for each residue in the relevant sequence region of p53. The two hatched bars indicate acetylated lysine residues.

each interaction, there is a very high degree of overlap, with a span of 7 core residues being the same (Figure 3A). Interestingly, the four complexes display all three major secondary structure types. The core span becomes a helix when binding to S100ββ, a sheet when binding to sirtuin, and a coil with two distinct backbone trajectories when binding to CBP and cyclin A2 (Figure 3A).

Because the secondary structures are distinct, it seems likely that p53 utilizes different residues for the interactions with these four different partners. To examine this, the buried surface area for each residue in each interaction was quantified by calculating the ΔASA (Figure 3B). Different amino acid interaction profiles are seen for each of the interactions, showing that the same residues are used to different extents in the four interfaces. The particularly large ΔASA peaks for K382 in complexes with CBP and sirtiun (indicated by the hatched bar) are due to extra buried areas arising from the acetylation of this residue. This highlights the importance of posttranslational modification for altering PPI networks.

Analysis of multiple specificities of the p53 DBD

The p53 molecule contains another set of overlapping interactions that contrasts with those at the C-terminus. These interactions are mediated by the DNA binding domain and include interactions with DNA [74], the BRCT domain of 53BP1 [75], the SH3 domain of 53BP2 [76], and the large T-antigen (LTag) of Simian Virus 40 [79]. Here we compare these four interactions using the methods described in Figure 4.

The structures of the p53-DNA, the p53-53BP1, the p53-53BP2, and the p53-LTag complexes are shown (Figure 4B). While all of the ligands are different, they all bind to basically the same region of p53.

Comparison of the interface profiles of the four complexes (Figure 4A) shows a large difference in the pattern of interface residues used by p53. For instance, there are several residues at the N-terminal end of the DBD which are only found in interaction with DNA. Similarly, interface residues near the C-terminal end participate in binding to different extents in three interactions, but not at all in the p53-53BP1 interaction. The differing usage of residues in each interaction is the most prevalent feature of this data. However, there are also several residues contributing an exceptionally large amount of surface area in each complex (*e.g.*, M243 and R248).

While the focus of this paper is on the roles of disorder in the interactions involving two different hub proteins, the DNA binding domain of p53 presents the opportunity to study structural changes involving one structured region binding to several different structured partners. For this purpose, we compiled a 4-panel set of plots for characterizing the induced fit as one protein binds to different partners (Figure 5). These panels show the average interface area (Figure 5A), the standard deviation of the interface area (Figure 5B), the differences in side chain conformation (Figure 5C), and differences in backbone conformation (Figure 5D). Furthermore, regions that are highly exposed to solvent are also indicated (Figure 5, blue shading), so that structural differences due to interactions can be distinguished from those due from intrinsic flexibility – disordered loops – or crystallization artifacts.

This induced-fit profiles exhibit a number of interesting features (Figure 5). The most striking of these is the region from residue 240 to residue 250. This region shows a large and variable interaction interface, which is associated with large side chain and backbone conformational differences. This is true also of a smaller region around residue 120. Other interaction regions show only side chain conformational differences associated with variable interface areas. Other conformational differences observed are limited regions of high solvent exposure, which suggests that these changes are due the details of the crystallization conditions more than interaction with a particular binding partner.

Together, these results suggest that multiple partners of p53 are accommodated by reusing similar binding interfaces. This is facilitated by small scale or large scale structural differences, which range from differences in side chain conformation to backbone rearrangements. It should be noted that this differs from our finding in a more limited analysis on only the p53-53BP1 and -53BP complexes [80].

Analysis of the multiple specificities of 14-3-3

Five different 3-D structures of the 14-3-3ζ protein bound to distinct partners were found in PDB. These partners include a peptide from the tail of histone H3 [128], serotonin N-acetyltransferase (AANAT) [77], a phage display-derived peptide (R18) [78], and motif 1 and 2 peptides (m1 and m2, respectively) [61]. For AANAT, only the region within the canonical 14-3-3 binding site is included in our analysis with the globular region being deleted. Two additional structures were not included because they were either unsuitable for structural analysis or were highly redundant with another structure. All peptides are phosphorylated in their respective structures except R18, which contains a glutamate in place of the phosphoserine.

The five bound peptides sequences were aligned structurally as described in the methods. Likewise, the 14-3-3 domain structures were independently aligned, without considering the bound peptides. Next the 14-3-3 align-

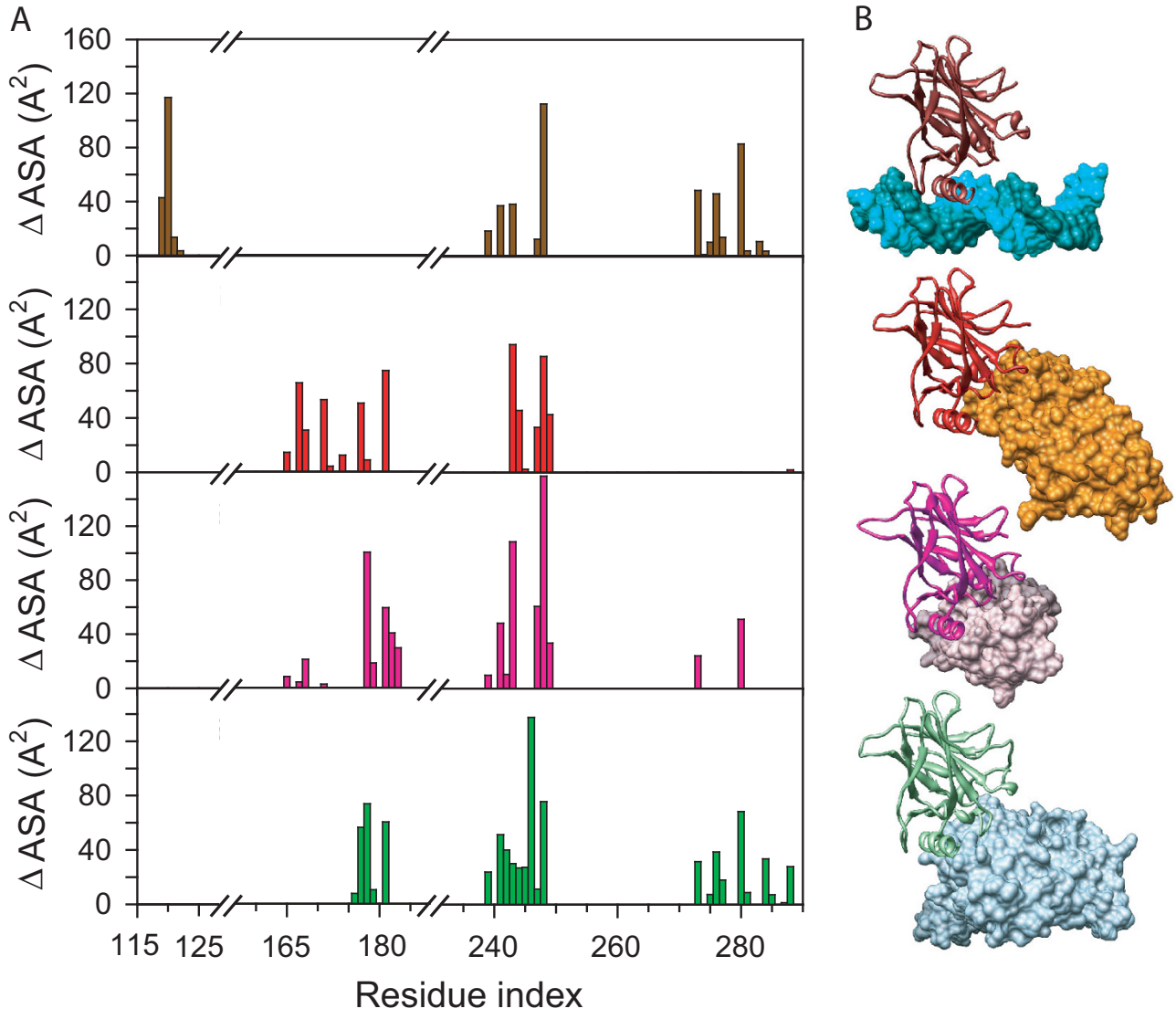

Figure 4
p53 DBD interaction with different binding partners. The interaction profiles (A) and rendered structures (B) for the four unique complexes of the p53 DBD. Rendered structures depict p53 as a ribbon and each interaction partner as a molecular surface. The interaction profile-structure pairs are (from top to bottom): p53-DNA, p53-53BP1, p53-53BP2, and p53-sv40.

ment was anchored manually by the observed correspondence of the bound peptide C_α atoms at the 0 and -1 positions and by extending the alignment without gaps from the anchor positions, thereby giving the final structural alignment (Figure 6A). In terms of sequence, the R18 sequence has no identical positions to any other peptide. The number of identities between the other peptides range from 1 to 4.

The high overlap in the backbone trajectories of the 5 peptides from position -3 to 1 but large divergences at either end of the structural alignment (Figure 6A). This divergence at the ends is apparent qualitatively in the superimposed structures of the five peptides (Figure 6B). Structural divergence and sequence variability are loosely correlated, where positions with 3 identical residues have a lower divergence than those with no identical residues. This suggests that 14-3-3 may use different binding pocket residues to interact with different peptide residues. The R18 sequence, which is divergent from the others, makes a large contribution to the estimated RMSF values (indicated by the cross-hatched bars, Figure 6A).

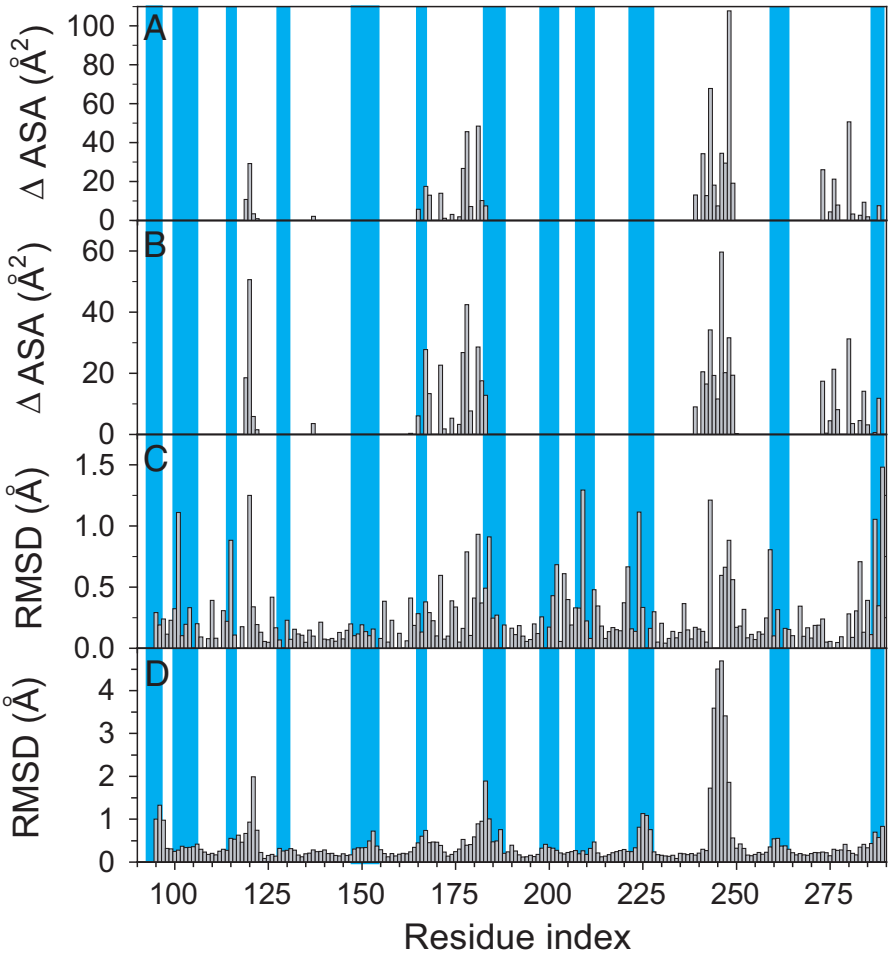

Figure 5

Comparison of residue interactions with structural differences for bound p53 DBD. The average (A) and standard deviations (B) were calculated over the four interaction profiles of the p53 DBD shown in Figure 4. These are shown aligned with the side chain RMSF (C) and the backbone RMSF (D) calculated from the four structures of bound p53 DBD. Regions of residues that are highly exposed to solvent in all complex structures are indicated by the blue-shaded regions.

The factors contributing to the ability of 14-3-3 to bind to distinct peptides were estimated by a detailed structural analysis. The peptide binding residues of 14-3-3 are located primarily in a central cleft, made up of four helices (Figure 7A), which has been noted previously by several researchers. The standard deviation of ΔASA for the peptide binding residues (Figure 7B) show that the residues with the most binding variability are located at either end of the central cleft, which is consistent with the variation of peptide backbone trajectories in these regions. Backbone variability in bound 14-3-3 structures (Figure 7C) is restricted to the ends of most of the binding cleft helices. These observations suggest that large a conformational change in 14-3-3 is not necessary for multiple specificities, although some small adjustments at the ends of binding helices may be necessary.

To assess the role of side chain conformational changes in peptide binding, the RMSF of side chain atoms was calculated (Figure 7D). The side chain RMSF and standard deviation of ΔASAs give similar indications for many binding site residues, where residues used inconsistently across multiple complexes are the most likely to undergo conformational rearrangement. These are the same residues that are located at the broadest parts of the binding site. However, a few residues deep in the binding grove show both consistent participation in the binding interface and variable side chain conformation. These observations suggests that the primary, high level mechanisms of 14-3-3 multiple specificity are a broad binding site that allows multiple trajectories (and therefore interaction with different residues) and side chain rearrangement to accommodate different peptide sequences.

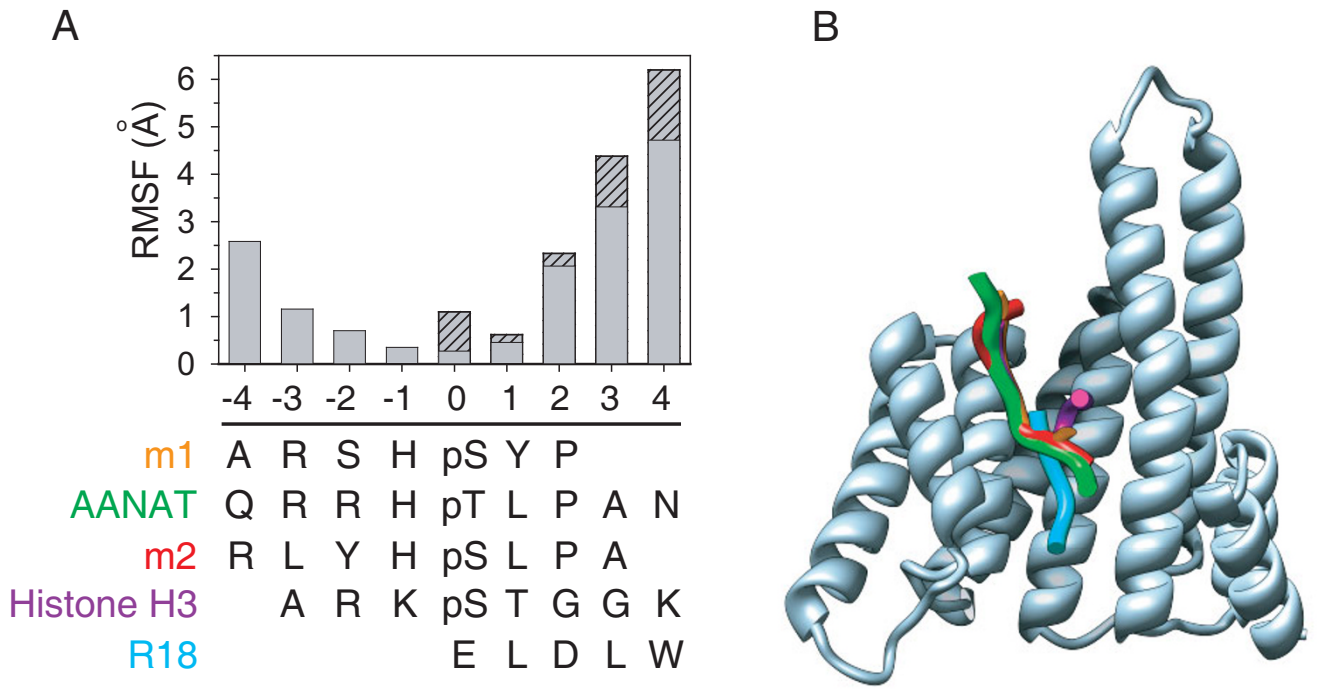

Figure 6

Sequence and structure for five peptides bound to 14-3-3ζ. (A) Sequence alignment of the bound peptides and the RMSF of their conformations. Solid grey bars give the RMSF for four peptides – excluding R18 – and the hatched bars give the RMSF for all five peptides. (B) Aligned ribbon representations of the structures of the five peptides, which were aligned through multiple alignment of their respectively bound 14-3-3 domains, show along with a representative ribbon representation of a 14-3-3 domain.

To further analyze the conformational changes in 14-3-3 upon binding to its multiple partners, we show the 4-panel induced-fit profile described above (Figure 8). Contrary to the results seen for the p53 DBD, 14-3-3 is much more static in its multiple interactions. All regions displaying large conformational differences across bound complexes are also highly exposed to solvent and play no direct role in mediating binding to any peptide. The plots do show several small scale structural differences – side chain rearrangements – associated with variable participation in peptide binding, particularly in the regions 40-60 and 215-230.

14-3-3 binding to two different partners

To gain further insight into 14-3-3 binding to different partners, we compared a pair of 14-3-3 binding peptides in detail. These two peptides, m1 and m2, were derived from two motifs, identified through the screening of peptide libraries for sequences that bound to all 14-3-3 isoforms [62]. These two peptide structures have been compared previously [61], but here we reanalyze these structural data from the order-disorder point of view.

As noted previously [61], the backbone traces of the two peptides are noticeably different even though the m1 and m2 peptides bind to essentially the same region of 14-3-3ζ (Figure 9A and B, respectively). Examining the side chain interactions of these peptides with specific 14-3-3 residues (Figure 9C and D) shows that there is difference in the location and identity of the residues involved, which is consistent with the aggregate findings (Figures 7 and 8). Similarly, distinctive hydrogen bonding patterns are exhibited between the two peptides and 14-3-3ζ and between the two peptides and bound water (Figure 9C and D). Since a cardinal feature of a structured protein is internal satisfaction of hydrogen bond donors and acceptors, these data are both consistent with the peptides being from unstructured regions of protein before binding.

The above data on the complexes suggest that 14-3-3ζ has distinct conformations when bound to the two different peptides. Overlaying the backbone structures of the four binding helices from both complexes – based on a pairwise alignment of the complete domains – shows only minor variability in conformation, with the most occurring at the helix spanning residues 216 to 228 (Figure 9E).

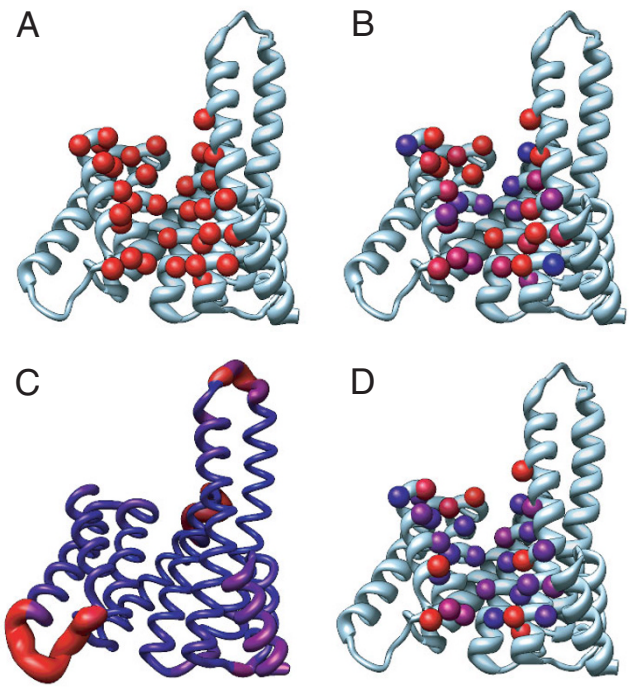

Figure 7
Peptide binding residues of 14-3-3ζ. (A) The C_β atoms of all residues involved in binding in any of the five peptide bound structures are shown (red) along with the rest of the backbone (light blue ribbon). (B) The standard deviation in the area bound on complex formation is displayed by coloring the C_β atoms of peptide binding residues on a gradient, from a standard deviation of 0Å^2 (blue) to 10Å^2 and greater (red). (C) The backbone RMSF of the 14-3-3 domain calculated over C_α atoms displayed as a color and radius gradient, from an RMSF of 0Å (blue, 0.25Å) to an RMSF of 2.0Å and greater (red, 2.0Å). (D) The side chain RMSF is displayed by coloring the C_β atoms of peptide binding residues on a gradient, from a RMSF of 0Å (blue) to an RMSF of 0.50Å and greater (red). All parameters were calculated using all five of the peptide-14-3-3 complexes.

Finally, comparison of side chain conformation in the two complexes shows significant differences in several of 14-3-3ζ side chains (Figure 7F, residues outlined in red show significant movement) and several other minor differences. Overall, these data suggest that a difference in the conformations of some side chains with rather less difference in backbone conformations is sufficient to accommodate the binding of two different phosphopeptides by the 14-3-3ζ molecule.

Discussion

Use of disordered regions for binding

The large majority of the binding sites on the p53 sequence map to the disordered regions of this protein (Figure 1), indicating that intrinsic disorder commonly provides the binding sites for the various partners that associate with p53. Recent bioinformatics investigations suggest that the majority use of disorder for binding to multiple partners is quite likely to be a general result [37-41].

The p53 binding sites are often indicated on the order-disorder predictions as dips, in other words as short segments with structure tendency flanked by regions of disorder tendency on both sides. Starting from this observation, we previously developed a predictor of such regions, which we called molecular recognition features, or MoRFs, because such regions "morph" from disorder to order upon binding [87,88]. Others have used the PONDR VL-XT order/disorder plots or MoRF predictors to identify potential binding sites that were subsequently verified by laboratory experiments [89,129]. Indeed, for some of these predicted examples, the regions did indeed form helix upon binding to their partners [130,131]. By greatly enlarging the training set, we recently improved the MoRF predictor. Interestingly, when tested against several order-disorder predictors including ones from other laboratories, PONDR VL-XT, gave the clearest indication of binding sites within disordered regions [90].

Others developed a sequence-based approach to identify short, conserved recognition sites, called eukaryotic linear motifs (ELMs) [132,132-134]. While MoRFs are identified by general order/disorder tendencies and while ELMs are identified by motif discovery from sequence analysis, the resulting binding sites identified by both methods share several features [135]. The use of different residues in the same disordered fragment for one-to-many signaling leads to a potential problem with the ELM model. That is, the concept behind ELMs is that each ELM uses a common set of amino acids for binding to different partners. These common amino acids therefore show up as an over-represented pattern leading to a "linear motif". What if a region used to bind to multiple partners uses different secondary structures and different amino acids? In such a case, the residues in the "linear motif" would not necessarily be over-represented. It will be interesting over time to determine whether ELMs having stronger signals use a reduced set of structures for their interactions.

While the observed binding sites in the disordered regions of p53 have a localized tendency for ordered structure, not all disorder-associated binding sites exhibit such features. We have found many binding sites that are associated with high disorder prediction values across the entire spans of the binding sites, one example of which was recently published [136]. Many of these dipless MoRFs form irregular structures upon binding with their partners, and often such binding regions are rich in proline. Our recent study of the complexes that form when various dis-

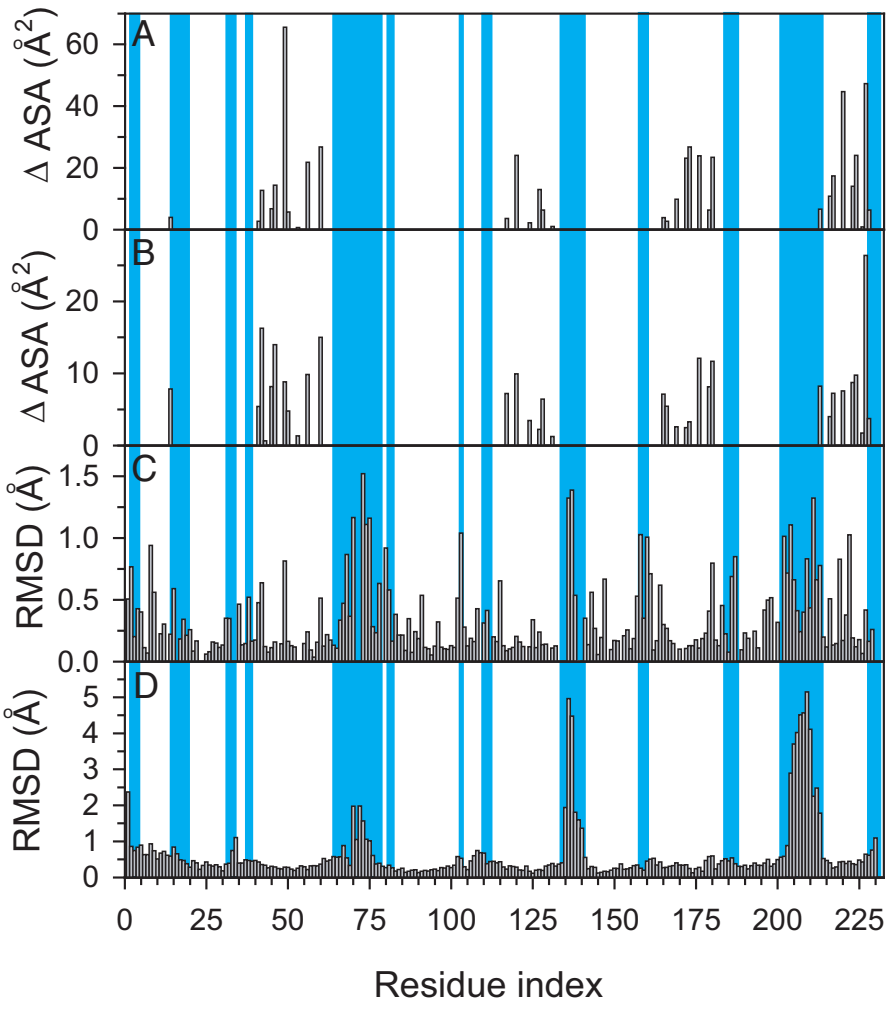

Figure 8
Comparison of residue interactions with structural differences for bound 14-3-3ζ. The average (A) and standard deviations (B) were calculated over the five 14-3-3ζ-peptide interaction profiles. These are shown aligned with the side chain RMSF (C) and the backbone RMSF (D) calculated from the five structures of bound 14-3-3ζ. Regions of residues that are highly exposed to solvent in all complex structures are indicated by the blue-shaded regions.

ordered segments bind to ordered partners indicates that the disorder-associated binding regions have distinct sequence features, even when the bound structure is irregular or sheet instead of helix, and so it should be feasible to develop a specific predictor for each of the different types of MoRFs [137].

One-to-many signaling

Date hubs bind to different proteins at different times. Figure 3 shows how a single region of p53 binds to four different partners. The amino acids involved in each interaction show a significant overlap and no two of these interactions could exist simultaneously. Furthermore, the same residues adopt helix, sheet, and two different irregular structures when associated with the different partners. Finally, the same amino acids are buried to very different extends in each of the molecular associations. These results show very clearly how one segment of disordered protein can bind to multiple partners via the ability to adopt distinct conformations.

The idea that one segment of protein can adopt different secondary structures depending on the context is not new. Many unrelated proteins have identical subsequences of length six, and sometimes even up to length eight, with the same sequences often adopting different secondary structures in different contexts [138,138-140]. Such sequences have been called chameleons for their ability to adopt different structures in different environments [139-145]. Chameleon behavior could be an important feature

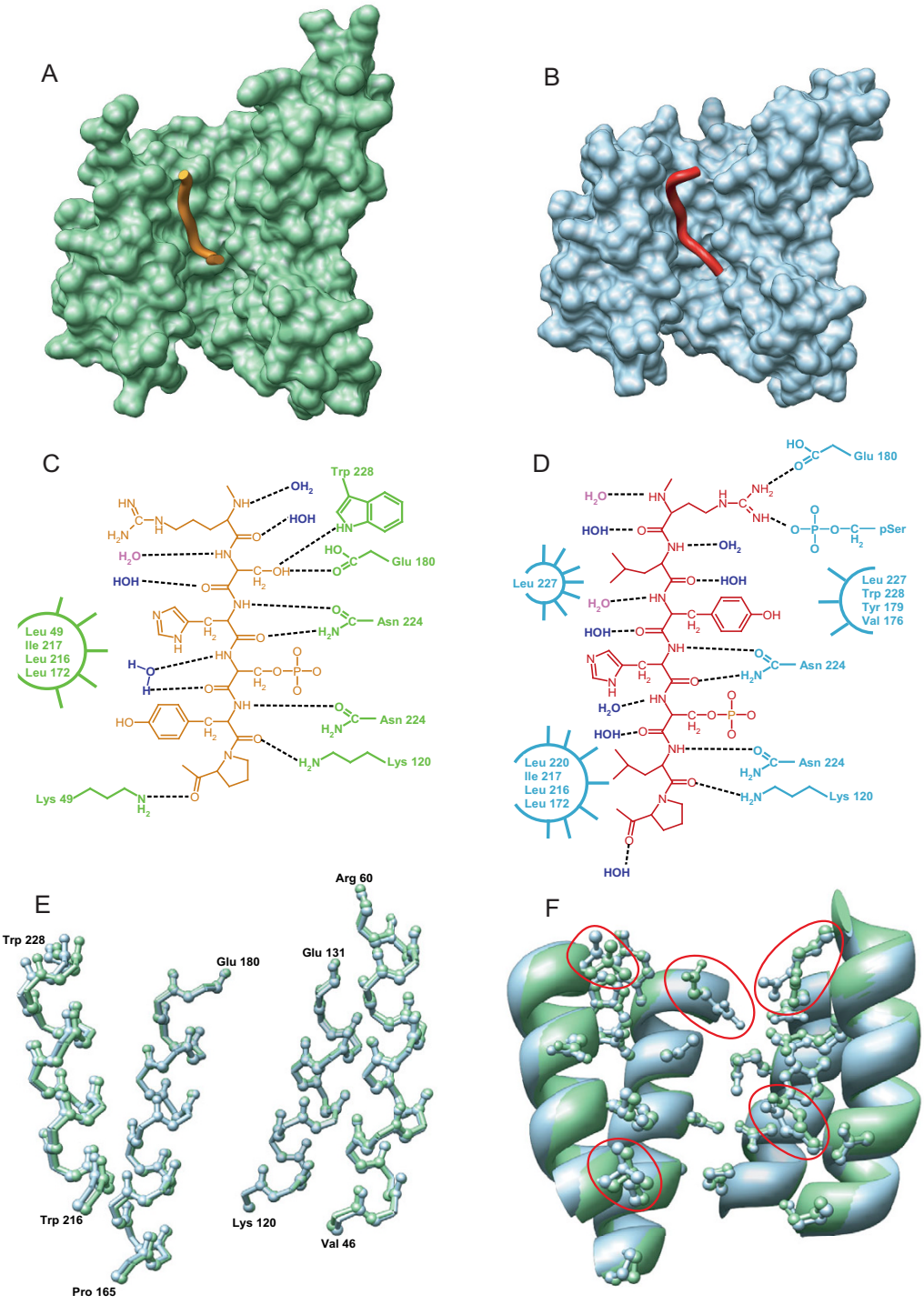

Figure 9

Detailed analysis of 14-3-3ζ peptide binding. The m1 peptide (A, orange ribbon) and m2 peptide (B, red ribbon) bound to 14-3-3 (A and B, shown by the green and blue surface, respectively). Details of 14-3-3 peptide binding are shown by a chemical schematic for the m1 peptide (C) and the m2 peptide (D), where both crystallographic waters (blue) and implicit waters (red) are shown. (E) Superposition of the backbone atoms from the 4 helices with the primary peptide binding residues for m1 (green) and m2 (blue) bound 14-3-3. (F) Superposition of ribbons of the same 4 helices showing the side chains of the residues that participate in m1 (green) and/or m2 (blue) binding.

that enables one disordered region to bind to multiple partners. With different secondary structures and with different side chain participation in the different complexes, it is as if one sequence can be "read" in multiple ways by the various binding partners.

Chameleon behavior occurs for short peptides (octamers), for longer protein fragments and even for entire proteins. For example, the 17 residues-long arginine-rich RNA binding domain (residues 65–81) of the Jembrana disease virus (JDV) Tat protein recognizes two different transactivating response element (TAR) RNA sites, from human and bovine immunodeficiency viruses (HIV and BIV, respectively). The JDV segment adopts different conformations in the two RNA contexts and uses different amino acids for recognition [142]. In addition to the above conformational differences, the JDV domain requires the cyclin T1 protein for high-affinity binding to HIV TAR, but not to BIV TAR [142]. Another protein with chameleon properties is human α-synuclein, which is implicated in Parkinson's disease and in a number of other neurodegenerative disorders known as synucleinopathies. This protein may remain substantially unfolded, or it may adopt an amyloidogenic partially folded structure, or it may fold into α-helical or β-strand species, including both monomeric and oligomeric species. In addition, this protein can form several morphologically different types of aggregates, including oligomers (spheres or doughnuts), amorphous aggregates, and amyloid-like fibrils [34].

Such chameleon sequences likely underlie the multiple specificity binding sites common in p53. For a quick calculation of the implied degree of interface overlap, assume that each residue in a region has equal probability to interact with a partner and consider the C-terminus of p53. The disordered C-terminus (~100 residues) associates with at least 44 distinct partners. The average length of a binding site in this region is ~14 residues, which means that on average only 100/14=7 partners bind at any given residue in the C-terminus. This simple back-of-the-envelope calculation suggests that multiple specificity sequences may be the rule for p53 interactions, rather than a curiosity of a single region. However, available data suggests that interactions do not overlap in a random fashion, but rather interactions are localized to specific regions. For example, consider that the majority of the structures available for the C-terminus of p53 involved the same region of sequence. Therefore, the back-of-the-envelope calculation provides an approximate minimum degree of overlap, where the actual degree of overlap is likely much higher. This idea, which is an extension of a previous proposal [117], further suggests a general mechanism by which hub proteins could bind to such a large multitude of partners, which cannot be explained from the view point of interaction between two structured proteins [51].

Finally, the p53 DBD offers a counter example to the disorder-based view of date hubs. That is, it uses the same or similar face of its globular structure to bind to multiple partners. While the p53 DBD is a folded protein, it does exhibit some remarkable structural differences when bound to difference partners. It seems unlikely that these *local* regions of the p53 DBD structure are well folded in isolation, otherwise the association rate of some or all of these complexes would be relatively low. This idea is supported by the finding that the p53 DBD is only marginally stable at physiological temperature [146]. Therefore, it is plausible that these regions of the monomeric DBD are only transiently folded in solution, where crystallization conditions cause a shift toward the folded state in monomeric crystal structures. The double NIA-NMA plot data (Figure 2B) does not contradict this idea, since it is limited to global analysis and this idea only applies to local regions of the DBD. This idea is conjecture and further experimental or simulation evidence is needed to test this idea. In any event, however, the p53 DBD demonstrates that even proteins generally thought to be well folded, structural changes can still occur in association with multiple specificity.

Many-to-one signaling

In 14-3-3, a common binding groove in a structured dimeric protein can be fitted by multiple, distinct sequences provided by many different binding partners. A recent bioinformatics study [63] found that14-3-3 proteins, and the 14-3-3 binding regions in particular, are predicted to be highly disordered by multiple disorder prediction methods. The authors proposed that 14-3-3 recognition generally involved coupled binding and folding of the recognition region. Our results support this conclusion because the backbone of m1 and m2 peptides are highly hydrated in the bound state (Figure 9C and D), indicating that the binding peptide is likely to be unstructured prior to binding [83].

One idea is that 14-3-3 holds its bound partner in a non-active state [63]. Even though 14-3-3 likely binds to disordered regions in its partners (data herein and [63]), this idea of blocking the active structure could still be true. For example, the productive state of 14-3-3's partner might involve the binding of the partner to a second partner via the same disordered region that binds to 14-3-3, in which case 14-3-3 binding would prevent the formation of the productive complex. Another possibility is that the disordered region exhibits an equilibrium between a bound state that activates the protein and an unbound state that inactivates the protein. The association of the unbound disordered region with 14-3-3 would then hold 14-3-3's

partner in the non-productive state as proposed previously.

We previously suggested that disordered segments with different sequences could use their flexibility to bind to a common binding site, thereby facilitating many-to-one signaling [35]. The multiple recognition of 14-3-3 depends on this mechanism to a considerable degree, with the different peptides taking different paths through the binding cleft and interacting with binding site residues in distinct ways (Figure 6B).

In addition, structured proteins also have a degree of flexibility, and so the binding site backbone and side chain residues can undergo shifts (induced-fit mechanisms) to help accommodate interactions with distinct sequences (Figure 6 and 8). Thus, induced-fit mechanisms are important for structured protein interactions with different partners whether the partners are structured or intrinsically disordered.

The induced-fit mechanisms observed for 14-3-3 and the DNA binding domain of p53 are commonly observed in other situations. For example, tethering, in which a peptide is covalently linked to its protein target to allow detection of low affinity interactions, often results large-scale side chain movements concomitant with peptide binding [147]. Also, when many different MoRFs and their binding partners are examined, induced-fit movements in the structured partners are very commonly observed [137]. Similarly, small backbone shifts and side chain conformational changes are both important for 14-3-3's ability to bind multiple partners. For all of these examples, the associations involve coupled binding and folding for the disordered peptide partner coupled with a near universal classical induced fit for the structured side of the partnership.

One-to-many signaling vs. many-to-one signaling

The p53 C-terminus and 14-3-3 use intrinsic disorder differently with regard to enabling multiple binding specificities. In p53, drastic conformational changes enable distinct surfaces to be exposed to binding partners. In 14-3-3, subtle differences in 14-3-3 conformation and peptide binding locations enable multiple specificities. Why would nature use one mechanism rather than the other for a particular biological role? The interactions of p53 serve to activate or inhibit its primary role as a transcription regulator, while 14-3-3 alters the functions or subcellular localization of many proteins. From this, one can make some highly speculative proposals: (1) disorder binding regions play a passive role in regulation by providing a specific binding site – i.e. the disordered regions are the identification sites of the protein to be regulated [148] – and (2) ordered proteins play the active role – i.e. altering the activity of the proteins they bind to – where recognition of disordered regions allows for a generalized specificity so that a single protein can alter the activity of many others. Validation of the accuracy and generality of these ideas requires further study.

Conclusions

Here we have examined the mechanisms of multiple specificities in two date hub-like hub proteins. Evidence here and elsewhere [37-41], suggests that disordered regions may be an extremely common mechanism by which hub proteins bind to their multitude of partners. The specific examples of p53 and 14-3-3 contrasts the mechanisms by which disorder facilitates multiple recognition, where the former involves drastic conformational differences in a single disordered region and the later involves a variety of subtler changes in order to recognize multiple disordered regions. Finally, it is proposed that the differences between the binding of the disordered region of p53 and the binding of disordered regions to 14-3-3 may have implications for the biological roles of both types of interactions.

Methods

PONDRs VL-XT and VSL1

Predictions of intrinsic disorder in HPV proteins were performed using a set of PONDR® (Predictor Of Natural Disordered Regions) predictors, VL-XT and VSL2. PONDR® VL-XT integrates three feed forward neural networks: the Variously characterized Long, version 1 (VL1) predictor from Romero *et al.* 2001 [84], which predicts non-terminal residues, and the X-ray characterized N- and C- terminal predictors (XT) from Li *et al.* 1999 [149], which predicts terminal residues. Output for the VL1 predictor starts and ends 11 amino acids from the termini. The XT predictors output provides predictions up to 14 amino acids from their respective ends. A simple average is taken for the overlapping predictions; and a sliding window of 9 amino acids is used to smooth the prediction values along the length of the sequence. Unsmoothed prediction values from the XT predictors are used for the first and last 4 sequence positions.

The recently developed Various Short-Long, version 1 (PONDR®-VSL1) algorithm is an ensemble of logistic regression models that predict per-residue order-disorder [85,150]. Two models predict either long (>30 residues) or short (<15 residues) disordered regions based on features similar to those used by VL-XT. The algorithm calculates a weighted average of these predictions, where the weights are determined by a meta-predictor that approximates the likelihood of a long disordered region within its 61-residue window. Predictor inputs include PSI-blast profiles [151], and PHD [152], and PSI-pred [153] secondary structure predictions.

Structure surface and complex interface analysis

Solvent accessible surface area (ASA) was calculated from atomic protein structure numerically using the double cubic lattice method [154] as implemented in the Biochemical Algorithms Library [155]. Using this algorithm, ASA of residues and entire chains can be calculated.

To determine interface areas, for example between two chains, the ASA of each individual chain is calculated, as well as the ASA of the complex. The interface area is then calculated as the change in ASA (ΔASA), i.e. the sum of the individual chain ASA minus the complex ASA. Residues directly involved in interactions were identified from molecular structures as residues with a ΔASA greater than 1 Å^2 [112,113]. All calculations used a probe radius of 1.4 Å, which roughly corresponds to the size of a water molecule.

Order-disorder evaluation from known structure

The work of Gunasekaran et al. has previously shown that, in many cases, the order-disorder state of a protein prior to complex formation is reflected in the complex structure [117]. Specifically, a plot of the normalized monomer area (NMA) – ASA divided by the number of monomer residues – versus the normalized interface area (NIA) – ΔASA divided by the number of monomer residues – effectively distinguishes between ordered and disordered monomers using a linear boundary. This effectiveness of this NMA-NIA plot has been validated on an expanded dataset and an optimal linear boundary has been estimated and evaluated (Oldfield et al., manuscript in preparation). The equation for the novel boundary is:

$$< NMA > = 157.43 - 3.51 < NIA >$$

Since the NMA-NIA plot can only represent one partner of a complex, the double NMA-NIA plot was developed to simultaneously represent both monomers of a binary complex – or complexes that can be treated as binary, such as two monomers bound to a dimer. Rather than plotting the NMA and NIA directly, the Euclidean distance to the order-disorder boundary is calculated, where disordered monomers have a positive distance and ordered monomers have a negative distance. Then the boundary distances of each monomer in a binary complex can be plotted against each other to give an overall order-disorder prediction for the complex. The double NMA-NIA plot is covered in more detail elsewhere (Oldfield et al., manuscript in preparation).

Other structure calculations

The root mean squared fluctuation (RMSF) is a commonly used measure of variability across multiple structure alignments. Here, RMSF of the protein backbone is approximated as the RMSF of the C_α atoms. The equation used is

$$RMSF_i = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(C\alpha_{j,i} - \overline{C\alpha_i}\right)^2}$$

where $C\alpha_{j,i}$ is the position vector of the ith C_α atom of the jth complex and $\overline{C_{\alpha_i}}$ is the averaged position for the ith amino acid from the multiple sequence alignment of N structures. The program MultiProt [156] was used to generate the multiple sequence alignments for RMSF calculation and structure rendering.

To estimate side chain conformation variability among multiple protein structures, the RMSF of side chain residues was calculated. In this calculation, the residue atoms C_α,C_β, backbone carbonyl carbon, and backbone nitrogen were used to align a residue to a selected reference residue of the same type. Thus aligned, the RMSF was calculated over side chain carbons beyond the C_β. Consequently, no side chain RMSF was calculated for Glycine or Alanine residues. The RMSF was also corrected for the number of atoms in the side chain beyond the C_β.

The solvent accessibility of individual residues was calculated relative to an extended Gly-X-Gly model peptide [157], which gives a conservative estimate of relative solvent exposure, i.e. underestimates relative solvent exposure. Residues exposed to solvent were defined as those with an accessible surface area at least 40% of that of the reference area for that residue type. This cutoff is arbitrary, but cutoffs for solvent exposed residues as low as 20% have been used by others, e.g. [158]. Solvent exposures were calculated in the context of binary complexes, which is valid for p53 complexes. In 14-3-3 complexes, 14-3-3 forms homotypic dimmers in addition to binding to phosphopeptides, so residues found to be highly solvent exposed are either actually exposed to solvent or involved in the homodimer interface.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

CJO has done the computational analysis, designed figures and contributed to the manuscript writing. JM, MQY and JYY were involved in finding and analyzing p53 and 14-3-3 binding partners. VNU was involved in planning of experiments, contributed to the manuscript writing and revised the final version. AKD was involved in design and planning of all the experiments, drafted the manuscript and headed the project. All authors have read and approved the final manuscript.

Acknowledgements

This work was supported in part by the grants R01 LM007688-01A1 (to A.K.D and V.N.U.) and GM071714-01A2 (to A.K.D and V.N.U.) from the National Institutes of Health and the Programs of the Russian Academy of Sciences for the "Molecular and cellular biology" and "Fundamental science for medicine" (to V. N. U.).

This article has been published as part of *BMC Genomics* Volume 9 Supplement 1, 2008: The 2007 International Conference on Bioinformatics & Computational Biology (BIOCOMP'07). The full contents of the supplement are available online at http://www.biomedcentral.com/1471-2164/9?issue=S1.

References

1. Goh KI, Oh E, Jeong H, Kahng B, Kim D: **Classification of scale-free networks.** *Proc Natl Acad Sci U S A* 2002, **99(20):**12583-12588.
2. Watts DJ, Strogatz SH: **Collective dynamics of 'small-world' networks.** *Nature* 1998, **393(6684):**440-442.
3. Erdös P, R´nyi A: **On the evolution of random graphs.** *Publ Math Inst Hung Acad Sci* 1960, **5:**17-61.
4. Barabasi AL, Bonabeau E: **Scale-free networks.** *Sci Am* 2003, **288(5):**60-69.
5. Albert R, Jeong H, Barabasi AL: **Error and attack tolerance of complex networks.** *Nature* 2000, **406(6794):**378-382.
6. Jeong H, Mason SP, Barabasi AL, Oltvai ZN: **Lethality and centrality in protein networks.** *Nature* 2001, **411(6833):**41-42.
7. Milgram S: **The small world problem.** *Psycol Today* 1967, **2:**60-67.
8. Bork P, Jensen LJ, von Mering C, Ramani AK, Lee I, Marcotte EM: **Protein interaction networks from yeast to human.** *Curr Opin Struct Biol* 2004, **14(3):**292-299.
9. Barabasi AL, Oltvai ZN: **Network biology: understanding the cell's functional organization.** *Nat Rev Genet* 2004, **5(2):**101-113.
10. Wu CH, Huang H, Nikolskaya A, Hu Z, Barker WC: **The iProClass integrated database for protein functional analysis.** *Comput Biol Chem* 2004, **28(1):**87-96.
11. Huang TW, Tien AC, Huang WS, Lee YC, Peng CL, Tseng HH, Kao CY, Huang CY: **POINT: a database for the prediction of protein-protein interactions based on the orthologous interactome.** *Bioinformatics* 2004, **20(17):**3273-3276.
12. Kelley BP, Yuan B, Lewitter F, Sharan R, Stockwell BR, Ideker T: **PathBLAST: a tool for alignment of protein interaction networks.** *Nucleic Acids Res* 2004, **32:**W83-88. (Web Server issue)
13. von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, Jouffre N, Huynen MA, Bork P: **STRING: known and predicted protein- protein associations, integrated and transferred across organisms.** *Nucleic Acids Res* 2005, **33:**D433-437. (Database issue)
14. Hahn MW, Kern AD: **Comparative genomics of centrality and essentiality in three eukaryotic protein-interaction networks.** *Mol Biol Evol* 2005, **22(4):**803-806.
15. Huang S: **Back to the biology in systems biology, what can we learn from biomolecular networks?** *Brief Funct Genomic Proteomic* 2004, **2(4):**279-297.
16. Han JD, Bertin N, Hao T, Goldberg DS, Berriz GF, Zhang LV, Dupuy D, Walhout AJ, Cusick ME, Roth FP, *et al.*: **Evidence for dynamically organized modularity in the yeast protein-protein interaction network.** *Nature* 2004, **430(6995):**88-93.
17. Hartwell LH, Hopfield JJ, Leibler S, Murray AW: **From molecular to modular cell biology.** *Nature* 1999, **402(6761 Suppl):**C47-52.
18. Cesareni G, Ceol A, Gavrila C, Palazzi LM, Persico M, Schneider MV: **Comparative interactomics.** *FEBS Lett* 2005, **579(8):**1828-1833.
19. von Mering C, Krause R, Snel B, Cornell M, Oliver SG, Fields S, Bork P: **Comparative assessment of large-scale data sets of protein-protein interactions.** *Nature* 2002, **417(6887):**399-403.
20. Bader GD, Hogue CW: **Analyzing yeast protein-protein interaction data obtained from different sources.** *Nat Biotechnol* 2002, **20(10):**991-997.
21. Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y: **A comprehensive two-hybrid analysis to explore the yeast protein interactome.** *Proc Natl Acad Sci U S A* 2001, **98(8):**4569-4574.
22. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, *et al.*: **A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae.** *Nature* 2000, **403(6770):**623-627.
23. Li S, Armstrong CM, Bertin N, Ge H, Milstein S, Boxem M, Vidalain PO, Han JD, Chesneau A, Hao T, *et al.*: **A map of the interactome network of the metazoan C. elegans.** *Science* 2004, **303(5657):**540-543.
24. Giot L, Bader JS, Brouwer C, Chaudhuri A, Kuang B, Li Y, Hao YL, Ooi CE, Godwin B, Vitols E, *et al.*: **A protein interaction map of Drosophila melanogaster.** *Science* 2003, **302(5651):**1727-1736.
25. Han JD, Dupuy D, Bertin N, Cusick ME, Vidal M: **Effect of sampling on topology predictions of protein-protein interaction networks.** *Nat Biotechnol* 2005, **23(7):**839-844.
26. Hasty J, Collins JJ: **Protein interactions. Unspinning the web.** *Nature* 2001, **411(6833):**30-31.
27. Fischer E: **Einfluss der configuration auf die wirkung derenzyme.** *Ber Dt Chem Ges* 1894, **27:**2985-2993.
28. Koshland DE Jr, Ray WJ Jr, Erwin MJ: **Protein structure and enzyme action.** *Fed Proc* 1958, **17(4):**1145-1150.
29. Landsteiner K: **The Specificity of Serological Reactions.** *Mineola, New York: Courier Dover Publications*; 1936.
30. Pauling L: **A theory of the structure and process of formation of antibodies.** *J Am Chem Soc* 1940, **62:**2643-2657.
31. Karush F: **Heterogeneity of the binding sites of bovine serum albumin.** *J Am Chem Soc* 1950, **72:**2705-2713.
32. Meador WE, Means AR, Quiocho FA: **Modulation of calmodulin plasticity in molecular recognition on the basis of x-ray structures.** *Science* 1993, **262(5140):**1718-1721.
33. Kriwacki RW, Hengst L, Tennant L, Reed SI, Wright PE: **Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity.** *Proc Natl Acad Sci U S A* 1996, **93(21):**11504-11509.
34. Uversky VN: **A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders.** *J Biomol Struct Dyn* 2003, **21(2):**211-234.
35. Dunker AK, Garner E, Guilliot S, Romero P, Albrecht K, Hart J, Obradovic Z, Kissinger C, Villafranca JE: **Protein disorder and the evolution of molecular recognition: theory, predictions and observations.** *Pac Symp Biocomput* 1998:473-484.
36. Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN: **Flexible nets. The roles of intrinsic disorder in protein interaction networks.** *Febs J* 2005, **272(20):**5129-5148.
37. Patil A, Nakamura H: **Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks.** *FEBS Lett* 2006, **580(8):**2041-2045.
38. Haynes C, Oldfield CJ, Ji F, Klitgord N, Cusick ME, Radivojac P, Uversky VN, Vidal M, Iakoucheva LM: **Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes.** *PLoS Comput Biol* 2006, **2(8):**e100.
39. Ekman D, Light S, Bjorklund AK, Elofsson A: **What properties characterize the hub proteins of the protein-protein interaction network of Saccharomyces cerevisiae?** *Genome Biol* 2006, **7(6):**R45.
40. Dosztanyi Z, Chen J, Dunker AK, Simon I, Tompa P: **Disorder and sequence repeats in hub proteins and their implications for network evolution.** *J Proteome Res* 2006, **5(11):**2985-2995.
41. Singh GP, Ganapathi M, Sandhu KS, Dash D: **Intrinsic unstructuredness and abundance of PEST motifs in eukaryotic proteomes.** *Proteins* 2006, **62(2):**309-315.
42. Marinissen MJ, Gutkind JS: **Scaffold proteins dictate Rho GTPase-signaling specificity.** *Trends Biochem Sci* 2005, **30(8):**423-426.
43. Jaffe AB, Aspenstrom P, Hall A: **Human CNK1 acts as a scaffold protein, linking Rho and Ras signal transduction pathways.** *Mol Cell Biol* 2004, **24(4):**1736-1746.
44. Jaffe AB, Hall A: **Rho GTPases: biochemistry and biology.** *Annu Rev Cell Dev Biol* 2005, **21:**247-269.
45. Hohenstein P, Giles RH: **BRCA1: a scaffold for p53 response?** *Trends Genet* 2003, **19(9):**489-494.
46. Luo W, Lin SC: **Axin: a master scaffold for multiple signaling pathways.** *Neurosignals* 2004, **13(3):**99-113.
47. Rui Y, Xu Z, Lin S, Li Q, Rui H, Luo W, Zhou HM, Cheung PY, Wu Z, Ye Z, *et al.*: **Axin stimulates p53 functions by activation of HIPK2 kinase through multimeric complex formation.** *Embo J* 2004, **23(23):**4583-4594.

48. Salahshor S, Woodgett JR: **The links between axin and carcinogenesis.** *J Clin Pathol* 2005, **58(3):**225-236.
49. Wong W, Scott JD: **AKAP signalling complexes: focal points in space and time.** *Nat Rev Mol Cell Biol* 2004, **5(12):**959-970.
50. Carpousis AJ: **The RNA Degradosome of Escherichia coli: A Multiprotein mRNA-Degrading Machine Assembled on RNase E.** *Annu Rev Microbiol* 2007.
51. Kim PM, Lu LJ, Xia Y, Gerstein MB: **Relating three-dimensional structures to protein networks provides evolutionary insights.** *Science* 2006, **314(5807):**1938-1941.
52. Anderson CW, Appella E: **Signaling to the p53 tumor suppressor through pathways activated by genotoxic and nongenotoxic stress.** In In: *Handbook of Cell Signaling* Edited by: *Bradshaw RA; Dennis EA. New York: Academic Press*; 2004:237-247.
53. Hollstein M, Sidransky D, Vogelstein B, Harris CC: **p53 mutations in human cancers.** *Science* 1991, **253(5015):**49-53.
54. Zhao R, Gish K, Murphy M, Yin Y, Notterman D, Hoffman WH, Tom E, Mack DH, Levine AJ: **Analysis of p53-regulated gene expression patterns using oligonucleotide arrays.** *Genes Dev* 2000, **14(8):**981-993.
55. Anderson CW, Appella E: **Signaling to the p53 tumor suppressor through pathways activated by genotoxic and nongenotoxic stress.** In In: *Handbook of Cell Signaling* Edited by: *Bradshaw RA; Dennis EA. New York: Academic Press*; 2003:237-247.
56. Dougherty MK, Morrison DK: **Unlocking the code of 14-3-3.** *J Cell Sci* 2004, **117(Pt 10):**1875-1884.
57. Pozuelo Rubio M, Geraghty KM, Wong BH, Wood NT, Campbell DG, Morrice N, Mackintosh C: **14-3-3-affinity purification of over 200 human phosphoproteins reveals new links to regulation of cellular metabolism, proliferation and trafficking.** *Biochem J* 2004, **379(Pt 2):**395-408.
58. Meek SE, Lane WS, Piwnica-Worms H: **Comprehensive proteomic analysis of interphase and mitotic 14-3-3-binding proteins.** *J Biol Chem* 2004, **279(31):**32046-32054.
59. Jin J, Smith FD, Stark C, Wells CD, Fawcett JP, Kulkarni S, Metalnikov P, O'Donnell P, Taylor P, Taylor L, *et al.*: **Proteomic, functional, and domain-based analysis of in vivo 14-3-3 binding proteins involved in cytoskeletal regulation and cellular organization.** *Curr Biol* 2004, **14(16):**1436-1450.
60. Yaffe MB: **How do 14-3-3 proteins work?-- Gatekeeper phosphorylation and the molecular anvil hypothesis.** *FEBS Lett* 2002, **513(1):**53-57.
61. Rittinger K, Budman J, Xu J, Volinia S, Cantley LC, Smerdon SJ, Gamblin SJ, Yaffe MB: **Structural analysis of 14-3-3 phosphopeptide complexes identifies a dual role for the nuclear export signal of 14-3-3 in ligand binding.** *Mol Cell* 1999, **4(2):**153-166.
62. Yaffe MB, Rittinger K, Volinia S, Caron PR, Aitken A, Leffers H, Gamblin SJ, Smerdon SJ, Cantley LC: **The structural basis for 14-3-3: phosphopeptide binding specificity.** *Cell* 1997, **91(7):**961-971.
63. Bustos DM, Iglesias AA: **Intrinsic disorder is a key characteristic in partners that bind 14-3-3 proteins.** *Proteins* 2006, **63(1):**35-42.
64. Lowe ED, Tews I, Cheng KY, Brown NR, Gul S, Noble ME, Gamblin SJ, Johnson LN: **Specificity determinants of recruitment peptides bound to phospho-CDK2/cyclin A.** *Biochemistry* 2002, **41(52):**15625-15634.
65. Avalos JL, Celic I, Muhammad S, Cosgrove MS, Boeke JD, Wolberger C: **Structure of a Sir2 enzyme bound to an acetylated p53 peptide.** *Mol Cell* 2002, **10(3):**523-535.
66. Mujtaba S, He Y, Zeng L, Yan S, Plotnikova O, Sachchidanand , Sanchez R, Zeleznik-Le NJ, Ronai Z, Zhou MM: **Structural mechanism of the bromodomain of the coactivator CBP in p53 transcriptional activation.** *Mol Cell* 2004, **13(2):**251-263.
67. Wu H, Maciejewski MW, Marintchev A, Benashski SE, Mullen GP, King SM: **Solution structure of a dynein motor domain associated light chain.** *Nat Struct Biol* 2000, **7(7):**575-579.
68. Chuikov S, Kurash JK, Wilson JR, Xiao B, Justin N, Ivanov GS, McKinney K, Tempst P, Prives C, Gamblin SJ, *et al.*: **Regulation of p53 activity through lysine methylation.** *Nature* 2004, **432(7015):**353-360.
69. Poux AN, Marmorstein R: **Molecular basis for Gcn5/PCAF histone acetyltransferase selectivity for histone and nonhistone substrates.** *Biochemistry* 2003, **42(49):**14366-14374.
70. Bochkareva E, Kaustov L, Ayed A, Yi GS, Lu Y, Pineda-Lucena A, Liao JC, Okorokov AL, Milner J, Arrowsmith CH, *et al.*: **Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A.** *Proc Natl Acad Sci U S A* 2005, **102(43):**15412-15417.
71. Kussie PH, Gorina S, Marechal V, Elenbaas B, Moreau J, Levine AJ, Pavletich NP: **Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain.** *Science* 1996, **274(5289):**948-953.
72. Di Lello P, Jenkins LM, Jones TN, Nguyen BD, Hara T, Yamaguchi H, Dikeakos JD, Appella E, Legault P, Omichinski JG: **Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIIH and the activation domain of p53.** *Mol Cell* 2006, **22(6):**731-740.
73. Kuszewski J, Gronenborn AM, Clore GM: **Improving the packing and accuracy of NMR structures with a pseudopotential for the radius of gyration.** *Journal of the American Chemical Society* 1999, **121(10):**2337-2338.
74. Cho Y, Gorina S, Jeffrey PD, Pavletich NP: **Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations.** *Science* 1994, **265(5170):**346-355.
75. Joo WS, Jeffrey PD, Cantor SB, Finnin MS, Livingston DM, Pavletich NP: **Structure of the 53BP1 BRCT region bound to p53 and its comparison to the Brca1 BRCT structure.** *Genes Dev* 2002, **16(5):**583-593.
76. Gorina S, Pavletich NP: **Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2.** *Science* 1996, **274(5289):**1001-1005.
77. Obsil T, Ghirlando R, Klein DC, Ganguly S, Dyda F: **Crystal structure of the 14-3-3zeta: serotonin N-acetyltransferase complex. a role for scaffolding in enzyme regulation.** *Cell* 2001, **105(2):**257-267.
78. Petosa C, Masters SC, Bankston LA, Pohl J, Wang B, Fu H, Liddington RC: **14-3-3zeta binds a phosphorylated Raf peptide and an unphosphorylated peptide via its conserved amphipathic groove.** *J Biol Chem* 1998, **273(26):**16305-16310.
79. Lilyestrom W, Klein MG, Zhang R, Joachimiak A, Chen XS: **Crystal structure of SV40 large T-antigen bound to p53: interplay between a viral oncoprotein and a cellular tumor suppressor.** *Genes Dev* 2006, **20(17):**2373-2382.
80. Oldfield CJ, Meng J, Yang JY, Uversky VN, Dunker AK: **Intrinsic disorder in protein-protein interaction networks: Case studies of complexes involving p53 and 14-3-3.** In: *The 2007 International Conference on Bioinformatics and Computational Biology: 2007* 2007. In press.
81. Dawson R, Muller L, Dehner A, Klein C, Kessler H, Buchner J: **The N-terminal domain of p53 is natively unfolded.** *J Mol Biol* 2003, **332(5):**1131-1141.
82. Lee H, Mok KH, Muhandiram R, Park KH, Suk JE, Kim DH, Chang J, Sung YC, Choi KY, Han KH: **Local structural elements in the mostly unstructured transcriptional activation domain of human p53.** *J Biol Chem* 2000, **275(38):**29426-29432.
83. Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK: **The importance of intrinsic disorder for protein phosphorylation.** *Nucleic Acids Res* 2004, **32(3):**1037-1049.
84. Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, Dunker AK: **Sequence complexity of disordered protein.** *Proteins* 2001, **42(1):**38-48.
85. Obradovic Z, Peng K, Vucetic S, Radivojac P, Dunker AK: **Exploiting heterogeneous sequence properties improves prediction of protein disorder.** *Proteins* 2005, **61(Suppl 7):**176-182.
86. Peng K, Radivojac P, Vucetic S, Dunker AK, Obradovic Z: **Length-dependent prediction of protein intrinsic disorder.** *BMC Bioinformatics* 2006, **7:**208.
87. Garner E, Romero P, Dunker AK, Brown C, Obradovic Z: **Predicting Binding Regions within Disordered Proteins.** *Genome Inform Ser Workshop Genome Inform* 1999, **10:**41-50.
88. Oldfield CJ, Cheng Y, Cortese MS, Romero P, Uversky VN, Dunker AK: **Coupled folding and binding with alpha-helix-forming molecular recognition elements.** *Biochemistry* 2005, **44(37):**12454-12470.
89. Callaghan AJ, Aurikko JP, Ilag LL, Gunter Grossmann J, Chandran V, Kuhnel K, Poljak L, Carpousis AJ, Robinson CV, Symmons MF, *et al.*: **Studies of the RNA degradosome-organizing domain of the Escherichia coli ribonuclease RNase E.** *J Mol Biol* 2004, **340(5):**965-979.
90. Cheng Y, Oldfield CJ, Romero P, Uversky VN, Dunker AK: **Mining α-helix-forming molecular recognition features α-MoRFs**

with cross species sequence alignments. *Biochemistry* 2007. In press
91. Iakoucheva LM, Kimzey AL, Masselon CD, Bruce JE, Garner EC, Brown CJ, Dunker AK, Smith RD, Ackerman EJ: **Identification of intrinsic order and disorder in the DNA repair protein XPA.** *Protein Sci* 2001, **10(3):**560-571.
92. Adkins JN, Lumb KJ: **Intrinsic structural disorder and sequence features of the cell cycle inhibitor p57Kip2.** *Proteins* 2002, **46(1):**1-7.
93. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z: **Intrinsic disorder and protein function.** *Biochemistry* 2002, **41(21):**6573-6582.
94. Iakoucheva LM, Brown CJ, Lawson JD, Obradovic Z, Dunker AK: **Intrinsic disorder in cell-signaling and cancer-associated proteins.** *J Mol Biol* 2002, **323(3):**573-584.
95. Longhi S, Receveur-Brechot V, Karlin D, Johansson K, Darbon H, Bhella D, Yeo R, Finet S, Canard B: **The C-terminal domain of the measles virus nucleoprotein is intrinsically disordered and folds upon binding to the C-terminal moiety of the phosphoprotein.** *J Biol Chem* 2003, **278(20):**18638-18648.
96. Karlin D, Ferron F, Canard B, Longhi S: **Structural disorder and modular organization in Paramyxovirinae N and P.** *J Gen Virol* 2003, **84(Pt 12):**3239-3252.
97. Munishkina LA, Fink AL, Uversky VN: **Conformational prerequisites for formation of amyloid fibrils from histones.** *J Mol Biol* 2004, **342(4):**1305-1324.
98. Bandaru V, Cooper W, Wallace SS, Doublie S: **Overproduction, crystallization and preliminary crystallographic analysis of a novel human DNA-repair enzyme that recognizes oxidative DNA damage.** *Acta Crystallogr D Biol Crystallogr* 2004, **60(Pt 6):**1142-1144.
99. Oldfield CJ, Ulrich EL, Cheng Y, Dunker AK, Markley JL: **Addressing the intrinsic disorder bottleneck in structural proteomics.** *Proteins* 2005, **59(3):**444-453.
100. Hansen JC, Lu X, Ross ED, Woody RW: **Intrinsic protein disorder, amino acid composition, and histone terminal domains.** *J Biol Chem* 2006, **281(4):**1853-1856.
101. Haag Breese E, Uversky VN, Georgiadis MM, Harrington MA: **The disordered amino-terminus of SIMPL interacts with members of the 70-kDa heat-shock protein family.** *DNA Cell Biol* 2006, **25(12):**704-714.
102. Radivojac P, Vucetic S, O'Connor TR, Uversky VN, Obradovic Z, Dunker AK: **Calmodulin signaling: analysis and prediction of a disorder-dependent molecular recognition.** *Proteins* 2006, **63(2):**398-410.
103. Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VN, Dunker AK: **Intrinsic disorder in transcription factors.** *Biochemistry* 2006, **45(22):**6873-6888.
104. Uversky VN, Roman A, Oldfield CJ, Dunker AK: **Protein intrinsic disorder and human papillomaviruses: increased amount of disorder in E6 and E7 oncoproteins from high risk HPVs.** *J Proteome Res* 2006, **5(8):**1829-1842.
105. Cheng Y, LeGall T, Oldfield CJ, Dunker AK, Uversky VN: **Abundance of intrinsic disorder in protein associated with cardiovascular disease.** *Biochemistry* 2006, **45(35):**10448-10460.
106. Sigalov AB, Aivazian DA, Uversky VN, Stern LJ: **Lipid-Binding Activity of Intrinsically Unstructured Cytoplasmic Domains of Multichain Immune Recognition Receptor Signaling Subunits.** *Biochemistry* 2006, **45(51):**15731-15739.
107. Singh VK, Zhou Y, Marsh JA, Uversky VN, Forman-Kay JD, Liu J, Jia Z: **Synuclein-gamma targeting peptide inhibitor that enhances sensitivity of breast cancer cells to antimicrotubule drugs.** *Cancer Res* 2007, **67(2):**626-633.
108. Ng KP, Potikyan G, Savene RO, Denny CT, Uversky VN, Lee KA: **Multiple aromatic side chains within a disordered structure are critical for transcription and transforming activity of EWS family oncoproteins.** *Proc Natl Acad Sci U S A* 2007, **104(2):**479-484.
109. Radivojac P, Iakoucheva LM, Oldfield CJ, Obradovic Z, Uversky VN, Dunker AK: **Intrinsic disorder and functional proteomics.** *Biophys J* 2007, **92(5):**1439-1456.
110. Jones S, Thornton JM: **Protein-protein interactions: a review of protein dimer structures.** *Prog Biophys Mol Biol* 1995, **63(1):**31-65.
111. Jones S, Thornton JM: **Principles of protein-protein interactions.** *Proc Natl Acad Sci U S A* 1996, **93(1):**13-20.
112. Jones S, Thornton JM: **Analysis of protein-protein interaction sites using surface patches.** *J Mol Biol* 1997, **272(1):**121-132.
113. Jones S, Thornton JM: **Prediction of protein-protein interaction sites using patch analysis.** *J Mol Biol* 1997, **272(1):**133-143.
114. Larsen TA, Olson AJ, Goodsell DS: **Morphology of protein-protein interfaces.** *Structure* 1998, **6(4):**421-427.
115. Lo Conte L, Chothia C, Janin J: **The atomic structure of protein-protein recognition sites.** *J Mol Biol* 1999, **285(5):**2177-2198.
116. Smith GR, Sternberg MJ, Bates PA: **The relationship between the flexibility of proteins and their conformational states on forming protein-protein complexes with an application to protein-protein docking.** *J Mol Biol* 2005, **347(5):**1077-1101.
117. Gunasekaran K, Tsai CJ, Kumar S, Zanuy D, Nussinov R: **Extended disordered proteins: targeting function with less scaffold.** *Trends Biochem Sci* 2003, **28(2):**81-85.
118. Ho WC, Luo C, Zhao K, Chai X, Fitzgerald MX, Marmorstein R: **High-resolution structure of the p53 core domain: implications for binding small-molecule stabilizing compounds.** *Acta Crystallogr D Biol Crystallogr* 2006, **62(Pt 12):**1484-1493.
119. Williams RS, Green R, Glover JN: **Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1.** *Nat Struct Biol* 2001, **8(10):**838-842.
120. Ferreon JC, Volk DE, Luxon BA, Gorenstein DG, Hilser VJ: **Solution structure, dynamics, and thermodynamics of the native state ensemble of the Sem-5 C-terminal SH3 domain.** *Biochemistry* 2003, **42(19):**5582-5591.
121. Uhrinova S, Uhrin D, Powers H, Watt K, Zheleva D, Fischer P, McInnes C, Barlow PN: **Structure of free MDM2 N-terminal domain reveals conformational adjustments that accompany p53-binding.** *J Mol Biol* 2005, **350(3):**587-598.
122. Rojas JR, Trievel RC, Zhou J, Mo Y, Li X, Berger SL, Allis CD, Marmorstein R: **Structure of Tetrahymena GCN5 bound to coenzyme A and a histone H3 peptide.** *Nature* 1999, **401(6748):**93-98.
123. Jeffrey PD, Gorina S, Pavletich NP: **Crystal structure of the tetramerization domain of the p53 tumor suppressor at 1.7 angstroms.** *Science* 1995, **267(5203):**1498-1502.
124. Avalos JL, Boeke JD, Wolberger C: **Structural basis for the mechanism and regulation of Sir2 enzymes.** *Mol Cell* 2004, **13(5):**639-648.
125. Sachchidanand , Resnick-Silverman L, Yan S, Mutjaba S, Liu WJ, Zeng L, Manfredi JJ, Zhou MM: **Target structure-based discovery of small molecules that block human p53 and CREB binding protein association.** *Chem Biol* 2006, **13(1):**81-90.
126. Ostendorr TD: **Structure and function of the metal-binding protein S100B and its interaction with the receptor for advanced glycation end products.** In *Ph.D. Konstanz: University of Konstanz*; 2007.
127. Kwon T, Chang JH, Kwak E, Lee CW, Joachimiak A, Kim YC, Lee J, Cho Y: **Mechanism of histone lysine methyl transfer revealed by the structure of SET7/9-AdoMet.** *Embo J* 2003, **22(2):**292-303.
128. Macdonald N, Welburn JP, Noble ME, Nguyen A, Yaffe MB, Clynes D, Moggs JG, Orphanides G, Thomson S, Edmunds JW, *et al.*: **Molecular basis for the recognition of phosphorylated and phosphoacetylated histone h3 by 14-3-3.** *Mol Cell* 2005, **20(2):**199-211.
129. Bourhis JM, Johansson K, Receveur-Brechot V, Oldfield CJ, Dunker KA, Canard B, Longhi S: **The C-terminal domain of measles virus nucleoprotein belongs to the class of intrinsically disordered proteins that fold upon binding to their physiological partner.** *Virus Res* 2004, **99(2):**157-167.
130. Kingston RL, Hamel DJ, Gay LS, Dahlquist FW, Matthews BW: **Structural basis for the attachment of a paramyxoviral polymerase to its template.** *Proc Natl Acad Sci U S A* 2004, **101(22):**8301-8306.
131. Chandran V, Luisi BF: **Recognition of enolase in the Escherichia coli RNA degradosome.** *J Mol Biol* 2006, **358(1):**8-15.
132. Puntervoll P, Linding R, Gemund C, Chabanis-Davidson S, Mattingsdal M, Cameron S, Martin DM, Ausiello G, Brannetti B, Costantini A, *et al.*: **ELM server: A new resource for investigating short functional sites in modular eukaryotic proteins.** *Nucleic Acids Res* 2003, **31(13):**3625-3630.
133. Neduva V, Linding R, Su-Angrand I, Stark A, de Masi F, Gibson TJ, Lewis J, Serrano L, Russell RB: **Systematic discovery of new rec-**

ognition peptides mediating protein interaction networks. *PLoS Biol* 2005, **3(12):**e405.
134. Neduva V, Russell RB: **Linear motifs: evolutionary interaction switches.** *FEBS Lett* 2005, **579(15):**3342-3345.
135. Fuxreiter M, Tompa P, Simon I: **Local structural disorder imparts plasticity on linear motifs.** *Bioinformatics* 2007, **23(8):**950-956.
136. Dunker AK: **Another window into disordered protein function.** *Structure* 2007, **15(9):**1026-1028.
137. Vacic V, Oldfield CJ, Mohan A, Radivojac P, Cortese MS, Uversky VN, Dunker AK: **Characterization of molecular recognition features, MoRFs, and their binding partners.** *J Proteome Res* 2007, **6(6):**2351-2366.
138. Kabsch W, Sander C: **Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features.** *Biopolymers* 1983, **22(12):**2577-2637.
139. Minor DL Jr, Kim PS: **Context-dependent secondary structure formation of a designed protein sequence.** *Nature* 1996, **380(6576):**730-734.
140. Jacoboni I, Martelli PL, Fariselli P, Compiani M, Casadio R: **Predictions of protein segments with the same aminoacid sequence and different secondary structure: a benchmark for predictive methods.** *Proteins* 2000, **41(4):**535-544.
141. Mezei M: **Chameleon sequences in the PDB.** *Protein Eng* 1998, **11(6):**411-414.
142. Smith CA, Calabro V, Frankel AD: **An RNA-binding chameleon.** *Mol Cell* 2000, **6(5):**1067-1076.
143. Yoon S, Jung H: **Analysis of chameleon sequences by energy decomposition on a pairwise per-residue basis.** *Protein J* 2006, **25(5):**361-368.
144. Guo JT, Jaromczyk JW, Xu Y: **Analysis of chameleon sequences and their implications in biological processes.** *Proteins* 2007, **67(3):**548-558.
145. Takano K, Katagiri Y, Mukaiyama A, Chon H, Matsumura H, Koga Y, Kanaya S: **Conformational contagion in a protein: Structural properties of a chameleon sequence.** *Proteins* 2007.
146. Bullock AN, Henckel J, Fersht AR: **Quantitative analysis of residual folding and DNA binding in mutant p53 core domain: definition of mutant states for rescue in cancer therapy.** *Oncogene* 2000, **19(10):**1245-1256.
147. Erlanson DA, Wells JA, Braisted AC: **Tethering: fragment-based drug discovery.** *Annu Rev Biophys Biomol Struct* 2004, **33:**199-223.
148. Uversky VN, Oldfield CJ, Dunker AK: **Showing your ID: Intrinsic disorder as an ID for recognition, regulation and cell signalling.** *J Mol Recognit* 2005, **18(5):**343-384.
149. Li X, Romero P, Rani M, Dunker AK, Obradovic Z: **Predicting Protein Disorder for N-, C-, and Internal Regions.** *Genome Inform Ser Workshop Genome Inform* 1999, **10:**30-40.
150. Peng K, Vucetic S, Radivojac P, Brown CJ, Dunker AK, Obradovic Z: **Optimizing long intrinsic disorder predictors with protein evolutionary information.** *J Bioinform Comput Biol* 2005, **3(1):**35-60.
151. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25(17):**3389-3402.
152. Rost B, Sander C, Schneider R: **PHD--an automatic mail server for protein secondary structure prediction.** *Comput Appl Biosci* 1994, **10(1):**53-60.
153. Jones DT, Ward JJ: **Prediction of disordered regions in proteins from position specific score matrices.** *Proteins* 2003, **53(Suppl 6):**573-578.
154. Eisenhaber F, Lijnzaad P, Argos P, Sander C, Scharf M: **The Double Cubic Lattice Method—Efficient Approaches to Numerical-Integration of Surface-Area and Volume and to Dot Surface Contouring of Molecular Assemblies.** *J Comput Chem* 1995, **16(3):**273-284.
155. Kohlbacher O, Lenhof HP: **BALL--rapid software prototyping in computational molecular biology. Biochemicals Algorithms Library.** *Bioinformatics* 2000, **16(9):**815-824.
156. Shatsky M, Nussinov R, Wolfson HJ: **A method for simultaneous alignment of multiple protein structures.** *Proteins* 2004, **56(1):**143-156.
157. Chothia C: **Structural invariants in protein folding.** *Nature* 1975, **254(5498):**304-308.
158. Ma B, Elkayam T, Wolfson H, Nussinov R: **Protein-protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces.** *Proc Natl Acad Sci U S A* 2003, **100(10):**5772-5777.