Intrinsic Disorder Is a Key Characteristic in Partners That Bind 14-3-3 Proteins

Diego M. Bustos,[1]* and Alberto A. Iglesias[2]
[1]*Instituto Tecnológico de Chascomús - IIB-INTECH, Camino Circunvalación, Chascomús, Argentina*
[2]*Laboratorio de Enzimología Molecular, Facultad de Bioquímica y Cs. Biológicas, Universidad Nacional Litoral, Santa Fe, Argentina*

ABSTRACT **Proteins named 14-3-3 can bind more than 200 different proteins, mostly (but not exclusively) when they are at a phosphorylated state. These partner proteins are involved in different cellular processes, such as cell signaling, transcription factors, cellular morphology, and metabolism; this suggests pleiotropic functionality for 14-3-3 proteins. Recent efforts to establish a rational classification of 14-3-3 binding partners showed neither structural nor functional relatedness in this group of proteins. Using three natural predictors of disorder in proteins, and the structural available information, we show that >90% of 14-3-3 protein partners contain disordered regions. This percentage is significantly high when compared with recent studies on cell signaling and cancer-related proteins or RNA chaperons. More important, almost all 14-3-3-binding sites are inside disordered regions, this reinforcing the importance of structural disorder in this class of proteins. We also propose that a disorder-to-order transition occurs in the binding of 14-3-3 proteins with their partners. We discuss the consequences of the latter for consensus binding sequences, specificity, affinity, and thermodynamic control. Proteins 2006;63:35–42.**

Key words: 14-3-3-binding site; protein–protein interaction; intrinsically disordered protein

INTRODUCTION

The 14-3-3-protein family was firstly identified in 1967 by Moore and Perez[1] during a systematic classification of brain proteins. Since then, the importance of 14-3-3 proteins has risen to a key position in cell biology, because they were identified to contribute to a wide range of vital regulatory processes including signal transduction, apoptosis, cell cycle progression, DNA replication, and cell malignant transformation.[2] This protein family is highly conserved and ubiquitously expressed, with at least seven isoforms in mammals, up to 15 isoforms present in plants, and two isoforms identified in yeast, *Drosophila melanogaster*, and *Caenorhabditis elegans*.[2]

One proposed functional model is that 14-3-3 binds to the specific target as a "molecular anvil" that causes conformational changes in the partner that can alter its enzymatic (biological) activity, or mask or reveal specific motifs that regulate its localization, activity, phosphorylation state, and/or stability.[3] Nonetheless, 14-3-3 also exhibits its adaptor function, mediating interaction between two different binding partners. Many molecules have been found to associate with 14-3-3 proteins in a phosphorylation-dependent manner, mainly after the phosphoserine/phosphothreonine-binding specificity exhibited by the process,[4] as defined by Muslin and colleagues[5] in the mid-1990s. The number of these proteins has now surpassed 200 (for a complete list, see references 6–8), and it has been speculated that 14-3-3-interacting proteins potentially amount to approximately 0.6% of the human proteome.[8]

The 14-3-3 proteins bind to the consensus motif RSX-pSXP (X representing any amino acid, and pS phosphoserine).[5] Because binding is dependent on the phosphorylation of the central serine, this mechanism allows the conditional association of 14-3-3 proteins with protein partners containing this motif.[4] Yaffe and coworkers[9] further refined the structural features after screening of phosphopeptide libraries and identification of two different assigned binding motifs: the above indicated motif RSXpSXP (mode 1) and RXY/FXpSXP (mode 2). More recently, a mode 3 motif for 14-3-3-binding has been assigned to: pS/T [X(1-2)]-COOH (—COOH being the C-terminus).[10] However, numerous exceptions were found, with changes in the amino acids specified in each of the three binding motif modes. Bridges and Moorhead[11] have proposed an intuitive explanation for the capability of 14-3-3 to bind to targets with variations in the canonical sequences. They proposed that two "imperfect" sites could bind to the 14-3-3 dimer with affinity enough if they are in antiparallel orientation and are able to achieve the correct spacing. A binding mode involving two "imperfect" sites may be the more common interaction between 14-3-3 proteins and their targets, because they are more likely to

---

The Supplementary Material referred to in this article can be found at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/

Grant sponsor: ANPCyT; Grant numbers: PICTO 1-13241, PICT'03 1-14733, and PAV 137.

*Correspondence to: Diego M. Bustos, IIB-INTECH, Camino Circunv. Laguna km 6, Casilla de Correo 164, B7130IWA, Chascomús, Argentina. E-mail: dbustos@intech.gov.ar

Received 12 April 2005; Revised 21 October 2005; Accepted 24 October 2005

Published online 27 January 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20888

dissociate from those motifs than from the "perfect" consensus sequence mode.[11]

A great number of recent evidence[12–15] indicates that for many proteins (and protein domains) the native and functional state is intrinsically disordered/unstructured/unfolded. This means that these proteins adopt regular secondary structure but they lack fixed tertiary arrangement, because intrinsic disorder does not necessarily mean lack of secondary structure.[16] It was also proposed[16,17] that intrinsically disordered proteins or domains are involved in molecular recognition and protein–protein interactions. Molecular recognition involving intrinsically disordered proteins has two important features providing advantages for signaling and regulation. First, disordered regions can bind their targets with high specificity and low affinity. Second, intrinsic disorder promotes binding diversity by enabling proteins to interact with numerous partners (one-to-many or many-to-one).[16] Transitions among native, disordered state, and a globular structure (induced by phosphorylation or other type of interaction) may also provide thermodynamic regulation of binding.

Although the transition disorder-to-order is disfavored in terms of entropy, the formation of the complex could be driven by a large enthalpy change associated with the favorable hydrogen bonding interactions involving the phosphoryl group.[13] This unfavorable energy term uncouples binding strength from specificity and renders highly specific reversible interactions, which is fundamental in regulation. This effect was explicitly demonstrated by comparing the interaction of two different proteins with the KIX domain of the CBP protein. The interaction of one protein (c-Myb) with KIX is constitutive, whereas the binding of the other protein (CREB) is inducible and triggered by phosphorylation (only phosphoCREB binds).[13,14,18] Inducible binding to KIX was proved as linked to a shift from disorder-to-order, with a negative change in entropy upon binding of CREB. This is opposed to the binding of c-Myb, which is folded into a helical structure in its free state. Disorder-to-order transition has also been demonstrated for other intrinsically disordered proteins, such as FlgM,[19] protein kinase inhibitor,[20] eukaryotic translation initiation factor 4E binding protein,[21] and stathmin,[22] and it may also apply to all unstructured proteins functioning in molecular recognition.[14]

Many proteins acting as partners of 14-3-3-binding could undergo disorder-to-order transition. Consequently, an important extension of the gatekeeper/enhancer hypothesis[3] is that, unless the two molecules snap into place like puzzle pieces, the target is likely to undergo conformational deformations when its two phosphate groups are inserted into the 14-3-3 binding groove.[4] Herein, we apply three different predictors of intrinsically disordered protein regions to investigate disorder in partners for the binding of 14-3-3 proteins. We analyze relationships between disordered regions and 14-3-3-binding sites. The results support a general involvement of intrinsically unstructured proteins in binding to 14-3-3 proteins, and a mechanistic model for the interaction is proposed.

METHODS

Sequences and Datasets

The dataset was composed of well-know human proteins, one polypeptide from polyomavirus, and one from human immunodeficiency virus; all of them interacting with 14-3-3 proteins. These partners include soluble and membrane proteins of eight different classes: protein kinases, phosphatases, receptors, G-proteins and related proteins, apoptosis-regulating proteins, adaptor proteins, transcription factors, and nuclear proteins and enzymes.[23] We used UniqueProt[24] (run locally) under the dataset, in order to eliminate any source of bias by overrepresented families. HSSP-distance value used was 20, and for all other parameters default values were used. After analysis, five proteins were eliminated.

To test the predictability of intrinsically disordered regions in 14-3-3-binding proteins, a randomized reference sequence dataset was created. Because predictions depend on sequence attribute, we calculated the percentage of each amino acid in the 14-3-3-binding protein dataset. Then, using an automatic random sequences generator (*RandSeq*; http://au.expasy.org/tools/randseq.html) and the specified amino acid composition in percent, we generated random sequences and collected them in the randomized reference sequence dataset.

Disorder Predictors

We utilized VL3H, PONDR®, and Disopred2 methods for predicting disorder in proteins. These predictors differ not just methodologically, but also conceptually because of different definitions of disorder. VL3H and PONDR® were trained to distinguish experimentally verified disorder in globular proteins by using various machine-learning approaches. For Disopred, the definition of disorder was restrained to regions missed from X-ray structures, which were specifically recognized by training a support vector machine.

Briefly, VL3H used the same 20 attributes, predictor models, and postfiltering as VL3.[25] The available dataset with 152 disordered proteins used for VL3 was fairly small and likely to constrain the achievable accuracy of disorder prediction. To solve this, we included homologs of the disordered sequences to enhance the dataset. PONDR® was formed by merging three neural network predictors of disorder: one for N-terminal regions, a second for internal regions, and a third for C-terminal regions. The merger was accomplished by performing overlapping predictions, followed by averaging the outputs. The PONDR® training set included disordered segments of 40 or more amino acid residues as characterized by X-ray and nuclear magnetic resonance for the predictor of the internal regions, and segments of five or more amino acid residues for the predictors of the two terminal regions.[26]

The training set for Disopred2 was the same as that used to train the original version of the program, and it was composed of nonredundant chains with X-ray structures in the Protein Data Bank and $<25\%$ pair-wise sequence identity. Only structures with resolutions better than 2.0 Å were used to avoid missing regions caused by

poor model quality. Disordered residues were identified after aligning the sequence of the protein chain in the SEQRES records with the sequence specified by the ATOM records (α-carbon coordinates). Residues that were found in the SEQRES records but not in the ATOM records were classified as disordered. The final training set comprised 715 protein chains, in which a total of 176,550 residues were classified as ordered and 4,590 residues as disordered.[27]

Secondary Structure Predictions

Three methods have been applied to predict the secondary structure of 14-3-3 partner proteins. The three algorithms are based on entirely different principles: GOR[28] is a simple statistical method, ALB[29] estimates short- and long-range interactions based on physicochemical properties of the amino acid residues, whereas PROF[30] is an efficient neural network algorithm that uses evolutionary information. These algorithms were used because their accuracies for intrinsically disordered proteins have already been evaluated, by comparison of predictions with DSSP assignments of experimental structures.[31] For a comparison of the results, evaluation was performed on three conformational states H (helix), E (extended), and irregular structure C (coil). Hence, the four-state prediction results of GOR and ALB are converted to three states by regarding T (turns) as coils. In ALB, the "B" and "S" state were defined as E (extended), whereas the highly probable "&" state was regarded as H (helix), as in reference 31.

RESULTS

Predicted Disorder Regions of 14-3-3-Binding Proteins

We took a collection of 46 partners that interact with 14-3-3 proteins (see reference 23) to analyze internal disorder, as well as to systematically study the intrinsic disorder tendencies and to test for an association with the binding process. Three different predictors were utilized in our study: VL3H, PONDR®, and Disopred2. VL3H and PONDR® predict disordered regions based solely on the amino acid sequences with accuracies of 85.3 ± 1.4% and 71.6 ± 1.3%, respectively.[25] VL3H has a better performance because homologs of the disordered proteins were used in the training stage. Disopred2 was trained directly on protein sequences with inputs derived from profiles generated by PSI-BLAST. When compared with the other algorithms, Disopred2 has an improved accuracy.[27] The dataset was composed of 14-3-3-binding proteins mainly derived from humans, plus two viral proteins. As detailed in the sequences and datasets section of Methods (see above), the collection included soluble and membrane proteins from eight different functional classes.[23]

Predictions are summarized in Table I. The regions predicted by the three algorithms are essentially the same; however, whereas VL3H predicts longer and continuous regions, PONDR® divided them into more smaller regions. Although PONDR® and Disopred2 predicted regions with the same average length (Table I), the number of disor-

TABLE I. Description of Disordered Regions Protein Dataset and Comparison With Regions Containing 14-3-3-Binding Sites

Analysis criteria	No. of proteins	Maximum length	Average length	Maximum disorder region length			Average disorder region length			No. of disordered regions		
				VL3H	PONDR®	Disopred2	VL3H	PONDR®	Disopred2	VL3H	PONDR®	Disopred2
All the disordered regions												
All proteins in database	46	1374	653	381	215	256	86	45	41	164	215	145
Soluble	32	1374	670	367	184	256	81	45	39	120	157	101
Membrane	14	1243	607	381	215	102	92	46	42	44	58	46
Disordered regions containing 14-3-3-binding sites												
All proteins in database				381	215	197	159	71	53			
Soluble				367	184	197	120	157	101			
Membrane				381	215	68	44	58	46			
Randomized reference sequence dataset				107	45	35	60	38	33	52	12	4

Disordered regions shorter than 30 amino acid residues were eliminated from the analysis. Lengths are expressed in residues.

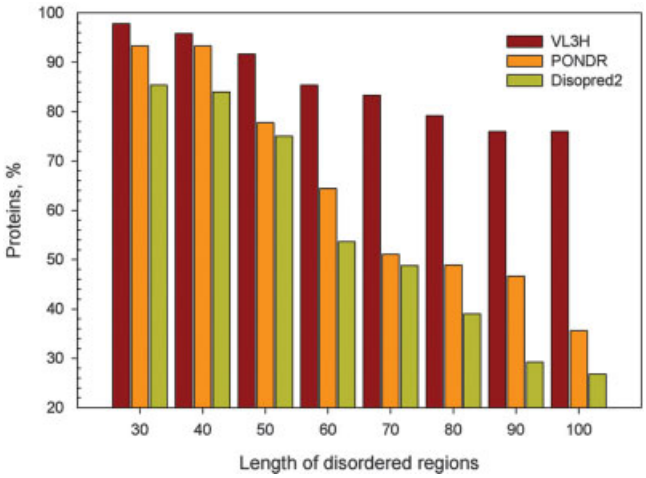

Fig. 1. Disorder prediction for 14-3-3-binding proteins in the dataset. Percentage of proteins with ≥30 to ≥100 consecutive residues (amino acids) predicted to be disordered.

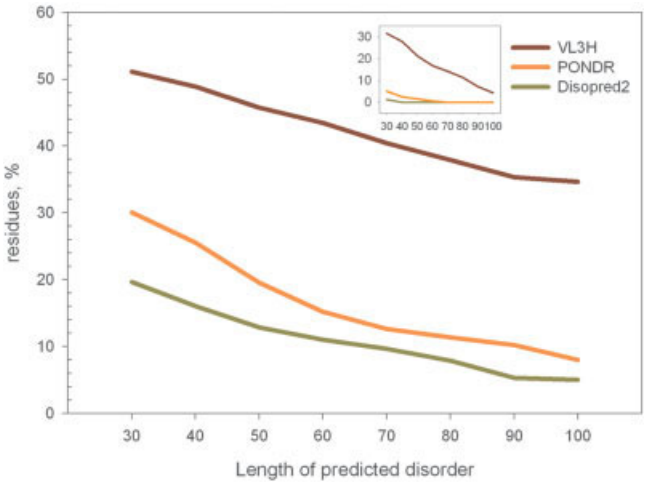

Fig. 2. Percentage of residues in the dataset predicted as disordered within segments of amino acids length of at least the value on the *x* axis. The main graph shows the results corresponding to 14-3-3-binding proteins, whereas, in the inset, a randomized reference sequence dataset was analyzed.

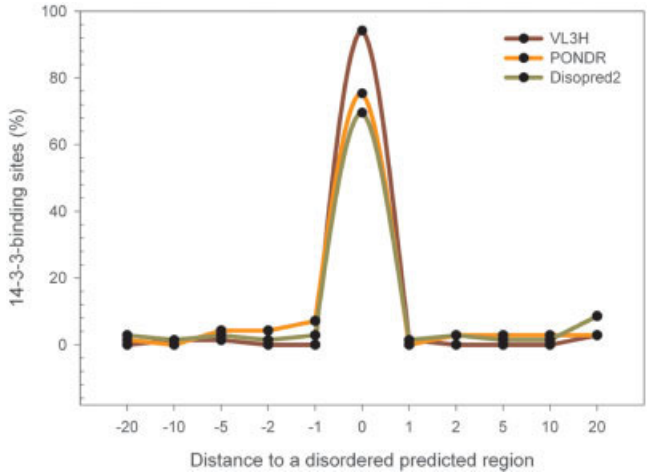

Fig. 3. Distribution of 14-3-3-binding sites within predicted disordered regions. Percentage of binding sites sorted by distance (amino acids, *x* axis) away from a disordered region. Negative or positive values mean that the binding site is not inside a disordered region and the nearest region is up or down stream, respectively. Zero in the *x* axis means that the binding site is (completely or partially) inside of a disordered region.

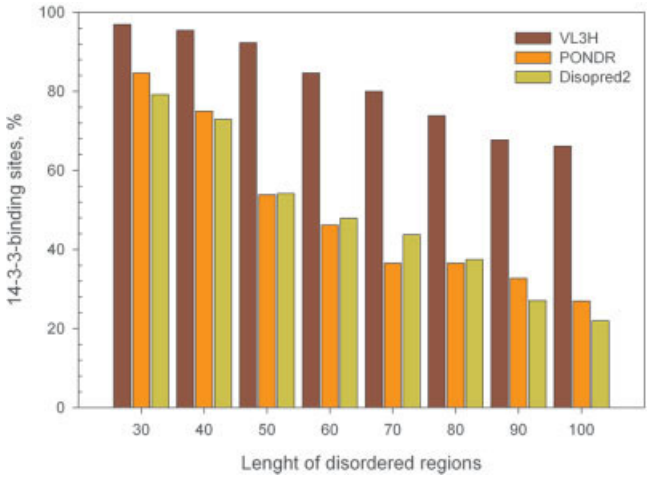

Fig. 4. Distribution of 14-3-3-binding sites along disordered regions sorted by amino acids length at least the value on the *x* axis.

dered regions obtained with Disopred2 was lower, essentially because it predicted many regions with <30 residues which were discarded in the analysis.

The percentage of 14-3-3-binding proteins containing long disordered regions (≥30 residues) was 97.8, 93.3, and 85.3% for VL3H, PONDR®, and Disopred2, respectively (Fig. 1). When the percentage of proteins with 60 (or more) consecutive disordered amino acids is compared, the differences between the three predictors are higher. Thus, according to VL3H, 85% of the proteins have disordered regions of at least 60 residues, but this percentage decreases to 64% for PONDR® and 54% for Disopred2 (Fig. 1).

The same ranking was observed when the results were presented as percentage of residues (Fig. 2). Based on prediction by VL3H, 50 or 35% of the residues belong to disordered regions of 30 or 100 amino acids, respectively. PONDR® and Disopred2 relatively predicted more residues in the short regions, with only 8 and 5% of the residues found in disordered regions longer than 100 amino acids (Fig. 2). Even with differences, all these predictive values are higher than those reported for globular (ordered) proteins,[26] especially when they are challenged with random sequences (inset Fig. 2). Thus, 14-3-3-binding proteins are richer in predicted disorder regions than typical eukaryotic proteins.[26,27,32]

Sites for 14-3-3 Binding as Disordered Regions

Almost all 14-3-3 proteins bind to phosphoserine/phosphothreonine-containing peptide motifs corresponding to the sequences RSXpSXP or RXXXpSXP. For the proteins selected in our study, their 14-3-3-binding sites

are well known after biochemical characterization. We utilized information in the subject extensively reviewed by Aitken and coworkers,[23] together with additional recent reports.[33–38] The most functionally significant observation in our analysis was that 94.2, 75.4, and 69.6% (for VL3H, PONDR®, and Disopred2, respectively) of the 14-3-3-binding sites were found to be within disordered regions (Fig. 3). The sequences that were experimentally demonstrated to bind to 14-3-3 proteins were predicted as disordered even for Disopred2, an algorithm that neither utilizes the biochemical nor the physicochemical properties of residues in its predictions. In VL3H and PONDR®, although arginine and serine residues are always positively correlated with disorder, a complex relationship of 38 different combinations of amino acids is used by these predictors.[39]

Table I also includes an analysis of disorder restricted to regions directly interacting with 14-3-3, showing an increase in the average length of them compared with all disordered regions in 14-3-3-binding proteins (in all of the proteins analyzed). The latter was not evident for membrane proteins, possibly because sites for 14-3-3 binding locate in regions connecting highly structured transmembrane domains. The average length of the disordered regions in soluble proteins increased almost 1.5-fold for VL3H (from 81 to 120 residues), 3.5-fold for PONDR® (from 45 to 157 residues), and 2.6-fold for Disopred2 (from 39 to 101 residues) (Table I). These data show a strong relationship between 14-3-3-binding sites and long disordered regions, thus reinforcing the possible biological role of disordered structure in the binding of 14-3-3 proteins to their partners.

Figure 4 illustrates the distribution of the sites for 14-3-3 binding all along different disordered regions classified by length. As expected, we found that VL3H predicted the higher percentages for the association between sites for the interaction with 14-3-3 and disordered structure, with values near 100 or 66% for region lengths up to 30 or 100 residues, respectively (Fig. 4). For the other two predictors, the distribution was quite uniform between them, as 85% (PONDR®) and 79% (Disopred2) of the 14-3-3-binding sites were found in regions with at least 30 amino acids. These percentage values decrease to about 50% for regions of up to 60 residues and to near 25% for region lengths of 100 residues (Fig. 4). These results correlate sequences with functional relevance in the binding of 14-3-3 proteins to their partners and putative long disordered regions. Even when the disordered regions have been implicated in protein–protein interactions[13] and phosphorylation,[40] this is the first time that a systematic approach has been used to evidence the relationship between these processes and 14-3-3 proteins.

Secondary Structure Propensity Into 14-3-3-Binding Site Containing Disordered Regions

To seek a more structure-oriented perspective, we investigated regular secondary structure (helix or extended) contents in the 14-3-3-binding proteins and regions by two different methods. Data were compared with those corresponding to regular secondary structure in both globular and intrinsically unfolded proteins.[31] The predictions are summarized in Table II. Algorithms ALB[29] and PROF[30] estimate similar percentages for helix and extended structures. It has been reported that accuracies of ALB and PROF methods are considerably higher than those of GOR[31] (a simple statistical method), and for this reason we discarded the latter in the analysis of secondary structures of disordered regions containing 14-3-3-binding sites.

TABLE II. Secondary Structures in Dataset†

Analysis criteria	Helix (%)	Extended (%)	Turns/coils (%)
14-3-3-Binding proteins			
ALB	24.80	11.53	63.67
PROF	25.57	11.16	63.27
Globular proteins[a]	35.3	21.90	42.80
Disordered regions containing 14-3-3-binding sites			
ALB	12.28	5.21	82.51
PROF	13.23	5.18	81.59

†Percentage of predicted secondary structures in 14-3-3-binding proteins and in disordered regions containing 14-3-3 binding sites.
[a]See reference [31]

Extended structures are almost equally settled in 14-3-3-binding proteins and intrinsically disordered proteins (Table II), but the percentage of helices predicted by ALB or PROF are significantly lower, indicating that 14-3-3-binding proteins do not have a similar probability to form 1 to 4 hydrogen bonds as intrinsically disordered or globular proteins (Table II). However, when the percentage of secondary structure in disordered regions containing 14-3-3-binding sites is considered, values for helices or extended conformations are 50% lower than those found for the entire protein (Table II). This suggests that 14-3-3 targets act as modular proteins, with low-complexity domains to bind the regulatory protein and other different domains to exert their physiological functions.

DISCUSSION

We compared the results obtained by using three methods for predicting disorder, which differ not just methodologically but also conceptually because of different definitions of disorder. PONDR® and Disopred2 predictors were used for the analysis of a high number of protein datasets. Iakoucheva et al.[26] used PONDR® to find the intrinsic disorder in cancer-associated and signaling proteins, and then compared their data with results for ordered proteins. They found that 85% of regulatory proteins and 79% of cancer-related proteins have at least 30 consecutive disordered residues.[26] In the same way, Tompa and Csermely[32] predicted that approximately 81% of a collection of 27 RNA chaperones have long internal disordered regions. Disopred2 was used to estimate the frequency of native disorder in several representative genomes from the three life kingdoms. Putative, long (≥30 residues)

TABLE III. Experimental Structural Information for 14-3-3-Binding Proteins and Comparison With Predicted Regions

Protein	Detection method	Disordered region (from–to)	Function	Predicted disordered region (from–to)			14-3-3-Binding site (from–to)
				VL3H	PONDR®	Disopred2	
p53	Circular dichroism	1–70	Protein interaction	1–103	1–18		375 QSTpSRH 380
		71–98	Unknown		34–99		
				164–178	152–165		
		285–324	Unknown	266–393	280–317	281–326	
		357–366	Unknown		326–329	353–393	
		367–388	Unknown		342–393		
		389–393	Unknown				
pAANAT	Cristallography		Protein interaction	1–45	1–28	20–30	28 RRHpTLP 33
			Protein interaction	196–207	200–207	205–207	202 RRpSGC 207

disordered segments were found to occur in 2.0% of archaean, 4.2% of eubacterial, and 33.0% of eukaryotic proteins.[27] The current scenario illustrates that in the eukaryotic proteome, 1 of 3 proteins has disordered regions longer than 30 residues.

Remarkably, results shown herein indicate that a high number of partners for the binding of 14-3-3 proteins exhibit long disordered regions in significantly higher proportion than any other protein class tested so far. These results were coherent independently of the predictor, which support a good confidence for the data (a comparison with a random sequence dataset can be found in Table I). Predictions of PONDR® can be compared directly with works of Iakoucheva et al.[26] and Tompa and Csermely,[32] where regulatory proteins, cancer-related proteins, and protein and RNA-chaperones are analyzed. Of the predictors we used, VL3H is the less extensively utilized at the present time, and thus its accuracy is hard to estimate. However, VL3H is an improved version of PONDR® and this supposes that the result obtained in our analysis is confirmatory of the high degree of disordered regions for partners interacting with 14-3-3 proteins.

From the modular nature of proteins it is logical to think that the large number of partners for 14-3-3 proteins could be grouped into structural subsets. However, a casual inspection of 14-3-3-binding regions reveals no extensive sequence similarities to support that notion. Thus, a direct proteomic and domain-based analysis of in vivo 14-3-3γ-binding partners in human embryonic kidney cells revealed that they lack a common domain to bind to the regulatory proteins.[8] The presence of disordered domains in such partners could be a plausible explanation for these results, which in addition retains the idea of a common mechanism for binding.

Most conspicuous, and perhaps functionally the most significant in results reported herein, is the occurrence of almost all of the 14-3-3-binding sites within disordered regions (Fig. 3). The average length of these regions was ~50% higher than the average for random sequences (Table I), suggesting that the longer, the more accurately predicted, and perhaps the more functional of the regions contain sites recognized by 14-3-3 proteins. These remarkable numbers underline the particular importance of structural disorder in this protein class.

The occurrence of disorder in partners that bind 14-3-3 proteins is strongly supported by the analysis of two proteins which structures have been characterized at the three-dimensional level (Table III). The p53 is a 14-3-3-binding protein[41] that was experimentally determined as disordered.[34,42] PONDR® predictions of p53 were analyzed previously (see reference 26). The three programs used in the present study predicted that the 14-3-3-binding site (residues 375–380) was disordered, and a disorder-to-order transition upon oligomerization and binding with the respective partner was experimentally determined.[26]

The shortened version of serotonin *N*-acetyltransferase ($pAANAT_{1–201}$) is the only partner of 14-3-3 proteins where the complex structure was determined by crystallography.[43] This enzyme has two binding sites (^{28}RRHpTLP^{33} and ^{204}RRNpSGC^{209}-COOH), both predicted as disordered in this work (Table III). Analysis of crystals suggests that actually the two 14-3-3-binding sites in pAANAT are disordered[43] and this characteristic makes both sites functional, as was recently confirmed.[10] Thus, these data derived from analysis of crystal structures completely agree with the main conclusion in the present study with respect to the association of disordered structure and interaction between 14-3-3 proteins and their partners.

Solved three-dimensional structures of some intrinsically disordered proteins have shown that in the binding to their respective partners they adopt an extended conformation.[14] The crystal structure of the 14-3-3ζ:$pAANAT_{1–201}$ complex shows that each monomer in the 14-3-3 dimer binds one molecule of $pAANAT_{1–201}$. Residues of this molecule lie in an extended conformation in the amphipathic groove for binding of 14-3-3ζ with the phosphate group of pThr-31 pointing into the positively charged depression in the middle of the groove. This is consistent with observations made with phosphopeptides, that determined the same extended conformation.[9,43,44] However, the region of AANAT described by Obsil et al.[43] as floppy loop 1, is a floppy element of the arylalkylamine-binding pocket of the protein that changes its conformation upon dual-site binding of AANAT to 14-3-3. The latter produces an optimal configuration of the loop for the binding of substrate with high affinity.[10] Although changes on this loop resemble the passage of a disordered region to an ordered one, the characteristics of loop 1 do not fit strictly

to the definition of intrinsically disordered regions. As is predicted in DisEMBL (http://dis.embl.de), loop 1 could be defined as loop/coils better than disordered region.

The model of molecular recognition involving disordered regions can explain binding between macromolecules occurring with high specificity, low affinity, and under thermodynamic control.[13] One example in such a way is the binding of phosphoCREB to the KIX domain of CBP (see Introduction in this work), which entropically disfavored ($-T\Delta S = 1.78$ kcal mol^{-1}) characteristics are consistent with the disorder-to-order transition (coil to helix folding) that must escort the process. Formation of the phosphoCREB-CBP complex is driven by a large enthalpy change ($\Delta H = -10.6$ kcal mol^{-1}), presumably associated with favorable hydrogen bonding interactions established by the phosphoryl group attached to a serine residue.[13,18] Thermodynamic parameters were determined for the ternary complex between the fungal phytotoxin fusicoccin, the C-terminus of the plant plasma membrane H^+-ATPase, and 14-3-3c.[45] The peptide binding to 14-3-3 is an enthalpy-driven ($\Delta H = -10$ kcal mol^{-1}) and entropically unfavorable ($-T\Delta S = 1.7$ kcal mol^{-1}) process. The fusicoccin increases the binding affinity, after decreasing $-T\Delta S$ to 0.4 kcal mol^{-1} and increasing the enthalpy change to -11.8 kcal mol^{-1}.[45] These values are in accordance with those previously discussed for phosphoCREB and KIX, where a transition disorder-to-order and the binding of a phosphoserine residue take place.

Specificity and affinity are concepts requiring a special analysis in the interaction between 14-3-3 proteins and their respective partners. It is clear that 14-3-3 proteins are highly specific, because the vast majority of 14-3-3 targets bind or increase their affinity for binding, in a phosphorylation-dependent manner.[6,11] This means that 14-3-3 proteins can "recognize" the phosphoryl group on a serine residue; it is the existence of anchor residues in 14-3-3 binding partners that allows a basal binding, as subjected by Rajamani et al.[46]

Three cases are known in which neither a phosphoryl group nor a consensus sequence is necessary. Glycoprotein Ibα,[47] ADP-ribosyltransferase toxin exoenzyme S of *Pseudomonas aeruginosa*,[48] and Tau protein[36] from brain can bind 14-3-3 proteins independently of phosphorylation. Whereas the sequence responsible to interact with 14-3-3 is unknown in Tau protein, in exoenzyme S, the binding-sequence of two unphosphorylated peptides (one derived from exoenzyme S and another obtained by phage display) inhibiting the interaction with 14-3-3 protein were objects of extensive studies.[49,50] These unphosphorylated peptides bind 14-3-3 in the same binding groove and with the same affinity ($K_D \sim 90$ nM) that the phosphorylated targets.[49,50] This apparent counterintuitive feature in the binding of 14-3-3 proteins could be explained by predictions of disorder and secondary structures of exoenzyme S of *P. aeruginosa*. The residues ^{424}DALDL^{428} (14-3-3-binding site) are within an α-helix in the ordered region of this protein (data not shown), opening the possibility that unphosphorylated targets bind 14-3-3 proteins in the same way that unphosphorylated c-Myb binds KIX.[13,18]

Based on results showed herein and the above details, a mechanism for the binding interaction between 14-3-3 proteins and their partners could be proposed. A disorder-to-order transition of partners or, in other words, an enthalpy-driven binding between 14-3-3 proteins and their protein targets, takes place by the hydrogen bonding interactions made by the phosphoryl group of phosphoserine or phosphothreonine residue ubiquitously present in the 14-3-3-binding sites. This could explain why 14-3-3 proteins are not able to bind targets with canonical sequences in an unphosphorylated manner. Such disorder-to-order hypothesis has a direct mechanistic relationship with the sequence of canonical binding sites (and their variations) found in the 14-3-3-binding partners. For example, the arginine or serine residues at positions -3 or -2 (from phosphoserine residue), respectively, could have only a stabilizing effect, but they are not essential. This means that others that confer similar properties could replace these residues. In this sense, the arginine residue (the most frequent amino acid in position -3) is naturally replaced by lysine (insulin receptor substrate 1 IRS-1), glutamine (breakpoint cluster region protein Bcr), glutamic acid (regulator of G-protein signaling RGS3), serine (insulin growth factor I receptor), or histidine (Ca^{+2}/calmodulin-dep myosin l-chain kinase, skeletal, b-chain of GM-CSF, interleukin-3 and -5 receptors), as main examples. This latter analysis explains the occurrence of "imperfect" sites for the binding of 14-3-3, as proposed by Bridges and Moorhead.[11] Importantly, the functioning of such sites are included in a more general mechanism in our model, having disorder-to-order transitions as a key issue.

In addition to proposing this mechanistic model, our finding could be the basis for a practical application associated with genomics and proteomics. In this context, the use in identification of disordered regions and sequence partners would be useful to analyze recent proteomic approaches,[6–8] and wide-genomic search of 14-3-3-binding partners in different organisms.

ACKNOWLEDGMENTS

Access to PONDR® was provided by Molecular Kinetics (6201 La Pas Trail - Ste 160, Indianapolis, IN 46268; 317-280-8737; E-mail: main@molecularkinetics.com). PONDR® is ©2004 by Molecular Kinetics, all rights reserved. Work in the authors' laboratory was granted by ANPCyT (PICTO 1-13241, PICT'03 1-14733, and PAV 137). The authors are members of the Investigator Career from CONICET (Argentina).

REFERENCES

1. Moore BW, Perez VJ. Specific acid proteins in the nervous system. Cliffs E, editor. New Jersey: Prentice-Hall; 1967.
2. Dougherty MK, Morrison DK. Unlocking the code of 14-3-3. J Cell Sci 2004;117:1875–1884.
3. Yaffe MB. How do 14-3-3 proteins work? Gatekeeper phosphorylation and the molecular anvil hypothesis. FEBS Lett 2002;513:53–57.
4. Mackintosh C. Dynamic interactions between 14-3-3 proteins and phosphoproteins regulate diverse cellular processes. Biochem J 2004;381(Pt 2):329–342.

5. Muslin AJ, Tanner JW, Allen PM, Shaw AS. Interaction of 14-3-3 with signaling proteins is mediated by the recognition of phosphoserine. Cell 1996;84(6):889–897.
6. Pozuelo Rubio M, Geraghty KM, Wong BHC, et al. 14-3-3-affinity purification of over 200 human phosphoproteins reveals new links to regulation of cellular metabolism, proliferation and trafficking. Biochem J 2004;273:395–408.
7. Meek SE, Lane WS, Piwnica-Worms H. Comprehensive proteomic analysis of interphase and mitotic 14-3-3-binding proteins. J Biol Chem 2004;279:32046–32054.
8. Jin J, Smith FD, Stark C, 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.
9. Yaffe MB, Rittinger K, Volinia S, et al. The structural basis for 14-3-3: phosphopeptide binding specificity. Cell 1997;91(7):961–971.
10. Ganguly S, Weller JL, Ho A, Chemineau P, Malpaux B, Klein DC. Melatonin synthesis: 14-3-3-dependent activation and inhibition of arylalkylamine N-acetyltransferase mediated by phosphoserine-205. Proc Natl Acad Sci USA 2005;102(4):1222–1227.
11. Bridges D, Moorhead GB. 14-3-3 proteins: a number of functions for a numbered protein. Sci STKE 2004;2004(242):re10.
12. Uversky VN. Natively unfolded proteins: a point where biology waits for physics. Protein Sci 2002;11(4):739–756.
13. Wright PE, Dyson HJ. Intrinsically unstructured proteins: reassessing the Protein Structure-Function Paradigm. J Mol Biol 1999;293:321–331.
14. Tompa P. The functional benefits of protein disorder. J Mol Struct (Theochem) 2003;666–667:361–371.
15. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z. Intrinsic disorder and protein function. Biochemistry 2002;41(21):6573–6582.
16. Dunker AK, Lawson JD, Brown CJ, et al. Intrinsically disordered protein. J Mol Graph Model 2001;19(1):26–59.
17. Dunker AK, Obradovic Z. The protein trinity-linking function and disorder. Nat Biotechnol 2001;19(9):805–806.
18. Parker D, Rivera M, Zor T, et al. Role of secondary structure in discrimination between constitutive and inducible activators. Mol Cell Biol 1999;19(8):5601–5607.
19. Daughdrill GW, Chadsey MS, Karlinsey JE, Hughes KT, Dahlquist FW. The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, sigma 28. Nat Struct Biol 1997;4(4):285–291.
20. Hauer JA, Taylor SS, Johnson DA. Binding-dependent disorder-order transition in PKI alpha: a fluorescence anisotropy study. Biochemistry 1999;38(21):6774–6780.
21. Fletcher CM, Wagner G. The interaction of eIF4E with 4E-BP1 is an induced fit to a completely disordered protein. Protein Sci 1998;7(7):1639–1642.
22. Wallon G, Rappsilber J, Mann M, Serrano L. Model for stathmin/OP18 binding to tubulin. EMBO J 2000;19(2):213–222.
23. Aitken A, Baxter H, Dubois T, et al. 14-3-3 proteins in cell regulation. Specificity of 14-3-3 isoform dimer interactions and phosphorylation. Biochem Soc Trans 2002;30(4):351–360.
24. Mika S, Rost B. UniqueProt: creating representative protein sequence sets. Nucleic Acids Res 2003;31(13):3789–3791.
25. 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 2004:1–23.
26. 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.
27. 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;337:635–645.
28. Garnier J, Osguthorpe DJ, Robson B. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J Mol Biol 1978;120(1):97–120.
29. Ptitsyn OB, Finkelstein AV. Theory of protein secondary structure and algorithm of its prediction. Biopolymers 1983;22(1):15–25.
30. Rost B, Sander C. Improved prediction of protein secondary structure by use of sequence profiles and neural networks. Proc Natl Acad Sci USA 1993;90(16):7558–7562.
31. Fuxreiter M, Simon I, Friedrich P, Tompa P. Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. J Mol Biol 2004;338(5):1015–1026.
32. Tompa P, Csermely P. The role of structural disorder in the function of RNA and protein chaperones. FASEB J 2004;18(11):1169–1175.
33. Pozuelo Rubio M, Peggie M, Wong BH, Morrice N, MacKintosh C. 14-3-3s regulate fructose-2,6-bisphosphate levels by binding to PKB-phosphorylated cardiac fructose-2,6-bisphosphate kinase/phosphatase. EMBO J 2003;22(14):3514–3523.
34. Chang J, Kim DH, Lee SW, Choi KY, Sung YC. Transactivation ability of p53 transcriptional activation domain is directly related to the binding affinity to TATA-binding protein. J Biol Chem 1995;270(42):25014–25019.
35. Kino T, Chrousos GP. Human immunodeficiency virus type-1 accessory protein Vpr: a causative agent of the AIDS-related insulin resistance/lipodystrophy syndrome? Ann NY Acad Sci 2004;1024:153–167.
36. Hashiguchi M, Sobue K, Paudel HK. 14-3-3ζ is an effector of tau protein phosphorylation. J Biol Chem 2000;275(33):25247–25254.
37. Morris MC, Heitz A, Mery J, Heitz F, Divita G. An essential phosphorylation-site domain of human cdc25C interacts with both 14-3-3 and cyclins. J Biol Chem 2000;275(37):28849–28857.
38. Johnson BA, Stehn JR, Yaffe MB, Blackwell TK. Cytoplasmic localization of tristetraprolin involves 14-3-3-dependent and -independent mechanisms. J Biol Chem 2002;277(20):18029–18036.
39. Xie Q, Arnold GE, Romero P, Obradovic Z, Garner E, Dunker AK. The sequence attribute method for determining relationships between sequence and protein disorder. Genome Inform Ser Workshop Genome Inform 1998;9:193–200.
40. Iakoucheva LM, Radivojac P, Brown CJ, et al. The importance of intrinsic disorder for protein phosphorylation. Nucleic Acids Res 2004;32(3):1037–1049.
41. Waterman MJ, Stavridi ES, Waterman JL, Halazonetis TD. ATM-dependent activation of p53 involves dephosphorylation and association with 14-3-3 proteins. Nat Genet 1998;19(2):175–178.
42. Kussie PH, Gorina S, Marechal V, et al. Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. Science 1996;274(5289):948–953.
43. 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.
44. Rittinger K, Budman J, Xu J, et al. 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.
45. Würtele M, Jelich-Ottmann C, Wittinghofer A, Oecking C. Structural view of a fungal toxin acting on a 14-3-3 regulatory complex. EMBO J 2003;22(5):987–994.
46. Rajamani D, Thiel S, Vajda S, Camacho CJ. Anchor residues in protein–protein interactions. Proc Natl Acad Sci USA 2004;101:11287–11292.
47. Gu M, Du X. A novel ligand-binding site in the zeta-form 14-3-3 protein recognizing the platelet glycoprotein Ibalpha and distinct from the c-Raf-binding site. J Biol Chem 1998;273(50):33465–33471.
48. Masters SC, Pederson KJ, Zhang L, Barbieri JT, Fu H. Interaction of 14-3-3 with a nonphosphorylated protein ligand, exoenzyme S of *Pseudomonas aeruginosa*. Biochemistry 1999;38(16):5216–5221.
49. Wang B, Yang H, Liu YC, et al. Isolation of high-affinity peptide antagonists of 14-3-3 proteins by phage display. Biochemistry 1999;38(38):12499–12504.
50. Petosa C, Masters SC, Bankston LA, et al. 14-3-3ζ binds a phosphorylated Raf peptide and an unphosphorylated peptide via its conserved amphipathic groove. J Biol Chem 1998;273(26):16305–16310.