





Article

Functional Analysis of Human Hub Proteins and Their Interactors Involved in the Intrinsic Disorder-Enriched Interactions

Gang Hu ^{1,†}, Zhonghua Wu ^{1,†}, Vladimir N. Uversky ^{2,3,*}  and Lukasz Kurgan ^{4,*} 

¹ School of Mathematical Sciences and LPMC, Nankai University, 300071 Tianjin, China; huggs@nankai.edu.cn (G.H.); wuzhh@nankai.edu.cn (Z.W.)

² Department of Molecular Medicine and Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA

³ Laboratory of New Methods in Biology, Institute for Biological Instrumentation of the Russian Academy of Sciences, Pushchino, Moscow region 142290, Russia

⁴ Department of Computer Science, Virginia Commonwealth University, Richmond, VA 23284, USA

* Correspondence: vuversky@health.usf.edu (V.N.U.); lkurgan@vcu.edu (L.K.)

† These authors contributed equally to this work.

Received: 1 November 2017; Accepted: 15 December 2017; Published: 19 December 2017

Abstract: Some of the intrinsically disordered proteins and protein regions are promiscuous interactors that are involved in one-to-many and many-to-one binding. Several studies have analyzed enrichment of intrinsic disorder among the promiscuous hub proteins. We extended these works by providing a detailed functional characterization of the disorder-enriched hub protein-protein interactions (PPIs), including both hubs and their interactors, and by analyzing their enrichment among disease-associated proteins. We focused on the human interactome, given its high degree of completeness and relevance to the analysis of the disease-linked proteins. We quantified and investigated numerous functional and structural characteristics of the disorder-enriched hub PPIs, including protein binding, structural stability, evolutionary conservation, several categories of functional sites, and presence of over twenty types of posttranslational modifications (PTMs). We showed that the disorder-enriched hub PPIs have a significantly enlarged number of disordered protein binding regions and long intrinsically disordered regions. They also include high numbers of targeting, catalytic, and many types of PTM sites. We empirically demonstrated that these hub PPIs are significantly enriched among 11 out of 18 considered classes of human diseases that are associated with at least 100 human proteins. Finally, we also illustrated how over a dozen specific human hubs utilize intrinsic disorder for their promiscuous PPIs.

Keywords: intrinsic disorder; protein-protein interactions; hub proteins; intrinsically disordered proteins; human proteome

1. Introduction

Although proteins are engaged in virtually all aspects of cellular functionality, they rarely act alone, being normally involved in multiple interactions (i.e., specific physical contacts that have functional consequences) with other proteins, peptides, nucleic acids, polysaccharides, various small molecules, lipids, etc. This is reflected in the fact that proteins in a cell are traditionally described in a form of specific pathways that represent a series of actions among molecules leading to a particular change in a cell, e.g., generation of a certain product. Therefore, mapping of protein-protein interactions (PPIs) represents an important step towards the better understanding of the complex molecular relationships in living systems [1]. In fact, the resulting intricate PPI networks, or interactomes, are considered now as a gateway into systems biology, since they represent an important means to understand how

biochemical and biophysical characteristics of proteins can be integrated into biological systems [2]. Interactomes are different from canonical pathways, being typically more complex and providing massive complementary information that reveals a multitude of links (or interactions) of each of the proteins included in a given network.

Interactomes have complex topologies, where proteins are depicted as nodes and PPIs are shown as links or connections between these nodes. Complex networks found in many diverse systems are typically assumed to be random [3]. Given a random placement of links in the random networks, most nodes have approximately the same number of links [3,4]. In other words, in a completely random network, the connectivity of nodes (also known as degree) follows a Poisson distribution with a bell shape. This means that it is difficult to find nodes that have significantly more or fewer links than the network average values [4]. On the other hand, PPI networks inside cells are classified as “scale-free” or “small-world” networks, as systematized in [5]. Other examples of such networks include the author-collaboration networks, the airline routes, the World Wide Web, the metabolic networks, and the protein domain networks. Degree distribution in these scale-free networks follows a power law [6], suggesting the presence of nodes with a degree that greatly exceeds the network average value. In other words, although in scale-free networks most nodes have just a few links, there are some popular nodes (hubs) with a very large number of connections to other nodes (they can have hundreds, thousands or even millions of links), as well as there are some nodes (ends) connected only to one other node [4,5,7]. Therefore, the geometry of scale-free networks combines the local clustering of connections characteristic of regular networks with occasional long-range connections between clusters, as can be expected to occur in random networks [7,8].

While there is no established consensus in the literature on how to define hub proteins, many studies define them as proteins with larger degrees in the PPI networks, when compared to the non-hub proteins. For instance, one of the early studies by Han et al. [9] defined hubs as proteins with degrees >5 , Haynes et al. [10] decided to increase the cut-off to over 10, while Batada et al. [11] defined hubs as the top 5% of proteins with the highest degrees. In more recent studies, Aragues et al. [12] specified hubs as proteins with degrees >20 , and Jin et al. [13] as the top 20% of proteins with the highest degrees. With respect to timing of PPIs in networks, some hubs have multiple, simultaneous interactions (“party hubs”), forming scaffolds that enable the assembly of functional modules [9], whereas others hubs have multiple sequential interactions (“date hubs”) [9] and are likely to play a role in connecting biological modules to each other [14]. Irrespective of all these issues and considerations, hubs are invariably characterized by an intrinsic ability to be engaged in multiple interactions with numerous (often unrelated) partners. This ability to utilize the “one-to-many” interaction mode moves hubs from the comfort of the traditional lock-and-key model, if each given PPI would be considered as one function. Such binding promiscuity also raises an important question on structural determinants of hubness.

The “one protein-one structure-one function” paradigm has morphed in recent years, as it is becoming progressively more evident that many proteins are biologically active despite the lack of a unique 3D-structure either entirely or in their significant parts [15–21]. These intrinsically disordered proteins (IDPs) and intrinsically disordered protein regions (IDPRs) represent a unique and important extension of the protein kingdom, being abundantly present in any given proteome, as evidenced by several systematic bioinformatics analyses [22–26]. Functions of IDPs/IDPRs are complementary to functions of ordered proteins and domains [18,27–29]. Furthermore, dysfunction and dysregulation of many IDPs and hybrid proteins (i.e., that have ordered domains and long IDPRs) results in protein misfolding and aggregation, loss of normal function, and gain of toxic function [30–32], which are associated with the development of various diseases [31,33–36].

Among the unique disorder-specific functional features of IDPs and IDPRs is their exceptional binding promiscuity, where one protein or region is able to bind to multiple partners [37]. Obviously, the classical molecular recognition mechanisms cannot explain the ability of IDPs/IDPRs to bind to multiple partners [38]. In fact, neither the lock-and-key [39] nor original induced-fit [40] mechanisms

that were developed to describe recognition behavior of ordered proteins give a reasonable description of this multi-binding capability. Furthermore, some IDPs/IDPRs were shown to adopt different structures upon binding to different partners [15,41–46], thereby playing a number of crucial roles in mediating PPIs [10,15,37,41–59].

Based on these observations and on the fact that IDPs/IDPRs are commonly involved in one-to-many and in many-to-one binding [15,60], it was proposed that molecular recognition via disorder-to-order transitions upon binding would be a reasonable mechanism of binding by hub proteins [61]. The overall importance of intrinsic disorder for functions of hub proteins was analyzed in several recent bioinformatics studies [10,47–50,59]. Although disorder is more clearly associated with date hubs than with party hubs [49,59], some protein complexes clearly use long IDPRs as scaffolds to assemble groups of interacting proteins [51–58,62].

Although several previous studies were specifically dedicated to the analysis of the enrichment of disorder in hubs [10,47–49,59], the functional peculiarities of PPIs involving hubs (hub PPIs) remains mostly unexplored. The goal of this study is to fill this gap and to provide detailed functional characterization of hub PPIs. We are the first to analyze abundance and functions of disorder for hub PPIs for both hubs and their interactors. We also investigate whether disease associated proteins are abundant among the proteins involved in the disorder-enriched hub PPIs. We focus our analysis on the human proteome given the relatively high degree of completeness of its PPI network [63] and relevance to the analysis of the disease-associated proteins.

2. Results and Discussion

2.1. Intrinsic Disorder in Hub Protein-Protein Interactions

Our aim was to annotate and functionally/structurally characterize human proteins involved in the hub PPIs that have higher than expected disorder content. To annotate these proteins, we compared distributions of the disorder content in the human hubs and their interactors between true and random human PPI networks. We utilized the random network to quantify the expected levels of disorder content. We defined the random network by generating PPIs between randomly chosen pairs of human proteins such that the number of interactions and the node densities are the same as in the true PPI network. Figure 1 compares the distribution of the PPIs in the two-dimensional space defined by the disorder content of the hubs (x -axis) and their interactors (y -axis), where the coloring represents a relative ratio between the density of the true and randomized interactions in the PPI networks. Using the color scale defined in Figure 1, the orange (light blue) corresponds to PPIs which are 0.5 times more (less) frequent in the true PPI network compared to the random network. The green isolines define the density of PPIs in that space.

The large rectangular region colored from orange to red in the top right quadrant of Figure 1 reveals a collection of about 3% of PPIs that have substantially higher than expected density in the two-dimensional intrinsic disorder space. The relative ratio of density is >0.5 in that region, which means that the density of the true PPIs that have the corresponding levels of intrinsic disorder in the hub proteins and its interactors is at least 50% higher than expected. More precisely, among the 141,346 PPIs that involve hubs (which we call “hub PPIs”), we found 4071 hub PPIs that have density enriched by at least 50% vs. 137,275 of the remaining hub PPIs (enrichment $<50\%$). We note that the 141,346 hub PPIs are a subset of all 151,461 PPIs since they exclude interactions between non-hub proteins. There are 481 hubs and 1580 interactors that are involved in the 4071 hub PPIs. In agreement with the prior works that suggest that hubs have higher than expected amounts of the intrinsic disorder, the proteins that have substantially higher than expected density in the two-dimensional disorder space are characterized by large amounts of disorder, between 50% and 90%. Our analysis suggests that only a relatively small fraction of all hubs (481 out of 2866 = 16.8%) are in this set, while the remaining majority of hubs participates in hub PPIs that do not have higher than expected density; the latter set of hubs corresponds to hub PPIs that are located in the regions colored blue, white, and light yellow in

Figure 1. Interestingly, the same is also true for the interactors of hubs. They follow a similar pattern of inclusion of only a relatively small fraction of proteins (1580 out of 13,963 = 11.3%) with substantially higher than expected density in the two-dimensional intrinsic disorder space. Like in the case of the hubs, these interactors are also characterized by large values of the disorder content. Moreover, our analysis indicates that there are fewer than expected hub PPIs that involve fully disordered proteins. These are the dark blue colored regions on the far right and at the top of Figure 1.

To sum up, only a relatively small portion of proteins involved in the hub PPIs in human, including 17% of hubs and 11% of their interactors, are enriched in the intrinsic disorder. This enrichment is in line with an overall increase in the disorder content among the hubs that was also reported in prior studies [10,47,48,59,61]. The proteins that participate in the high-density disorder-enriched hub PPIs harbor substantial amounts of disorder (over 50% of their residues are disordered) but they include fewer than expected fully disordered proteins.

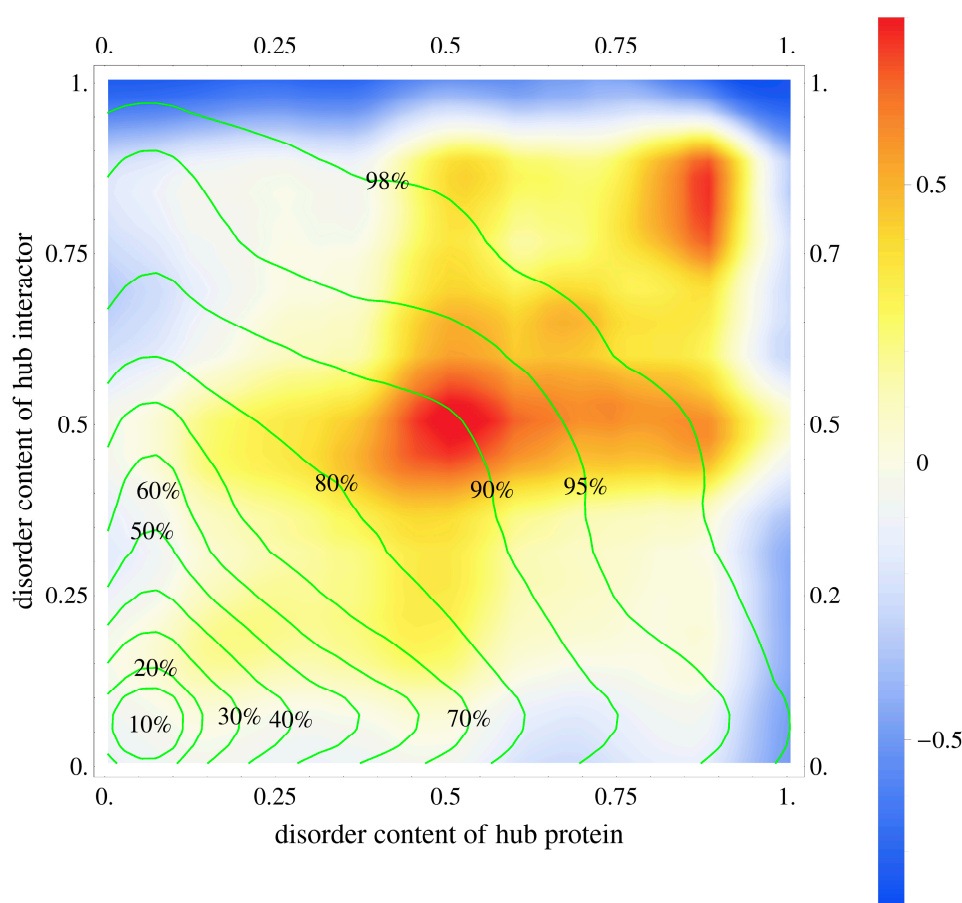


Figure 1. Enrichment in intrinsic disorder of human hub proteins and their interactors. The x - and y -axis show the amount of disorder content of the hubs and hub interactors, respectively. Each protein-protein interaction (PPI) is mapped into this two-dimensional plane and the density of these hub-interactor pairs is represented by green isolines. For instance, 40% of these pairs occupy the lower left corner where the disorder content of both hubs and interactors is below 0.25. The density was modelled with the Epanechnikov kernel function using Mathematica software. Next, we simulated a randomized PPI network that follows the same distribution of node density, i.e., we randomly assigned interactions between the human proteins to maintain the same density profile as in the true PPI network. Coloring of the inside of the two-dimensional plane reflects a relative ratio between the density of true (d_n) and randomized (d_r) interactions in the PPI networks calculated as $[d_n(x,y)-d_r(x,y)]/d_r(x,y)$. The color scale given on the right defines values of the ratio, e.g., orange corresponds to PPIs which are 0.5 times more frequent in the true PPI network compared to the random network.

Table 1. Analysis of enrichment in functional and structural characteristics of proteins. We compare hubs/hub interactors/hub interactors that exclude hubs that are associated with the disorder-enriched hub protein-protein interactions (PPIs) that have disorder content higher than expected by over 50% vs. the remaining hubs/hub interactors/hub interactors that exclude hubs. To ensure that results are statistically robust to represent diverse subpopulations of these proteins, we select 20% of proteins at random from each of the two protein sets, quantify a given characteristic for each set, and repeat this ten times. We report the average of the ten repetitions and the relative difference between averages for the two protein sets. We also evaluate the significance of the differences between these measurements. We use the t-test if the data are normal (we test normality with the Anderson-Darling test at p -value = 0.05); otherwise we use the Wilcoxon test. Bold font indicates large differences that are statistically significant (p -value < 0.001 and |relative difference| \geq 20%).

Type of Functional/Structural Characteristic ¹	Measure ¹	Results for Hubs Involved in Hub PPIs				Results for all Hub Interactors Involved in Hub PPIs				Results for Hub Interactors Involved in Hub PPIs that Exclude Hubs			
		Avg for Disorder-Enriched Proteins	Avg for Remaining Proteins	p -Value	Relative Difference	Avg for Disorder-Enriched Proteins	Avg for Remaining Proteins	p -Value	Relative Difference	Avg for Disorder-Enriched Proteins	Avg for Remaining Proteins	p -Value	Relative Difference
Structural properties	Disorder content	0.542	0.224	<0.001	141%	0.540	0.227	<0.001	138%	0.527	0.213	<0.001	148%
	Number of LIDPRs (per 1000 AAs)	4.428	1.900	<0.001	133%	4.436	1.910	<0.001	132%	4.863	1.954	<0.001	149%
	Number of disulfide links (per 1000 AAs)	0.765	1.240	<0.001	−62%	0.764	1.249	<0.001	−63%	0.973	1.828	<0.001	−47%
Sequence	Average protein chain length	736.8	637.6	<0.001	16%	734.5	653.2	<0.001	12%	791.5	654.2	<0.001	21%
Evolutionary conservation	Average conservation per protein	0.920	1.069	<0.001	−16%	0.930	1.053	<0.001	−13%	0.965	1.048	<0.001	−8%
	Average conservation of LIDPRs	0.735	0.742	0.113	−1%	0.762	0.763	0.811	0%	0.790	0.455	<0.001	74%
Functional regions	MoRF content	0.014	0.008	<0.001	68%	0.014	0.008	<0.001	67%	0.018	0.012	<0.001	53%
	Number of MoRF regions (per 1000 AAs)	1.843	1.090	<0.001	69%	1.834	1.096	<0.001	67%	2.398	1.577	<0.001	52%
	Disordered protein binding (DPB) content	0.230	0.092	<0.001	149%	0.231	0.093	<0.001	148%	0.201	0.074	<0.001	171%
	Number of DPB regions (per 1000 AAs)	8.764	3.665	<0.001	139%	8.784	3.663	<0.001	140%	8.402	3.270	<0.001	157%
Functional motifs (ELMs)	Proteolytic cleavage sites (per 1000 AAs)	0.023	0.033	0.017	−43%	0.022	0.032	0.039	−46%	0.000	0.003	0.044	−89%
	Degradation sites (per 1000 AAs)	0.014	0.019	0.051	−31%	0.014	0.018	0.016	−32%	0.003	0.004	0.342	−40%
	Docking sites for catalysis (per 1000 AAs)	0.137	0.074	<0.001	85%	0.137	0.074	<0.001	85%	0.010	0.005	0.031	88%
	Non-catalytic ligand binding sites (per 1000 AAs)	0.278	0.240	0.026	15%	0.257	0.234	0.061	10%	0.119	0.043	<0.001	174%
	PTM sites (per 1000 AAs)	0.180	0.084	<0.001	114%	0.183	0.087	<0.001	110%	0.028	0.013	0.022	121%
	Targeting sites for localization (per 1000 AAs)	0.149	0.053	<0.001	182%	0.146	0.058	<0.001	151%	0.000	0.008	<0.001	−100%
PTMs	Number of PTMs (per 1000 AAs)	413.3	275.3	<0.001	50%	412.5	275.8	<0.001	50%	410.4	277.9	<0.001	48%
	Number of residues with PTMs (per 1000 AAs)	325.9	226.5	<0.001	44%	325.7	227.1	<0.001	43%	327.3	230.0	<0.001	42%

Table 1. Cont.

Type of Functional/Structural Characteristic ¹	Measure ¹	Results for Hubs Involved in Hub PPIs				Results for all Hub Interactors Involved in Hub PPIs				Results for Hub Interactors Involved in Hub PPIs that Exclude Hubs			
		Avg for Disorder-Enriched Proteins	Avg for Remaining Proteins	p-Value	Relative Difference	Avg for Disorder-Enriched Proteins	Avg for Remaining Proteins	p-Value	Relative Difference	Avg for Disorder-Enriched Proteins	Avg for Remaining Proteins	p-Value	Relative Difference
PTMs	Acetylation sites (per 1000 AAs)	9.603	7.954	<0.001	21%	9.481	7.919	<0.001	20%	8.951	7.759	<0.001	15%
	ADP-ribosylation sites (per 1000 AAs)	32.02	18.52	<0.001	73%	31.69	18.39	<0.001	72%	30.97	17.50	<0.001	77%
	Amidation sites (per 1000 AAs)	48.67	36.64	<0.001	33%	49.05	36.59	<0.001	34%	49.62	39.75	<0.001	25%
	Carboxylation sites (per 1000 AAs)	39.24	23.42	<0.001	68%	39.22	23.43	<0.001	67%	38.22	22.28	<0.001	72%
	C-linked_glycosylation sites (per 1000 AAs)	0.200	0.261	<0.001	−31%	0.212	0.287	<0.001	−35%	0.262	0.331	0.004	−21%
	Farnesylation sites (per 1000 AAs)	0.007	0.008	0.303	−21%	0.006	0.008	0.059	−30%	0.008	0.021	0.083	−61%
	Geranylgeranylation sites (per 1000 AAs)	0.008	0.010	0.059	−32%	0.007	0.009	0.045	−40%	0.011	0.031	0.226	−63%
	GPI anchor amidation sites (per 1000 AAs)	1.695	1.621	0.050	5%	1.689	1.626	0.056	4%	1.917	1.865	0.500	3%
	Hydroxylation sites (per 1000 AAs)	26.75	12.56	<0.001	113%	26.98	13.16	<0.001	105%	26.18	12.99	<0.001	101%
	Methylation sites (per 1000 AAs)	15.41	10.27	<0.001	50%	15.33	10.27	<0.001	49%	13.98	9.97	<0.001	40%
	Myristoylation sites (per 1000 AAs)	0.012	0.017	0.002	−39%	0.014	0.013	0.908	1%	0.021	0.015	0.222	41%
	N-linked glycosylation sites (per 1000 AAs)	3.754	3.772	0.577	0%	3.748	3.781	0.356	−1%	3.474	3.660	0.020	−5%
	N-terminal acetylation sites (per 1000 AAs)	0.525	0.517	0.345	2%	0.542	0.515	0.035	5%	0.769	0.797	0.321	−3%
	O-linked glycosylation sites (per 1000 AAs)	14.22	8.18	<0.001	74%	14.71	8.13	<0.001	81%	14.21	7.81	<0.001	82%
	Palmitoylation sites (per 1000 AAs)	1.334	1.734	<0.001	−30%	1.342	1.786	<0.001	−33%	1.783	2.751	<0.001	−35%
	Phosphorylation sites (per 1000 AAs)	44.52	21.63	<0.001	106%	43.75	21.71	<0.001	102%	42.60	18.45	<0.001	131%
	PUPylation sites (per 1000 AAs)	1.949	2.603	<0.001	−34%	1.920	2.633	<0.001	−37%	1.838	2.436	<0.001	−25%
	Pyrrolidone carboxylic acid sites (per 1000 AAs)	4.803	4.357	<0.001	10%	4.813	4.366	<0.001	10%	4.998	4.883	0.197	2%
	Sulfation sites (per 1000 AAs)	3.589	2.291	<0.001	57%	3.603	2.305	<0.001	56%	3.148	2.026	<0.001	55%
	SUMOylation sites (per 1000 AAs)	6.844	6.484	0.003	6%	6.876	6.467	<0.001	6%	6.457	6.682	0.035	−3%
Ubiquitination sites (per 1000 AAs)	9.377	6.924	<0.001	35%	9.378	6.984	<0.001	34%	9.194	6.931	<0.001	33%	

¹ Abbreviations: long intrinsically disordered protein region (LIDPR); amino acid (AA); disordered protein binding (DPB); eukaryotic linear motif (ELM); molecular recognition feature (MoRF; short DPB region); posttranslational modification (PTM).

2.2. Functional and Structural Characteristics of Proteins Involved in the Disorder-Enriched Hub Protein-Protein Interactions

We focus our analysis on the hubs and their interactors that are involved in hub PPIs and which have higher than expected density in the two-dimensional intrinsic disorder space. This group of PPIs has a density that is higher than expected by at least 50% (regions colored from orange to red in Figure 1) and includes hubs and interactors that have high levels of intrinsic disorder. Furthermore, besides analyzing the complete set of all interactors, we also consider a subset of interactors that excludes hubs. In other words, in the latter case we exclude interactors involved in the hub-hub interactions. We analyze a comprehensive set of functional and structural characteristics to uncover which of them, if any, is specific to a given group of proteins (hubs, all interactors, and interactors that exclude hubs) that are involved in the disorder enriched hub PPIs. We accomplish that by comparing the quantity of a given characteristic between the proteins involved in these disorder-enriched hub PPIs and proteins involved in the remaining PPIs (Table 1).

As expected, both hubs and the two sets of hub interactors that are engaged in the high-density and disorder enriched hub PPIs are characterized by a significantly higher amount of intrinsic disorder (p -value < 0.001 ; see “Structural properties” rows in Table 1). The corresponding relative enrichment is $>130\%$ when quantifying the disorder by both content and normalized number of the long IDPRs. These observations are further supported by the significantly lower number of disulfide links, which can be used as a proxy for the formation of a stable tertiary structure, among the hubs and hub interactors that are engaged in the disorder-enriched hub PPIs (p -value < 0.001 , decrease by over 46%). We break down this analysis based on the subcellular location of these proteins. We annotate the locations based on the Universal Protein resource (UniProt) tags using a protocol defined in [64]. We analyze the results for all considered human proteins, hubs, hub interactors, and hub interactors that exclude hubs; the latter three proteins sets are subdivided into proteins associated with the disorder-enriched hub PPIs versus the remaining proteins. To ensure statistically sound estimates of the disorder content and number of disulfide bonds, our analysis includes 10 cellular locations for which we found at least 10 proteins for all considered protein sets, see Figure 2. As expected, we observe that the counts of the disulfide bonds vary between the subcellular locations, with substantially larger numbers for secreted proteins when compared to the other cellular compartments (black bars in Figure 2A). This is in agreement with prior research, which similarly showed enrichment in the eukaryotic proteins associated with secretory pathways and lower abundance in the other locations, in particular in the cytoplasm [65]. More importantly, the disorder-enriched protein sets are characterized by the higher numbers of the disulfides (blue, red, and green bars with horizontal stripes in Figure 2A) compared to the corresponding sets of the proteins that exclude the disorder-enriched proteins (solid blue, red, and green bars in Figure 2A). Figure 2B confirms the expected trend that proteins involved in the disorder-enriched PPIs are characterized by the increased levels of disorder (solid bars are higher than the corresponding striped bars). It also show that the levels of the disorder content vary between subcellular compartment, with the highest values being found in the nucleus and lowest levels being among the membrane and secreted proteins. This is in agreement with the results shown in several recent studies [25,50,66].

Interestingly, Table 1 reveals that proteins associated with high-density disorder-enriched hub PPIs are characterized by similar evolutionary conservation levels and size when compared to the other proteins; the difference is below 20%. However, while the levels of conservation of the long disordered regions are similar for the hubs and the set of all their interactors, they are substantially lower, by over 70%, for the long IDPRs in interactors that exclude hubs and that are not part of the high-density disorder enriched hub PPIs. This finding is related to a recent observation that disordered regions in hubs are enriched in the conserved residues [67]. Furthermore, our results suggest that the long IDPRs in these disorder-depleted interactors are not only less evolutionarily conserved but also less likely to be involved in the PPIs; their content of the disordered protein binding (DPB) regions in Table 1 is much lower than that for the other protein sets.

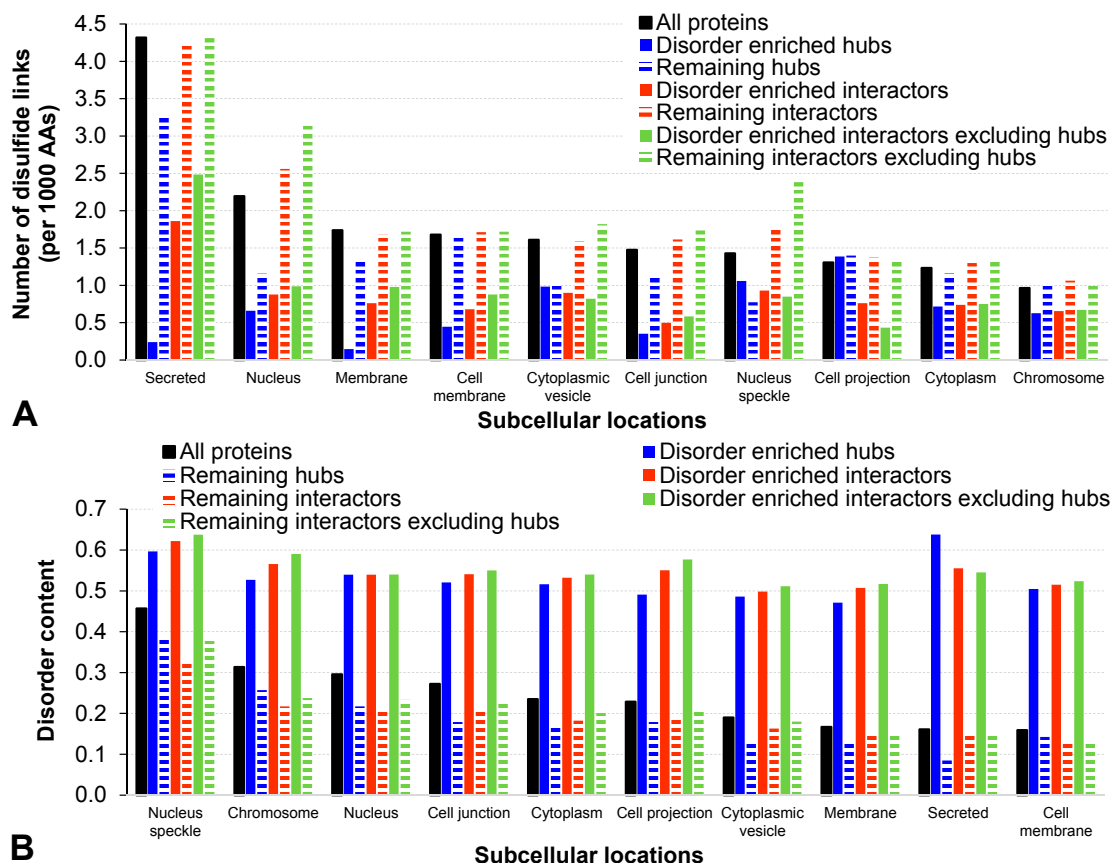


Figure 2. Number of disulfide bonds panel (A) and disorder content panel (B) in specific subcellular locations. We report the values for all human proteins (black bars), hubs (solid blue), hub interactors (solid red), and hub interactors that exclude hubs (solid green) that are associated with the disorder-enriched hub PPIs vs the remaining hubs (blue horizontal stripes), hub interactors (red horizontal stripes), and hub interactors that exclude hubs (green horizontal stripes), respectively, for each location. We consider all locations that include at least 10 proteins for each of the seven protein sets. The locations in panel (A,B) are sorted in descending order by the values of the number of disulfide bonds (disorder content) for all proteins (black bars).

The high-density disorder-enriched hubs and their interactors have significantly more MoRF (molecular recognition feature) and DPB (disordered protein binding) regions (p -value < 0.001; increase by over 50% and over 130%, respectively). This is likely due to the overall enrichment in the intrinsic disorder. As expected, the content of putative MoRFs is overall substantially lower than the content of the putative DPB regions. This is because MoRFs are a subset of the DPB regions. We emphasize that the abundance of the DPB regions in the interactors that exclude hubs is lower when compared to the hubs and the interactors that include hubs. More specifically, the DPB content equals 0.23 (the normalized number of protein binding regions equals 3.7) for the latter protein sets vs. 0.20 (3.3) for the interactors that exclude hubs. We also further investigate the relation between the MoRFs and the DPB regions given their close functional relation. The average point biserial correlation coefficients, which quantify relations for binary variables, that are measured between the predicted MoRF and DPB residues over the hub proteins, their interactors and the interactors that exclude hubs are equal to 0.24, 0.22, and 0.21, respectively. Furthermore, the Pearson correlation coefficients calculated between the per protein MoRF and DPB content value for the hubs, interactors, and interactors that exclude hubs are equal to 0.32, 0.33, and 0.33, respectively. Both sets of measurements suggest a weak correlation between the two related functional aspects of IDPRs. This expected observation to some extent validates the corresponding predictions collected from the two methods, MoRFpred and DisoRDPbind.

Furthermore, results based on the analysis of the abundance of a wide range of functional motifs defined with ELMs show that proteins linked with the high-density disorder enriched hub PPIs feature a modest decrease by 30% to 40% in the number of cleavage and degradation sites. More importantly, they also have significantly more catalytic sites (p -value < 0.001 , increase by over 80%) and posttranslational modification (PTM) sites (p -value < 0.001 , increase by over 110%). Furthermore, we observe divergent results concerning the non-catalytic binding sites and targeting sites between the interactors that exclude hubs and the other two sets of proteins, hubs and interactors that include hubs. While the abundance of the non-catalytic binding sites does not differ between the disorder-enriched and remaining proteins for the hubs and the set of all interactors, it drops significantly by over 170% for the disorder-depleted interactors that exclude hubs. The number of targeting sites for the interactors that exclude hubs is very low (near zero) when compared to a relatively high number of these sites for hubs. Furthermore, the number of the targeting sites is significantly increased in the disordered-enriched hubs/hub-including interactors compared to the remaining hubs/hub-including interactors (p -value < 0.001 , increase by 150%).

PTMs extend the range of physico-chemical properties of amino acids, thereby diversifying structures and functions of target proteins, affecting their conformational stabilities, modulating their interactions with nucleic acids, membranes, ligands or other proteins, and targeting modified proteins to specific subcellular compartments [68–70]. In fact, the presence of various PTMs (as many as 300 different PTMs are known to occur physiologically [71]) extends the repertoire of amino acids found in proteins from 20 primary residues utilized during protein biosynthesis to more than 140 physico-chemically different residues, thereby dramatically extending the complexity of eukaryotic proteomes [68]. PTMs can affect both structured proteins/regions and IDPs/IDPRs. PTMs that preferentially affect IDPs/IDPRs are common in proteins engaged in regulatory and signaling functions. These disorder-associated PTMs were shown to include acetylation, acylation, adenylation, ADP ribosylation, amidation, carboxylation, formylation, glycosylation, methylation, sulfation, phosphorylation, prenylation, ubiquitination, and Ubl-conjugation [72]. Previous bioinformatics analyses revealed that protein phosphorylation sites are often found in IDPRs [73–75], and there is a high correspondence between the prediction of disorder and the occurrence of phosphorylation [75]. In addition to phosphorylation, several other types of PTMs, such as acetylation, protease digestion, ubiquitination, fatty acid acylation, and methylation, have also been observed to occur in IDPRs [72–74,76].

The analysis using the PTMs annotated with ModPred [77], which is independent from and extends the analysis based on ELMs, confirms the enrichment in the number of PTM sites and the number of residues that have PTM sites. We computed both quantities since some residues may have multiple PTM sites. Table 1 provides a detailed breakdown of this enrichment across the 21 major PTMs. The hubs and their interactors (including the set of interactors without hubs) that are involved in the high-density disorder-enriched hub PPIs have a significantly larger number of acetylation, ADP-ribosylation, amidation, carboxylation, hydroxylation, methylation, O-linked glycosylation, phosphorylation, sulfation, and ubiquitination sites. Our results specifically concerning proteins that participate in the disorder-enriched hub PPIs are in good agreement with the prior analyses of a more generic set of IDPRs [72–76]. Interestingly, MoRFs in the eukaryotic proteins were shown to be significantly enriched in PTM sites, suggesting that PTMs are crucial for the modulation of MoRF-driven protein recognition related to transcriptional and developmental processes [77]. This provides a feasible explanation for our observation that both MoRFs and certain PTM sites are significantly more abundant among our proteins of interest.

Figure 3 summarizes the results from Table 1 and reveals that they are similar between the hubs and their interactors that include hubs. It lists the considered functional and structural features that are sorted in descending order by the value of their relative difference between the proteins involved in the disorder-enriched hub PPIs and the other proteins. Figure 3A shows the characteristics that have the statistically significant and highest levels of enrichment among the disorder-enriched hubs and interactors that include hubs contain a substantial increase in the amount of DPB regions, a larger

number of long IDPRs, and a significantly increased content of intrinsic disorder. These characteristics also include considerably increased rate of targeting, catalytic, and PTM sites. Figure 3B reveals that the hydroxylation, phosphorylation, O-linked glycosylation, ADP-ribosylation, carboxylation, sulfation, and methylation sites are the most enriched PTMs among the hubs and all their interactors that participate in the disorder-enriched hub PPIs. Furthermore, the far right part of Figure 3 reveals that the patterns of enrichment are different for the interactors that exclude hubs. They are characterized by low abundance of the targeting sites, which nears zero for the disorder-enriched interactors. These disorder-enriched interactors also have significantly more ligand binding sites and are more conserved when compared to the disorder-depleted interactors. However, the characteristics of the PTM sites are similar.

HUBS			Type and name of functional and structural characteristics	ALL INTERACTORS		INTERACTORS excluding HUBS	
Relative difference	Statistical significance			Statistical significance	Relative difference	Statistical significance	Relative difference
182%	+	functional	Targeting sites for localization	+	151%	+	-100%
149%	+	functional	Disordered protein binding content	+	148%	+	171%
141%	+	structural	Disorder content	+	138%	+	148%
139%	+	functional	Disordered protein binding segments	+	140%	+	157%
133%	+	structural	Long IDPRs	+	132%	+	149%
114%	+	functional	PTM sites	+	110%	=	121%
85%	+	functional	Docking sites for catalysis	+	85%	=	88%
69%	+	functional	MoRF segments	+	67%	+	52%
68%	+	functional	MoRF content	+	67%	+	53%
50%	+	functional	PTMs	+	50%	+	48%
44%	+	functional	Residues with PTMs	+	43%	+	42%
16%	+	sequence	Average chain length	+	12%	+	21%
15%	=	functional	Ligand binding sites	=	10%	+	174%
-1%	=	conserv.	Conservation of long IDPRs	=	0%	+	74%
-16%	-	conserv.	Conservation of proteins	-	-13%	-	-8%
-31%	=	functional	Degradation sites	=	-32%	=	-40%
-43%	=	functional	Proteolytic cleavage sites	=	-46%	=	-89%
-62%	-	structural	Disulfide links	-	-63%	-	-47%

(A)

HUBS			PTM type	ALL INTERACTORS		INTERACTORS excluding HUBS	
Relative difference	Statistical significance			Statistical significance	Relative difference	Statistical significance	Relative difference
113%	+		Hydroxylation	+	105%	+	101%
106%	+		Phosphorylation	+	102%	+	131%
74%	+		O-linked glycosylation	+	81%	+	82%
73%	+		ADP-ribosylation	+	72%	+	77%
68%	+		Carboxylation	+	67%	+	72%
57%	+		Sulfation	+	56%	+	55%
50%	+		Methylation	+	49%	+	40%
35%	+		Ubiquitination	+	34%	+	33%
33%	+		Amidation	+	34%	+	25%
21%	+		Acetylation	+	20%	+	15%
10%	+		Pyrrolidone carboxylic acid	+	10%	=	2%
6%	+		SUMOylation	+	6%	=	-3%
5%	=		GPI anchor amidation	=	4%	=	3%
2%	=		N-terminal acetylation	=	5%	=	-3%
0%	=		N-linked glycosylation	=	-1%	=	-5%
-21%	=		Farnesylation	=	-30%	=	-61%
-30%	-		Palmitoylation	-	-33%	-	-35%
-31%	-		C-linked glycosylation	-	-35%	=	-21%
-32%	=		Geranylgeranylation	=	-40%	=	-63%
-34%	-		PUPylation	-	-37%	-	-25%
-39%	=		Myristoylation	=	1%	=	41%

(B)

Figure 3. Significance of the differences in the functional and structural characteristics between hubs (on the left), all hub interactors, and hub interactors that exclude hubs (on the right) that are associated with the disorder-enriched hub PPIs when compared to the remaining hubs (on the left) and the corresponding remaining interactors (on the right). Panel (A) summarizes the results concerning structural characteristics, functional regions and motifs, evolutionary conservation and the overall abundance of PTMs. Panel (B) gives detailed results for specific types of PTMs. We reported relative differences and their statistical significance. The characteristics are sorted in descending order by their relative differences for the hubs. The characteristics are color-coded as follows: green for large (relative difference > 20%) and statistically significant (p -value < 0.001) enrichment; red for large (relative difference < -20%) and statistically significant (p -value < 0.001) depletion; and blue for lack of large and significant differences ($|$ relative difference $|$ < 20% or p -value over 0.001). Abbreviations: Eukaryotic linear motif (ELM); molecular recognition feature (MoRF; short disordered protein binding region); and posttranslational modification (PTM).

2.3. Proteins Involved in the Disorder-Enriched Hub Protein-Protein Interactions and Human Diseases

We analyzed whether the abundance of hubs and their interactors that participate in the disorder-enriched hub PPIs is biased among a subset of human proteins that are associated with diseases. Table 2 summarizes these results. When considering the complete set of over 4000 disease-linked proteins, we show that they are significantly enriched among the proteins that participate in the high-density disorder-enriched hub PPIs (p -value < 0.001, 20% increase when compared to the human proteome). Analysis that focuses on the 18 major classes of diseases, each associating with at least 100 human proteins, reveals a significant enrichment for 11 disease classes, significant depletion for five classes and no significant differences for the remaining two classes. We observed the largest enrichment by over 600% in the disorder-enriched hub PPIs for the neoplasm-centered diseases that include a variety of cancers (MeSH class C04). Several other classes of diseases are also characterized by high values of over 50% enrichment. These include stomatognathic maladies (C07) and diseases of the endocrine system (C19), digestive system (C06), respiratory tract (C08), female urogenital system (C13), nervous system (C10), and musculoskeletal diseases (C05).

Table 2. Analysis of enrichment in the disease associated proteins. We compare the abundance of the disease-linked proteins, measured as the number of the corresponding PPIs that include these proteins that is computed per protein, which are among the hubs and interactors involved in the disorder-enriched hub PPIs vs. their enrichment among PPIs for all human proteins. The 4296 diseases that are associated with human proteins are grouped into disease classes using the MeSH hierarchy. The classes are sorted in descending order by the value of the relative difference between the abundance among the disorder-enriched hub PPIs and all PPIs. To ensure that results are statistically robust to represent diverse subpopulations of these proteins, we select 50% of proteins at random from each of the two protein sets, quantify a given characteristic for each set, and repeat this ten times. We used a larger fraction to define the populations when compared to the results in Table 1 to accommodate the small number of proteins associated with some of the disease classes. We report the average of these repetitions and relative difference between averages for the two protein sets. We also evaluate the significance of the differences between these measurements. We use the t -test if the data are normal (we test normality with the Anderson-Darling test at p -value = 0.05); otherwise we use the Wilcoxon test.

Disease Class (ID at the Second MeSH Level)	Average for Disorder-Enriched Proteins	Average for all Human Proteins	p -Value	Relative Difference
ALL Diseases	0.519	0.433	<0.001	20%
Neoplasms, including cancers (C04)	3.150	0.436	<0.001	622%
Stomatognathic Diseases (C07)	1.833	0.437	<0.001	320%
Endocrine System Diseases (C19)	1.594	0.432	<0.001	269%
Digestive System Diseases (C06)	1.440	0.436	<0.001	230%
Respiratory Tract Diseases (C08)	1.060	0.437	<0.001	143%
Female Urogenital Diseases and Pregnancy Complications (C13)	0.855	0.437	<0.001	96%
Nervous System Diseases (C10)	0.775	0.432	<0.001	79%
Musculoskeletal Diseases (C05)	0.654	0.433	<0.001	51%
Specific disease classes	0.543	0.434	<0.001	25%
Hemic and Lymphatic Diseases (C15)	0.543	0.434	<0.001	25%
Pathological Conditions, Signs and Symptoms (C23)	0.467	0.434	<0.001	8%
Congenital, Hereditary, and Neonatal Diseases and Abnormalities (C16)	0.456	0.433	<0.001	5%
Male Urogenital Diseases (C12)	0.459	0.436	0.009	5%
Immune System Diseases (C20)	0.409	0.435	0.81	−6%
Eye Diseases (C11)	0.373	0.437	<0.001	−17%
Cardiovascular Diseases (C14)	0.347	0.437	<0.001	−26%
Nutritional and Metabolic Diseases (C18)	0.248	0.433	<0.001	−75%
Skin and Connective Tissue Diseases (C17)	0.247	0.435	<0.001	−76%
Otorhinolaryngologic Diseases (C09)	0.162	0.431	<0.001	−165%

These observations are in line with the results of prior rigorous investigations of IDP/IDPR functions and dysfunctions that already have revealed that intrinsic disorder is prevalent among the disease-related proteins, with dysfunction of IDPs being often related to the development of various

pathological conditions [30–33,35,36,78,79]. Many diseases are correlated with proteins predicted to be disordered, whereas no disease-associated proteins were found to be correlated with the absence of disorder [72,80,81]. In agreement with our results, the prevalence of IDPs/IDPRs was illustrated for various human diseases, such as cancer [82] (MeSH class C04), cystic fibrosis [83] (C06, C08 and C16) and neurodegenerative diseases [30] (C10). Among the cancer associated proteins with experimentally confirmed IDPRs are p53 [84], BRCA1 [85], EWS [86], HPV proteins [87], PTEN [88], and cancer/testis antigens (CTAs) [89]. The most notable disordered players associated with the pathogenesis of neurodegeneration are α -synuclein (Parkinson's disease (PD), dementia with Lewy bodies, Alzheimer's disease (AD), Down's syndrome, and other synucleinopathies) [46], amyloid β and tau protein (AD), prions (Creutzfeldt-Jakob disease, scrapie, bovine spongiform encephalopathy), ataxin (Spinocerebellar ataxia) [30], as well as TAR-DNA binding protein-43 (TDP-43), RNA-binding protein FUS, polypeptides generated as a result of their intronic hexanucleotide expansions of the cofilin-binding protein C9orf72, and profilin-1 (PFN1) associated with amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration [90].

High prevalence of intrinsic disorder in proteins associated with the various human diseases is related (at least in part) to the functional roles of IDPs/IDPRs in regulation and control of biological processes. In fact, it was pointed out that the abundant involvement of IDPs/IDPRs in various signaling, regulation, and recognition processes is defined by a broad set of functional advantages of IDPs/IDPRs over their ordered counterparts [21,91,92]. These advantages include (but are not limited to) the ability of IDPs/IDPRs to react easily and quickly and change their conformation in response to changes in the environment, their structural plasticity and conformational adaptability, their binding promiscuity, and their unique capability to fold differently while interacting with different binding partners [60,93]. These same factors also define the capability of IDPs/IDPRs to play diverse roles in the modulation and control of functions of their binding partners and in promotion of the assembly of supra-molecular complexes. Furthermore, the ability of IDPs/IDPRs to return to the highly flexible conformations after the completion of a particular function and their predisposition to adopt different conformations depending on their environment are their unique physiological properties. They define the ability of IDPs/IDPRs to have different functions in different cellular contexts according to a specific conformational state [21]. Furthermore, in line with our observations in this study, biological activities of IDPs/IDPRs can be precisely and tightly controlled and regulated by various PTMs [21,77,94,95] and by alternative splicing (AS) [29,96,97]. In fact, the ability of AS to generate multiple protein isoforms with highly diverse sets of regulatory elements [96] is determined by the mosaic structure of IDPs/IDPRs that are known to contain multiple relatively short, functional elements, which, being spread within the amino acid sequences, define the multi-functionality of these proteins [98]. Clearly, AS-driven removal of pieces of an IDP/IDPR sequence containing different functional elements could dramatically reshuffle such multi-functionality [29,97,99].

The aforementioned factors define the capacity of IDPs/IDPRs to serve as hubs and as important controllers and regulators of various biological processes. However, all these factors also suggest that IDPs and hybrid proteins have to be tightly regulated and controlled themselves, since controllers/regulators have to be available in appropriate amounts and not to be present longer than needed [34,36,98]. In fact, there are numerous studies that not only emphasized the important roles of disordered regulators in signaling [100,101], regulation [28,102–108], cell protection [109], protein protection [110,111], and cellular homeostasis [112,113], but which also showed that IDPs/IDPRs are concisely controlled by themselves via multiple mechanisms, such as interaction with chaperones [114–118] or nanny proteins [119], partner binding [120–126], various PTMs [127–131], and regulated degradation [132–137]. Obviously, the consequences of a tightly controlled process, which suddenly goes out of control, could be disastrous. Similarly, the consequences of the misbehavior of an important regulator are disastrous too, and the development of particular pathological conditions is often linked to the failure of a specific peptide or protein (or a set proteins and peptides) to adopt its functional conformational state. These maladies are commonly known as

conformational diseases, each originating from the dysfunction and misbehavior of a particular protein (or a set of proteins). Some disease-related proteins are characterized by the ability to spontaneously form pathologic conformation(s), whereas misbehavior of other proteins is determined by their mutations or (impaired) interactions with chaperones, intracellular or extracellular matrices, other proteins, nucleic acids, small molecules, other endogenous factors, or by the exposure to internal or external toxins. Furthermore, impaired PTMs, wrecked AS, an increased probability of degradation, impaired trafficking, loss of binding partners, or oxidative damage can also be pathogenic. Finally, one should keep in mind that all these factors can act independently, additively, or synergistically [33].

2.4. Several Illustrative Examples of Proteins Involved in the Disorder-Enriched Hub Protein-Protein Interactions

The next few sections provide description of several illustrative examples of proteins involved in the disorder-enriched hub PPIs. Supplementary Materials Tables S1 and S2 list the 100 most and the 100 least disordered-enriched human hubs, respectively, and include information about their putative disorder content, normalized number of DPB regions, and degree (promiscuity). These tables also provide a list of all their interactors, interactors involved in the disorder enriched hub PPIs, and the disorder content for each listed interactor. Supplementary Materials Table S1 shows that binding promiscuity N is not directly correlated with protein disorderedness. In fact, one of the most disordered proteins in this dataset, the intracellular hyaluronan-binding protein 4 (UniProt ID: Q5JVS0) with the disorder content of 0.88, is characterized by a relatively low number of interactors ($N = 29$), for a hub protein. On the other hand, the most promiscuous binder in this set ($N = 415$), BAG family molecular chaperone regulator 3 (UniProt ID: O95817), is characterized by the disorder content of 0.86. Supplementary Materials Table S2 reveals that the most ordered hubs have anywhere between 25 and 8548 interactors. Therefore, we describe 15 representative hubs that are grouped into the three following categories: (a) the most disordered human hubs; (b) highly disordered human hubs with the largest number of interactors; and (c) highly connected ordered hubs. We note that the latter set of hubs by definition is not involved in the disorder enriched hub PPIs. This is why the Supplementary Materials Table S2 does not include the associated list of the interactors that participate in the disorder enriched hub PPIs.

Furthermore, we analyzed the putative DPB regions and disorder content for these proteins using not only the same tools that were used to characterize the entire human interactome, but also a second set of independent tools. Besides using DisoRDPbind [138,139], we also utilized ANCHOR [140] to predict putative disordered protein binding regions; we call these ANCHOR-identified binding sites (AIBSs). We annotated the putative disorder and computed disorder content using the MobiDB platform that applies eight methods to derive a consensus-based prediction [141–143], on the top of the analysis that is based on the consensus of five versions of IUPred [144] and ESpritz [145] methods that we utilized to annotate the human PPI network. The results of the disorder analysis with MobiDB for the selected 15 proteins are summarized in Figure 4. These proteins include five of the most disordered human hubs (Figure 4A–E), five highly disordered human hubs with the highest promiscuity (Figure 4F–J), and five highly connected ordered hubs (Figure 4K–O). The results with MobiDB are in good agreement with the results based on the other two predictors. The average disorder content equals 0.88 (for IUPred and Espritz-based analysis) vs. 0.79 (for MobiDB-based analysis) for the five most disordered hubs, 0.74 vs. 0.61 for the most promiscuous disordered hubs and 0 vs. 0 for the highly connected ordered hubs.

We also note the average disorder content values of the interactors of the selected hubs, which were computed with the consensus of five versions of IUPred and Espritz (Supplementary Materials Tables S1 and S2). These values range between 0.23 and 0.38 for the most disordered hubs, 0.19 and 0.34 for the most promiscuous hubs, and between 0.20 and 0.33 for the highly connected ordered hubs. As expected, the average disorder content of the interactors that participate in the disorder enriched hub PPIs is much higher and varies from 0.84 to 0.88 for the most disordered hubs and from 0.50 to 0.89 for the most promiscuous hubs.

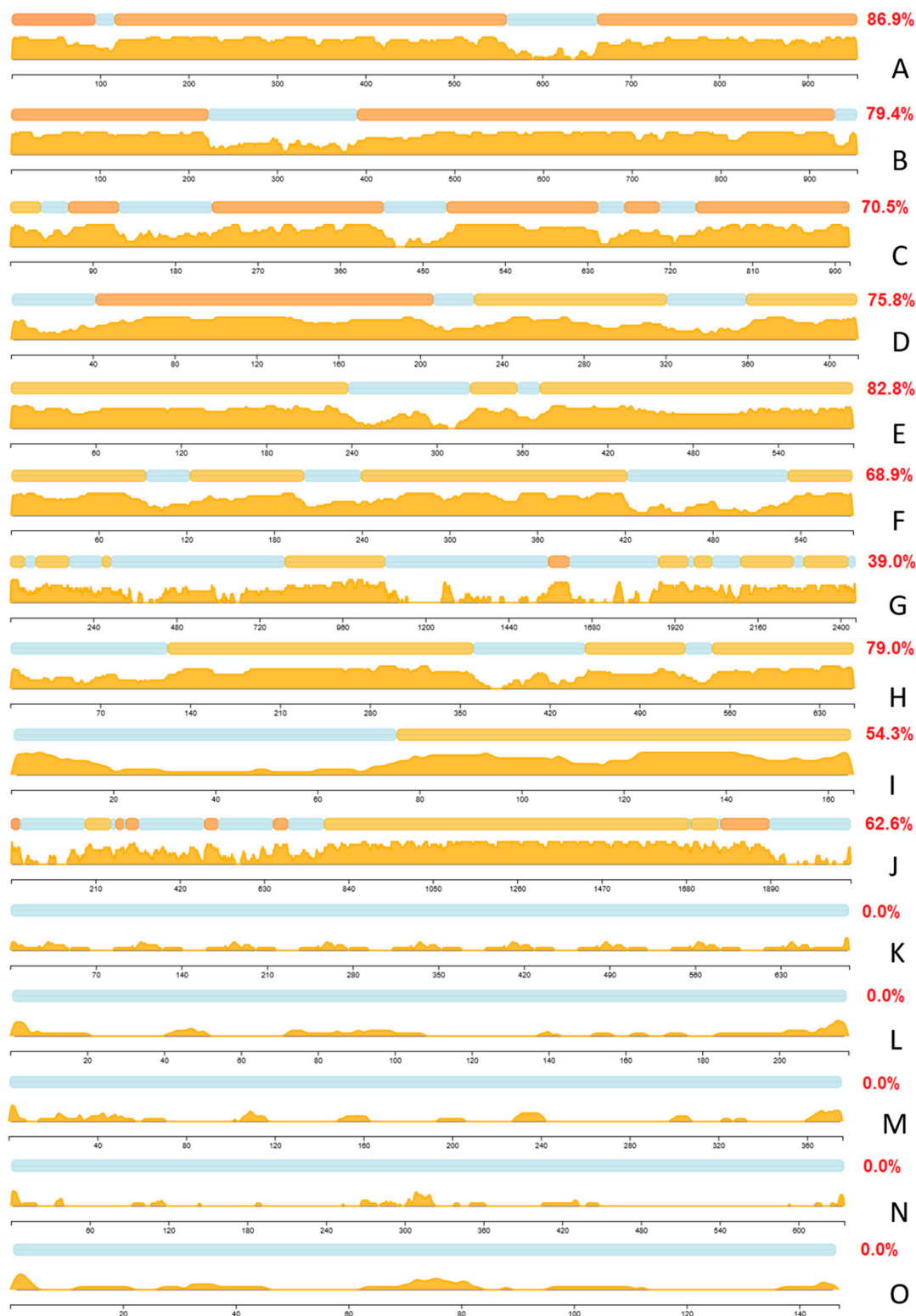


Figure 4. Intrinsic disorder levels in the disordered human hubs characterized by the highest levels of disorder (A–E), highly disordered hubs characterized by the highest levels of interactability (F–J), and ordered hubs with the highest interactability levels (K–O). The disorder was annotated using the MobiDB platform [141–143]; disorder content is shown in red font. Each plot represents disorder tendencies in two forms—by bar plots showing location of IDPRs and by area plots showing sequence distribution of consensus disorder scores evaluated by MobiDB lite disorder predictor [143].

(A) Thyroid hormone receptor-associated protein 3 (UniProt ID: Q9Y2W1). (B) Zinc finger CCCH domain-containing protein 18 (UniProt ID: Q86VM9). (C) Scaffold attachment factor B1 (UniProt ID: Q15424). (D) Intracellular hyaluronan-binding protein 4 (UniProt ID: Q5JVS0). (E) TATA-binding protein-associated factor 2N (UniProt ID: Q92804). (F) BAG family molecular chaperone regulator 3 (UniProt ID: O95817). (G) CREB-binding protein (UniProt ID: Q92793). (H) RNA-binding protein EWS (UniProt ID: Q01844). (I) Cyclin-dependent kinase inhibitor 1 (UniProt ID: P38936). (J) Mediator of DNA damage checkpoint protein 1 (UniProt ID: Q14676). (K) Ubiquitin (UniProt ID: P0CG48; 8548 interactors). (L) Growth factor receptor-bound protein 2 (UniProt ID: P62993; 804 interactors). (M) Actin (UniProt ID: P60709; 263 interactors). (N) Protection of telomeres protein 1 (UniProt ID: Q9NUX5; 200 interactors). (O) Protein mago nashi homolog (UniProt ID: P61326; 190 interactors).

2.4.1. The Most Disordered Hubs

Curiously, the most disordered human hubs considered in this section are nuclear proteins. This observation is in line with the known fact that the nuclei of eukaryotic cells are often enriched in IDPs or hybrid proteins with long functional IDPRs [50,66]. Binding promiscuity of these mostly disordered nuclear hubs is driven by the presence of multiple long disorder-based protein binding regions. Supplementary Materials Table S1 shows that the 31 hubs that have disorder content >0.8 on average have 13 such regions per unit of 1000 residues while the average for the complete set of 100 most disordered-enriched human hubs equals 11. The normalized (per 1000 residues) number of these putative long DPB regions for the top 100 hubs is modestly correlated with the disorder content (Pearson correlation coefficient = 0.26). The subsequently discussed five proteins with the highest disorder content are sorted in descending order by their disorder content.

Thyroid hormone receptor-associated protein 3 (Supplementary Materials Table S1: UniProt ID Q9Y2W1; disorder content = 0.89; 72 interactors; average disorder content of all (disorder enriched) interactors = 0.29 (0.85)). RNA processing factor THRAP3 is a 955 residue-long arginine/serine-rich domain-containing nuclear protein involved in pre-mRNA splicing [146]. As many other highly disordered proteins, THRAP3 (which is also known as Trap150) has several functions ranging from regulation of signal-induced alternative splicing [147], nuclear mRNA decay [146], response to DNA damage [148], transcriptional co-activation [149], adipocyte differentiation [150], to the regulation of the circadian clock [151]. Together with another arginine/serine-rich domain-containing protein BCLAF1 (Bcl-2-associated transcription factor 1), as well as Hakai, Virilizer homolog, KIAA0853 (also known as zinc finger CCCH domain-containing protein 13, ZC3H13), RBM15 (RNA-binding motif protein 15), several general splicing regulators, and Wilms' tumor 1-associating protein (WTAP), THRAP3 was shown to form the WTAP complex that serves as an important component of the RNA processing machinery [152]. Furthermore, THRAP3 seems to control WTAP localization to nuclear speckles [152]. Recently, the important role of the aberrant THRAP3 phosphorylation in androgen-independent prostate cancer cell growth was demonstrated including the presence of a remarkable difference in the interactability of the non-phosphorylated and phosphorylated forms of this protein [153]. In addition to phosphorylation, THRAP3 is known to undergo extensive and variable PTMs at numerous sites. N-terminal third of this protein shows strong compositional biases, being serine-arginine enriched, and therefore is known as arginine/serine-rich domain. According to the ANCHOR-based analysis, THRAP3 has 25 disorder-based protein-protein binding sites (AIBSs) that range in length from 6 to 58 residues (residues 1–56, 65–122, 162–180, 188–196, 220–265, 276–317, 445–351, 374–394, 408–426, 450–460, 470–479, 513–530, 549–569, 586–597, 609–614, 626–659, 675–694, 704–725, 784–793, 809–822, 830–844, 876–883, 903–919, and 921–928) and that cover 56.3% of its sequence. Therefore, with more than half of its sequence serving as a set of potential protein-protein binding sites and with high arginine and lysine content (11.9% and 10.5%, respectively), it is not surprising that THRAP3 serves as a highly connected hub protein that is also able to efficiently interact with RNA. Figure 4A represents a disorder profile generated for THRAP3 with MobiDB database. It is clear that THRAP3 is a highly

disordered protein (disorder content = 0.869), which is in agreement with the disorder computed with IUPred and Espritz (disorder content = 0.892).

Zinc finger CCCH domain-containing protein 18 (Supplementary Materials Table S1: UniProt ID Q86VM9; disorder content = 0.89; 40 interactors; average disorder content of all (disorder enriched) interactors = 0.31 (0.88)). ZC3H18 is a 953 residue-long metal binding protein containing a C₃H₁-type zinc finger motif (residues 219–245) and interacting with RNA, likely due to its high arginine (11.6%) and lysine (8.2%) content. This protein contains three coiled coil regions (residues 105–134, 399–464, and 921–950), a poly-asparagine region (residues 201–204), and two regions of high compositional bias (proline- and serine-regions at residues 259–296 and 532–746, respectively). Similar to THRAP3, ZC3H18 protein is enriched in disorder-based protein-protein interaction sites. It contains 22 AIBSs that range in length from 8 to 78 residues, are uniformly spread over the ZC3H18 sequence, and which account for 60.8% of its amino acids (residues 1–11, 15–60, 78–102, 108–154, 167–183, 195–203, 222–258, 273–287, 304–315, 365–393, 502–516, 522–553, 561–569, 574–604, 616–623, 630–641, 648–656, 680–751, 781–793, 803–838, 856–871, and 876–953). There are multiple different PTMs such as phosphorylation, acetylation and SUMOylation in ZC3H18, suggesting that interactivity of this highly disordered promiscuous hub (see Figure 4B) is PTM-controlled. In fact, ZC3H18 was shown to interact with the hepatitis B virus post-transcriptional element (PRE), which is one of the viral RNA elements facilitating mRNA export, and with proteins of the conserved TREX complex [154], which is engaged in coupling of the transcription with messenger RNA export [155].

Scaffold attachment factor B1 (Supplementary Materials Table S1: UniProt ID Q15424; disorder content = 0.88; 65 interactors; average disorder content of all (disorder enriched) interactors = 0.31 (0.87)). SAFB1 is a 915 residue-long DNA binding protein that is known to interact with the S/MAR DNA (scaffold/matrix attachment region). This interaction is responsible for coupling of transcription and RNA processing via the formation of a molecular assembly point needed for the formation of a transcriptosomal complex that includes RNA polymerase II and SR proteins involved in the RNA splicing and containing domains with long serine-arginine repeats [156]. SAFB1 is a member of a protein family that includes SAFB2 and the SAFB-like transcriptional modulator SLTM [156]. Human SAFB1 is predicted with MobiDB to be highly disordered (see Figure 4C) and has 18 AIBSs, some as long as 95 residues (residues 1–6, 12–21, 31–43, 54–65, 116–210, 212–228, 235–297, 306–341, 353–367, 404–414, 422–438, 446–455, 463–482, 560–582, 598–609, 728–747, 797–813, and 832–915). These disordered protein-protein binding regions cover 68.2% of SAFB1 sequence and are responsible for the ability of this protein to interact with multiple partners. In fact, C-terminally located domain (residues 528–792) is known to interact with the subunit A of RNA polymerase II (POLR2A) [157], whereas interaction with SAFB2 relies on an overlapping region spanning residues 599–915. Curiously, both of these regions contain multiple AIBSs. On the other hand, the N-terminal half of SAFB1 contains a SAP motif (residues 31–65) that includes two AIBSs and, in other nuclear proteins, is involved in transcription, DNA repair, RNA processing and apoptotic chromatin degradation. The presence of the two AIBSs in SAP motif is in line with the observation that this motif might form two alpha helices separated by a turn [158]. Also, SAFB1 contains a RNA recognition motif (RRM) located at the middle of the protein (residues 408–484) and possessing several AIBSs. Finally, there are three regions with strong composition biases, residues 612–831, 625–705, and 785–903, which are enriched in arginines, glutamates, and glycines, respectively. In addition to the canonical form, human SAFB1 might have three isoforms produced by alternative splicing and has multiple PTM sites.

Intracellular hyaluronan-binding protein 4 (Supplementary Materials Table S1: UniProt ID Q5JVS0; disorder content = 0.88; 29 interactors; average disorder content of all (disorder enriched) interactors = 0.38 (0.84)). IHABP4 (also known as Ki-1/57 intracellular antigen) is a 413 residue-long highly disordered (according to the analysis with MobiDB, see Figure 4) cytoplasmic and nuclear protein that plays a role in several nuclear functions, such as the chromatin remodeling and the transcription regulation [159,160]. The role of IHABP4 in regulation of transcription is supported by the ability of this protein to interact with the adaptor protein RACK1, the transcription factor MEF2C, and the

chromatin remodeling factor CHD3 [160]. Furthermore, IHABP4 is engaged in interaction with eleven p53-binding proteins and can interact with p53 itself [159]. For its functional regulation, IHABP4 can be phosphorylated on several serine and threonine residues [160]. It has a coiled-coil domain (residues 40–64) and is predicted to have 11 AIBSs (residues 1–45, 56–88, 94–119, 133–169, 201–246, 254–262, 276–299, 317–336, 338–365, 384–400, and 408–413) that range from 6 to 46 residues in length and cover 70.5% of the IHABP4 sequence. It seems that all these AIBSs can be engaged in interaction with various binding partners. In fact, N-terminal region (residues 1–150) where shown to bind to UBC9, GADD34, YB-1, and SF2p32, whereas the C-terminal region (residues 122–413) is responsible for interaction with RACK1, CHD-3, p53-binding protein, LUN, DAXX, PIAS-3, HNG-protein 2-like 1, and Tip-60 [159]. On the other hand, the entire IHABP4 sequence is needed for interaction of this protein with p53 [159].

TATA-binding protein-associated factor 2N (Supplementary Materials Table S1: UniProt ID Q92804; disorder content = 0.88; 54 interactors; average disorder content of all (disorder enriched) interactors = 0.24 (0.87)). TAF15 (formerly known as TAF_{II}68) is a 592 residue-long nuclear RNA-binding protein that can also interact with single stranded DNA (ssDNA). Functionally, TAF15 contains an N-terminal activation domain, an RNA recognition motif (RRM, residues 234–320), and numerous (at least 21) Arg-Gly-Gly (RGG) repeats that spread throughout its C-terminal end (residues 407–575) [161]. These RGG repeats are needed for RNA binding of this protein [162]. TAF15 is structurally related to TLS/FUS and EWS proteins, and together these multifunctional proteins constitute the TET family [163]. Each of the TET family members is able to contribute a potent transcriptional activation domain to oncogenic fusion proteins originating due to the chromosomal translocation and often serving as the primary causes of their associated cancers [163]. In fact, in certain human chondrosarcomas, the N-terminal activation domain of TAF15 serves as an essential transforming domain in the fusion oncoproteins created by chromosomal translocation [164]. This N-terminal transactivation domain of TAF15 was shown to possess transactivation and oncogenic properties [165]. Figure 4E shows that TAF15 is another highly disordered protein according to the MobiDB-based analysis. There are eight AIBSs in TAF15, which are exclusively located within the N-terminal transactivation domain (residues 16–22, 28–84, 98–192, 204–218, 234–261, 297–321, 358–373, and 394–401). Therefore, although AIBSs cover 42.3% of the full-length TAF15, these disorder-based protein-protein interaction sites occupy 62.6% of its 401 residue-long activation domain.

2.4.2. Highly Connected Hubs Enriched in Disorder

The section below analyses the most connected human hubs that contain high levels of intrinsic disorder. Disordered hubs in this section are sorted by their interactability (number of interactors).

BAG family molecular chaperone regulator 3 (Supplementary Materials Table S1: UniProt ID O95817; 415 interactors; disorder content = 0.86; average disorder content of all (disorder enriched) interactors = 0.19 (0.85)). High binding promiscuity of this highly disordered BAG3 protein (see Figure 4F for the MobiDB-based result) can be defined by the fact that this 575 residue-long protein serves as a co-chaperone for HSP70 and HSC70 chaperone proteins [166]. BAG3 is a nucleotide-exchange factor (NEF), the major function of which is the promotion of the release of ADP from the HSP70 and HSC70 proteins required for triggering the release of client/substrate proteins [167]. BAG3 acts as a two-sided sword, promoting nucleotide release via binding to the nucleotide-binding domain (NBD) of HSPA8/HSC70, and mediating the substrate release via interaction with the substrate-binding domain (SBD) of HSPA8/HSC70 [166,167]. Human BAG3 contains two WW domains (residues 20–54 and 124–154, also known as rsp5 or WWP domains), which are shown to be engaged in interaction with proteins containing specific proline-motifs, [AP]-P-P-[AP]-Y and phosphorylated serine/threonine-proline sites, often mediating the assembly of multiprotein networks [168]. Among other specific features of the human BAG3 amino acid sequence are the presence of a short poly-serine region (residues 180–187) and another functional domain, BAG domain (residues 421–498), needed for interaction with HSP70 and HSC70 and positive or negative

regulation of these chaperone proteins [166]. According to the ANCHOR analysis, 73.6% of BAG3 sequence can be engaged in disordered protein-protein interactions. There are twelve AIBSs in human BAG3 (residues 1–17, 22–46, 81–141, 147–164, 173–314, 329–350, 362–371, 383–401, 420–439, 450–466, 490–537, and 552–575). Importantly, the known functional domains of BAG3 either overlap with AIBSs (e.g., WW-1 and WW-2) or include such disorder-based binding sites (e.g., BAG domain that contains AIBSs spanning residues 420–439 and 450–466).

CREB-binding protein (Supplementary Materials Table S1: UniProt ID Q92793; 262 interactors; disorder content = 0.63; average disorder content of all (disorder enriched) interactors = 0.34 (0.51)). CREBBP (or CBP) is a 2442 residue-long protein found both in the nucleus and cytoplasm with a wide spectrum of important biological functions. These functions range from histone acetyltransferase (HAT) activity [169], to acetylation of proliferating cell nuclear antigen (PCNA) [170] and many other non-histone proteins, such as general transcription factors [171], erythroid Krueppel-like factor (EKLF) [172], forkhead transcription factor FOXO1 [173] and many other proteins, to regulation of the cell cycle [174], to promiscuous interaction with multiple protein partners, including CREB [175], as well as hypoxia-inducible factor 1 α (HIF-1 α), NF- κ B, p53, and STAT proteins [176], among many others. Because of this multifunctionality and binding promiscuity, CREBBP and its paralogue p300 are considered as central nodes in eukaryotic transcriptional regulatory networks [177]. Human CREBBP is a large multidomain protein containing several ordered domains (TAZ1 (residues 352–432), KIX (residues 587–666), bromodomain BRD (residues 1094–1180), CH2 (residues 1192–1318), HAT (residues 1323–1700), ZZ (residues 1701–1742), TAZ2 (residues 1770–1843), and nuclear coactivator binding domain (NCBD, residues 2059–2115) accounting for ~950 of its residues, with the remaining ~1500 residues being intrinsically disordered [176] (see Figure 4G). The importance of intrinsic disorder for functions and interactions of CREBBP was systemized in a recent review [176]. The ability of CREBBP to be engaged in numerous protein-protein interactions is defined by the presence of 34 AIBSs (residues 1–7, 17–102, 109–140, 145–165, 171–221, 232–247, 254–283, 295–314, 353–365, 479–500, 515–578, 599–613, 628–635, 649–662, 673–683, 707–768, 776–879, 881–898, 902–1004, 1017–1059, 1084–1102, 1121–1127, 1537–1551, 1616–1639, 1844–1851, 1857–1878, 1911–1943, 1952–2028, 2032–2054, 2060–2077, 2084–2114, 2132–2200, 2217–1353, and 2371–2442) that involve 1304 (or 53.4%) residues. AIBSs are abundantly present in disordered linker regions connecting globular domains of this important protein. In fact, it seems that almost all residues in linker regions can be used for the protein-protein interactions. Furthermore, it was emphasized that even ordered domains of CREBBP show high levels of interactability serving as scaffolds that bind disordered transactivation domains of a wide variety of partners [176].

RNA-binding protein EWS (Supplementary Materials Table S1: UniProt ID Q01844; 227 interactors; disorder content = 0.81; average disorder content of all (disorder enriched) interactors = 0.27 (0.84)). EWS oncogene (also known as Ewing sarcoma breakpoint region 1 protein) is a 656 residue-long nuclear and cytoplasmic protein that can also be found in cell membrane [178]. Similar to TAF15, EWS is highly disordered (see Figure 4H; disorder content = 0.79 according to the MobiDB analysis) member of the TET protein family [163]. N-terminal half of this protein contains 31 approximate tandem repeats and is considered as EAD (Gln/Pro/Thr-rich) region. EWS also has a short IQ domain (residues 256–286), an RRM domain (residues 361–447), and three C-terminally located Arg/Gly/Pro-rich regions (residues 300–340, 454–513, and 559–640). Also, similar to TAF15, EWS often serves as an important part of the fusion oncoproteins created by chromosomal translocation [179–181]. High disorder levels in EWS and TAF15 are in line with the important notion that many oncogenic fusion proteins are enriched in intrinsic disorder, and that this high disorderedness enables these proteins to evade cellular surveillance mechanisms evolved for the elimination of misfolded proteins [182]. According to the ANCHOR analysis, human EWS contains 14 AIBSs (residues 89–98, 110–131, 156–185, 200–234, 242–293, 300–336, 343–350, 360–388, 396–403, 423–460, 465–473, 482–501, 512–559, and 570–644) that cover 64.2% of its sequence. Since almost the entire sequence of EWS (except to the 88 N-terminal

residues) is densely packed with these disorder-based binding sites, it is likely that they do play important roles in the promiscuity of this protein.

Cyclin-dependent kinase inhibitor 1 (Supplementary Materials Table S1: UniProt ID P38936; 206 interactors; disorder content = 0.63; average disorder content of all (disorder enriched) interactors = 0.20 (0.50)). CDK-interacting protein 1 or p21^{Waf1/Cip1/Sdi1} is a highly disordered (see Figure 4I; disorder content = 0.54 according to the MobiDB analysis) [183], 164 residue-long protein mediating G1/S-phase arrest through inhibition of cyclin-dependent kinases (Cdks), transcriptional regulation, modulation or inhibition of apoptosis, and playing an important role in controlling the p53-dependent cell growth and regulating the p53-mediated inhibition of cellular proliferation in response to DNA damage [102,105,106,184,185]. The prevalence of intrinsic disorder in p21 and another cyclin-dependent kinase (Cdk) regulator, p27, is well documented, and the importance of intrinsic disorder for functionality of these two proteins was systematically analyzed; see [102,105,106] for the topical reviews. Human p21 contains five AIBSs (residues 36–41, 68–79, 100–105, 109–123, and 145–164) accounting for 36.0% of its amino acid sequence. The very C-terminal AIBS coincides with the PIP-box K+4 motif that is known to mediate the p21 interaction with PCNA [186].

Mediator of DNA damage checkpoint protein 1 (Supplementary Materials Table S1: UniProt ID Q14676; 176 interactors; disorder content = 0.74; average disorder content of all (disorder enriched) interactors = 0.24 (0.89)). MDC1 (also known as nuclear factor with BRCT domains protein 1, NFBFD1 or KIAA0170) is a 2089 residue-long, highly disordered (see Figure 4J; disorder content = 0.63 according to the MobiDB analysis) nuclear protein that serves as a key regulator of the DNA damage response [187]. As it follows from its name, MDC1 plays a crucial role in controlling the checkpoint mediated cell cycle arrest in response to DNA damage within both the S phase and G2/M phases of the cell cycle [188–194]. MDC1 contains an N-terminal FHA (forkhead-associated) domain (residues 54–105) and a tandem repeat of BRCT (breast cancer susceptibility gene-1 C terminus) domains (residues 1892–1970 and 1991–2082). The tandemly repeated BRCT domains of MDC1 are required for localization of this protein to chromatin [193]. There are several functional regions in this protein needed for MDC1 interaction with the CHK2 (residues 1–150), the MRN complex (residues 1–220), or the PRKDC complex (residues 1148–1610). Furthermore, there are two nuclear localization signals (NLS1 and NLS2, residues 145–568, and NLS2, residues 1698–2089) and a proline-rich region (residues 1034–1469). Among all illustrative proteins considered in this section, human MDC1 is predicted to contain the largest number of AIBSs (43) that cover 1270 of its residues (or 60.8%). In fact, according to ANCHOR analysis, the DPB regions densely populate almost entire MDC1 sequence, with the noticeable exception for its N-terminal region (residues 1–200) that is predicted to be mostly ordered and to have only two short AIBSs (residues 1–8 and 32–37). On the other hand, the remaining 90% of this protein have AIBSs that range in length from 6 to 81 residues and are characterized by the mean length of 30 residues (residues 202–220, 235–265, 276–293, 314–332, 338–375, 384–403, 409–431, 434–445, 463–485, 503–522, 531–542, 556–577, 587–595, 604–618, 627–656, 686–760, 767–786, 796–824, 834–846, 862–868, 889–897, 907–946, 968–1045, 1051–1126, 1132–1174, 1182–1216, 1226–1255, 1260–1297, 1311–1341, 1352–1377, 1389–1421, 1430–1461, 1472–1507, 1518–1542, 1557–1615, 1628–1708, 1713–1747, 1754–1821, 1831–1841, 1893–1902, and 1909–1914).

2.4.3. Highly Connected Hubs Depleted in Disorder

The proteins considered below are very different from the highly disordered human hubs presented in two preceding sections. They are examples of highly connected ordered hubs with no putative disorder; the disorder content is predicted to be 0. At first glance, the presence of such highly connected but ordered hubs may seem a contradiction to the major motto of this article “disorder is crucial for binding promiscuity”. However, one of the first studies dedicated to the analysis of roles of intrinsic disorder in PPI networks pointed out that intrinsic disorder can contribute to protein binding promiscuity in three general ways: “First, intrinsic disorder can serve as the structural basis for hub protein promiscuity; secondly, intrinsically disordered proteins can bind to structured hub proteins; and thirdly, intrinsic disorder can provide flexible linkers between functional domains with the linkers

enabling mechanisms that facilitate binding diversity” [61]. We emphasize the second mechanism that provides an explanation for how structured hubs can be highly promiscuous by interacting with IDPs and IDPRs. To this end, Supplementary Materials Table S2 reveals that interactors of the highly connected structured hubs have substantial levels of disorder. The average computed over the average disorder content of interactors of the top 100 structured hubs equals 0.19. Moreover, the average disorder content for the interactors of the five protein ranges between 0.20 and 0.33 is described below.

Ubiquitin (Supplementary Materials Table S2: UniProt ID P0CG48; 8548 interactors; average disorder content of interactors = 0.20). Among proteins considered in this article, ubiquitin is characterized by the largest PPI network. This is not too surprising, since protein ubiquitination is one of the more common PTMs in eukaryotic cells, and, as a matter of fact, is present in bacteria and other prokaryotes, where target proteins are modified with ubiquitin-like proteins, such as recently discovered ubiquitin Bacterial (UBact) [195] and prokaryotic ubiquitin-like protein (Pup) [196], respectively. The biological importance of ubiquitin is reflected in an intriguing fact that this protein is encoded by four different genes, *UBA52*, *RPS27A*, *UBB*, and *UBC*. A single copy of ubiquitin fused to the ribosomal proteins L40 and S27a is encoded by *UBA52* and *RPS27A* genes, respectively, whereas *UBB* and *UBC* genes code for polyubiquitin precursors with exact head to tail repeats of three and nine identical ubiquitin sequences. Also, ubiquitin is one of the more abundant eukaryotic proteins.

It is known that mature ubiquitin (which is a 76 residue-long globular protein) can exist in unbound form, or be covalently bound to target proteins, which can be mono- and polyubiquitinated. Ubiquitination is intensively used in several biological processes, such as directing proteasomal protein degradation [197,198], regulation of budding of retroviral virions [199], modulation of the activity of transcription factors [200], control of the receptor endocytosis and lysosomal trafficking [201], mediation of the Arc-dependent synaptic plasticity [202], regulation of the repair of DNA double strand breaks (DSBs) by the ubiquitin-dependent molecular unfoldase/segregase p97 [203], regulation of various signaling pathways, such as insulin [204] and TGF- β signaling [205], as well as initiating immune response [206], and many others functions. In fact, it is close to impossible to provide an exhaustive list of biological activities regulated by ubiquitination, since as of October 24, 2017 there were 19,541 related entries in PubMed.

Although human ubiquitin is predicted to contain 0% of disordered residues (see Figure 4K), solution NMR analysis of this protein revealed that human ubiquitin is characterized by considerable conformational heterogeneity and can be described as a rather dynamic conformational ensemble (see Figure 5A) [207]. It was also pointed out that “the interior atoms of the protein are tightly packed in each individual conformation that contributes to the ensemble but their overall behaviour can be described as having a significant degree of liquid-like character” [207]. Curiously, although human ubiquitin has a unique structure with considerable conformational flexibility and heterogeneity, its prokaryotic homologue, Pup, is intrinsically disordered [208,209]. Finally, in relation to the subject of this study, it was indicated that the ubiquitination sites of target proteins are commonly located within IDPRs [76].

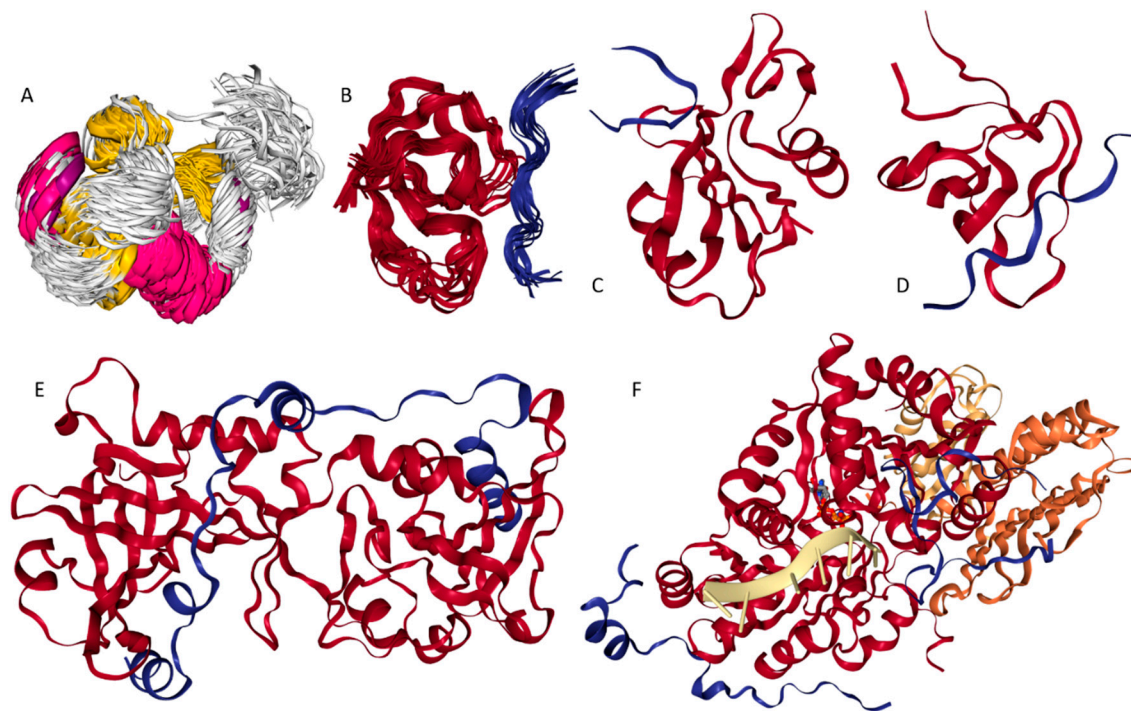


Figure 5. Structural characterization of highly connected ordered hubs. (A) Solution NMR structure of human ubiquitin (PDB ID: 1XQQ) [207]. (B) Solution NMR structure of a complex between the N-terminal SH3 domain of GRB2 (residues 1–56, red ribbons) and a peptide from SOS (blue ribbons) (PDB ID: 1AZE) [210]. (C) Minimized mean solution NMR structure of a complex between the SH2 domain of human GRB2 (residues 49–168, red ribbon) and a KPFY*VNVEF peptide (blue ribbon) (PDB ID: 1BMB) [211]. (D) Minimized mean solution NMR structure of a complex between the C-terminal SH3 domain of human GRB2 (residues 159–215, red ribbon) and a ligand peptide (blue ribbon) (PDB ID: 1IO6). (E) Crystal structure of a telomeric shelterin complex between the POT1 C-terminal domain (POT1C, residues 330–634, red ribbon) and POT1-binding region (residues 254–336, blue ribbon) of the adrenocortical dysplasia protein homolog (PDB ID: 5JUN7) [212]. (F) Crystal structure of a core EJC complex containing the complex of MAGOH (full length, dark orange ribbon), Y14 (residues 66–174, light orange ribbon), eIF4AIII (full length, red ribbon), Btz (the SELOR domain, residues 137–286, two blue ribbons), a non-hydrolyzable ATP analog (AMPPNP, bound to eIF4AIII), and U₁₅ RNA (yellow ribbon) (PDB ID: 2J0Q) [213].

Growth factor receptor-bound protein 2 (Supplementary Materials Table S2: UniProt ID P62993; 804 interactors; average disorder content of interactors = 0.24). SH2/SH3 adapter protein GRB2 is a 217 residue-long regulatory protein found in nucleus, Golgi apparatus, endosomes, and cytoplasm [214]. The ability of GRB2 to affect various biological processes and modulate multiple proteins defines its position as a critical link between the cell surface growth factor receptors and downstream cellular targets [215,216]. Furthermore, GRB2 is able to inhibit receptor tyrosine kinases (RTKs, such as hepatocyte growth factor receptor and platelet derived growth factor receptor) that have important biological functions serving as basic communication systems between the extracellular milieu and intracellular signaling pathways [217], as well as it affecting non-receptor tyrosine kinases [218]. GRB2 conducts its regulatory activities by binding to specific phosphotyrosine-containing and proline-rich sequence motifs of countless target proteins. Among numerous binding partners of GRB2 are tyrosine phosphorylated son of sevenless (SOS) [219,220], tyrosine phosphorylated EGF and PDGF receptors [221,222], tyrosine phosphorylated SIT1, IRS1, IRS4, SHC, and LNK [223–226], phosphorylated C-terminus of SH2B2 [227], as well as phosphorylated

LAX1, LAT, LAT2, and LIME1 [228–231]. GRB2 can also bind to NISCH, PTPNS1, REPS2 [232–234], and PIK3C2B [235].

Despite its remarkable binding promiscuity human GRB2 is characterized by a rather simple architecture, containing three functional domains, N-terminal SH3 domain (residues 1–58), SH2 domain (residues 60–152), and C-terminal SH3 domain (residues 156–217). Figure 5B–D shows NMR solution structures of these domains complexed with the corresponding target peptides. It is clearly seen that peptides bound to the functional domains of GRB2 are characterized by extended configuration, suggesting that they are mostly disordered in their unbound forms. This is in agreement with the known fact that SH3 and SH2 domains of human GRB2 prefer to bind to phosphotyrosine-containing and proline-rich sequence motifs, which are expected to be highly disordered. Therefore, ordered domains of GRB2 serve as an example of disorder-based many-to-one binding modes, where numerous disordered partners interact with the same ordered protein/domain.

Actin (Supplementary Materials Table S2: UniProt ID P60709; 263 interactors; average disorder content of interactors = 0.20). Actin, a 375 residue-long globular protein, is one of the most highly abundant proteins in every living cell. For example, in the muscle cells, actin concentration ranges from 230 to 960 μM [236]. Actin is known to possess unusual structural characteristics. Although its unique 3D-structure is stabilized by binding of one ATP molecule and one Ca^{2+} ion (or Mg^{2+} In Vivo), it is thermodynamically unstable and quasi-stationary [237,238], being formed in vivo via complex posttranslational folding processes that are assisted by Hsp 70, prefoldin, and chaperonin CCT and require the ATP energy expenditure. Among important functional features of actin is its ability to polymerize (to undergo G-F transition) and to interact with the other main muscle proteins, such as myosin, as well as with the regulatory proteins controlling muscle relaxation and contraction [239,240]. Furthermore, monomeric actin and actin-containing microfilaments are critically involved in numerous interactions with a wide spectrum of unrelated actin-binding proteins (ABPs) [241,242], many of which are characterized by high levels of intrinsic disorder [241].

Protection of telomeres protein 1 (Supplementary Materials Table S2: UniProt ID Q9NUX5; 200 interactors; average disorder content of interactors = 0.27). POT1 is a 634 residue-long nuclear protein that serves as an important constituent of the telomerase ribonucleoprotein (RNP) complex, also known as the shelterin complex (telosome), essential for the regulation and maintenance of telomere length and containing TERF1, TERF2, TIN2, TERF2IP, ACD, and POT1, as well as arrays of single stranded DNA TTAGGG repeats at the ends of human chromosomes [243,244]. POT1 is also found in the telomerase holoenzyme complex, which in addition to POT1 contains TERT, DKC1, WRAP53/TCAB1, NOP10, NHP2, GAR1, TEP1, EST1A, and a telomerase RNA template component (TERC) [245]. Furthermore, POT1 can homodimerize and homo-oligomerize. Human POT1 contains an N-terminal DNA-binding domain (residues 1–300) and a C-terminal protein interaction domain (residues 330–634). Figure 5E shows a crystal structure of a complex between the C-terminal domain of POT1 (POT1C, residues 330–634) and the POT1-binding region of the adrenocortical dysplasia protein homolog (residues 254–336) [212], which is also known as POT1 and TIN2-interacting protein, PIP1, PTOP, TINT1, and TPP1. It is apparent that POT1C forms a bilobal globular structure consisting of an OB-fold and a holiday junction resolvase domain, whereas a POT1-binding motif of TPP1 wraps around POT1C forming a set of loops and disjointed helices involved in extensive interactions with POT1C [212]. This structure of the POT1C-TPP1 complex provides an important illustration of how globular POT1C can bind at least some of its disordered partners.

Protein mago nashi homolog (Supplementary Materials Table S2: UniProt ID P61326; 190 interactors; average disorder content of interactors = 0.33). MAGOH is a 146 residue-long globular protein found in nucleus, nuclear speckles, and cytoplasm [246]. MAGOH serves as a crucial component of the exon junction complex (EJC) and is also found in the spliceosome C complex [247]. EJC, which plays a major role in posttranscriptional regulation of mRNA, is deposited in a sequence-independent manner at a fixed distance (20–24 nucleotides) upstream of the exon-exon junction of mRNAs and can be involved in the control of alternative splicing [248]. It was pointed out that EJC complex is characterized by a

highly dynamic nature, consisting of several core proteins (such as eIF4AIII, Barentsz [Btz], MAGOH, and Y14) and several peripherally associated proteins, such as factors involved in cytoplasmic mRNA surveillance (UPF3 and UPF2), mRNA export factors (REF/Aly, TAP-p15), and splicing coactivators (Srm160, Pinin, RNPS1) [249]. Figure 5F represents a crystal structure of the core EJC complex containing MAGOH (full length), Y14 (residues 66–174), eIF4AIII (full length), Btz (the SELOR domain, residues 137–286), a non-hydrolyzable ATP analog (AMPPNP) and a U₁₅ RNA (yellow ribbon) [213]. Similar to POT1C-TTP1 interaction, in the EJC structure, an extended Btz does not have a globular fold, its two separate stretches wrap around both domains of eIF4AIII and also contact MAGOH [213]. In this complex, MAGOH is characterized by a flat antiparallel β -sheet flanked on one side by two parallel α -helices engaged in interaction with the RNA binding domain (RBD) of Y14. Importantly, the β -sheet surface of MAGOH is entirely exposed to solvent and therefore can be engaged in additional PPIs [213].

3. Materials and Methods

3.1. Dataset and Annotation of Hubs

We collected a set of currently available PPIs for human proteins and used this information to annotate hub proteins and their partners/interactors. The raw data on the PPI were obtained from the *mentha* resource [63]. This resource integrates expert curated PPIs obtained from a comprehensive set of five primary databases: MINT [250], IntAct [251,252], DIP [253], MatrixDB [254], and BioGRID [255]. We extracted all records from taxon 9606 (*Homo sapiens*) for this analysis. These records include 154,395 interactions that cover 15,870 proteins. Next, we mapped these proteins into UniProt [256] to secure their reference sequences and relevant annotations. We were able to map 14,512 proteins that cover 151,461 interactions into the UniProt. The most relevant to our study characteristics of the PPI networks is the degree of the proteins (nodes in the network), which is defined as the number of partners of a given protein. Figure 6 compares the distribution of the degrees between the original set of 15,870 proteins and the set of 14,512 mapped proteins. Figure 6A reveals that the fractions of proteins with a given degree are very similar between the two datasets of proteins. The same observation is true in Figure 6B that compares the two sets of cumulative fractions of proteins with a given degree. This means that the loss of 8.5% of proteins due to mapping into the UniProt, which was necessary to collect the functional and structural annotations for these proteins, did not bias the main characteristics of the underlying network.

As has already been emphasized (see Introduction section), there are several definitions of hub proteins. We adopted the one recent definition of hubs as the top 20% of proteins with the highest degrees (more most connected proteins) [13]. This definition was also utilized in many other studies [257–259]. Figure 6B reveals that the top 20% most connected proteins have degrees ≥ 24 , which is also in close agreement with the hub definition by Aragues et al. [12]. Consequently, the human PPI network includes 2866 hubs, where 13,963 (96.22%) proteins interact with these hubs, 141,346 (93.32%) interactions include at least one hub protein, and 80,557 (53.19%) of these interactions are between hubs. Furthermore, we decided not to sub-divide hubs into subtypes (obligate/party vs. transient/date), since the quality of such categorization was disputed in several works [11,258,260,261]. Instead, we aimed to characterize the entire class of 2866 hub proteins in the context of their already established enrichment in the intrinsic disorder [10,47,48,59,61].

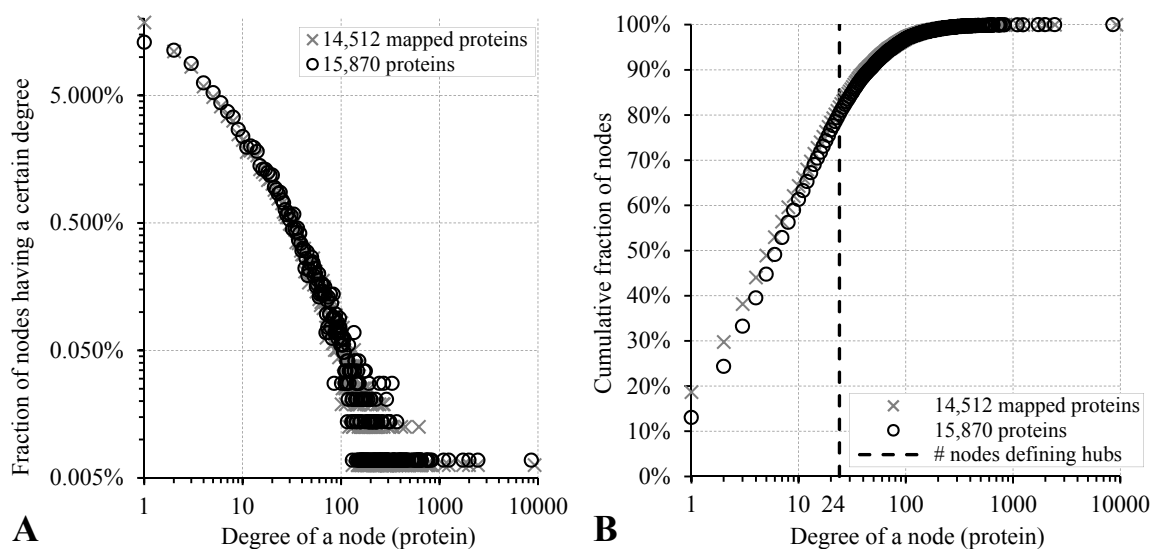


Figure 6. Connectivity of proteins in the human PPI network. Panel (A) summarizes the fraction of proteins (nodes in the PPI network) with a given number of PPI interactions (degree). Panel (B) gives the cumulative fraction of proteins having a degree less than the corresponding value on the x-axis. The circles (crosses) show the results for the original PPI network collected from mentha (the network where proteins were mapped to UniProt). The dashed vertical line in panel (B) shows the degree that demarcates hubs, which is defined as the 20% of the most connected nodes.

3.2. Functional and Structural Characterization of Proteins

We characterized functionally and structurally each protein in the human PPI network. This includes the annotations of the presence of intrinsic disorder, structural stability, evolutionary conservation, ligand binding, catalytic activity, degradation, proteolytic and subcellular targeting sites and annotation of a wide array of post-translational modifications (PTMs).

We comprehensively annotated the intrinsic disorder based on the presence of disordered regions, DPB regions, and MoRF regions. We predicted intrinsic disorder using a consensus of two popular methods: IUPred [144] and ESpritz [145]. We selected these methods based on their favorable predictive quality [262–264], the runtime which is sufficiently short to process the 15 thousand proteins, and complementary designs. To the latter point, the consensus includes two versions of IUPred that were designed to specialize in prediction of long disordered regions (30 or more consecutive residues) and short disordered segments (typically present in structured, globular proteins), and three versions of ESpritz that were designed for three types of annotations of disordered residues: using crystal structures, nuclear magnetic resonance structures, and annotations from the DisProt database [265,266]. We implemented the consensus using the majority vote, where three or more out of the five methods must predict the disorder for a given residue to be predicted as disordered. This was motivated by an observation that a consensus secures better predictive quality, when compared to the use of individual predictors [267,268]. The same consensus was utilized in a number of other studies [25,50,66,269–272]. Similar, consensus-based putative annotations of disorder can be also obtained from the MobiDB [141,142] and D²P² [273] databases. In agreement with these works and conventions in this field of research [274], we removed putative disordered segments with less than four consecutive residues.

We quantified disorder with the disorder content (fraction of disordered residues in a given protein sequence) and the normalized number of long IDPRs (30 or more consecutive residues). We normalized the latter by the unit of length of protein sequence (per 1000 amino acids) to alleviate bias due to the length of the protein chains, i.e., we computed the number of long IDPRs per 1000 amino acids. We considered the long IDPRs since they are recognized as biologically functional disordered

domains [24,275,276]. We predicted the DPB regions with the DisoRDPbind method [138,139]. DisoRDPbind was recently shown to outperform other similar methods [138], such as MoRFpred [277], DISOPRED3 [278] and ANCHOR [140]. It is also sufficiently runtime-efficient to process our large dataset. As with all IDPRs, we eliminated putative DPB regions that are shorter than four consecutive residues. We quantified the abundance of these regions with the content of DPB residues (fraction of the DPB residues in a given sequence) and with the normalized number of these regions (number of DPB regions per 1000 amino acids). We also annotated putative MoRF regions, defined as short (between 5 and 25 consecutive residues) DPB regions that undergo disorder-to-order transition upon binding to protein partners. We predicted these regions with a popular and accurate MoRFpred method [264,277]. The same as with the disorder, we assessed the abundance of MoRFs with the content of MoRF residues and with the normalized (per 1000 residues) number of MoRF regions.

On the other end of the protein structure spectrum, we estimated structural stability of these proteins with the number of putative disulfide linkage bonds obtained with the ModPred software [77]. We normalized the number of these bonds per unit of protein chain length (1000 residues). Furthermore, we also assessed other characteristics of proteins in our dataset including the length of their sequences (used as a proxy of protein size) and their evolutionary conservation. We measured the conservation for each residue in an input protein chain using relative entropy [279]:

$$\text{relative entropy} = \sum_i^N p_i \log_2 \left(\frac{p_i}{p_{ib}} \right) \quad (1)$$

where N is the number of amino acid types in the column representing an alignment position in the position specific scoring matrix (PSSM), p_i is the observed frequency of the i th amino acid type in the aligned column, and p_{ib} is the background frequency of the i th amino acid found in naturally occurring protein sequences; we computed the latter using the *nr* dataset. PSSM was collected by running the input protein against the *nr* database with the PSI-BLAST algorithm [280] using its default parameters. The relative entropy was empirically shown to be more sensitive to detect functional sites when compared to entropy (which does not utilize p_{ib} values) and several other conservation measures [279]. This is relevant since functional residues are among the key targets of our analysis. We quantified the conservation per protein and per each putative long IDPR by computing the average of the conservation scores over all residues in a given protein chain and over all residues in a given long disordered segment, respectively.

Next, we annotated and quantified abundance of eukaryotic linear motifs (ELMs) [281], also referred to as short linear motifs (SLiMs). These are short segments of adjacent amino acids that facilitate a wide array of cellular functions, such as protein-ligand interactions, catalytic activity, targeting of subcellular locations, proteolysis, and degradation [281–283]. Our focus on these motifs is motivated by an observation that ELMs are often found in IDPRs [284,285]. We extracted ELMs using the ELM resource [282,286,287]. This resource categorizes ELMs into six types: (1) proteolytic cleavage sites (CLV); (2) degradation sites (DEG); (3) posttranslational modification (PTM) sites (MOD); (4) docking sites (DOC) which are associated with the catalytic activity; (5) ligand binding sites (LIG) that mediate non-catalytic protein-ligand interactions; and (6) targeting sites (TRG) that facilitate the subcellular localization of proteins. We quantified abundance of these ELM categories by computing the normalized (per 1000 residues) number of corresponding ELMs sites.

Finally, we annotated and assessed the amount of specific PTM types in our protein set. This analysis extends beyond the analysis based on ELMs that focuses specially on the PTMs associated with SLiMs, by encompassing a broader set of PTMs and quantifying 21 specific PTM types. We annotated the PTMs with the ModPred software [77] that is tuned for eukaryotic organisms. The considered PTM types include acetylation, ADP-ribosylation, amidation, carboxylation, C-linked glycosylation, farnesylation, geranylgeranylation, GPI-anchor amidation, hydroxylation, methylation, myristoylation, N-linked glycosylation, N-terminal acetylation, O-linked glycosylation, palmitoylation, phosphorylation, pupylation, pyrrolidone carboxylic acid, sulfation, sumoylation, and ubiquitination.

We computed the overall normalized number of PTMs (across all PTM types) per unit of protein sequence length (1000 residues) where a residue might be counted multiple times if it is annotated to have multiple types of PTMs; the normalized number of residues that have PTMs per 1000 residues where a residue with multiple PTMs is counted just once; and the normalized number of residues with a specific PTM type per 1000 residues.

3.3. Annotation of Disease-Linked Proteins

We used the UniProt resource to annotate all diseases that each of the considered 14,512 human proteins is linked with. We mapped the accession numbers of each human protein to the corresponding list of 4296 disease accession numbers. The number of proteins associated with a given disease ranges between a single protein, which occurs for 4085 diseases, and twenty proteins for the non-insulin-dependent diabetes mellitus (accession: DI-02060). Only seven diseases are associated with ten or more proteins. Thus, to ensure sufficient statistical power, we clustered the diseases into broader disease classes. We used the National Library of Medicine's Medical Subject Headings (MeSH) vocabulary thesaurus to group the 4296 diseases into 26 major classes defined at the second level of the MeSH's hierarchical structure. The largest number of 4003 proteins are coupled with the MeSH category C16 "congenital, hereditary, and neonatal diseases and abnormalities", and 18 other classes that we consider in this study have at least 100 associated proteins each. The remaining eight classes include diseases with either a low protein count or that are not related to the human proteins, such as animal and parasitic diseases.

3.4. Statistical Analysis

We analyzed whether the considered functional, structural, and disease-association characteristics differ between proteins involved in the hub PPIs (PPIs that involve hubs) that have higher than expected disorder content when compared to proteins associated with the remainder of the hub PPIs. We performed this analysis for both hubs and their interactors. First, we derived the two sets of hubs PPIs; we describe this in Section 3.1. Next, we measured the magnitude and statistical significance of the differences of the considered characteristics between the two sets of proteins. To accomplish that we selected at random 20% of proteins from each of the two sets, quantified a given characteristic for each protein set, and repeated this ten times. This ensures that the measured ten times quantity covers a variety of diverse subpopulations of proteins from a given set. We represented each set of ten measurements with a mean value and we quantified the magnitude of the difference between the two sets of proteins by the relative difference (*Rdiff*) between the two corresponding means defined as:

$$Rdiff = \frac{mean_{disorder\ enriched\ hub\ PPIs} - mean_{remaining\ hub\ PPIs}}{\min\{mean_{disorder\ enriched\ hub\ PPIs}, mean_{remaining\ hub\ PPIs}\}} \times 100\% \quad (2)$$

Positive *Rdiff* values correspond to functional/structural characteristics that are enriched in the hubs (hub interactors) that have higher than expected disorder content in their hub PPIs while negative value are for the functional/structural characteristics that are enriched for the remaining hubs (hub interactors). *Rdiff* quantifies the percentage of the increase or decrease, e.g., *Rdiff* = 100% means that the corresponding characteristic is 100% larger in hubs (hub interactors) that have higher than expected disorder content in their hub PPIs compared to the remaining hubs (hub interactors). We also evaluated the statistical significance of the differences between the two sets of ten values. We used the *t*-test if the measurements were normal (we tested normality with the Anderson-Darling test at *p*-value = 0.05); otherwise we used the Wilcoxon test to generate the corresponding *p*-values. We assumed that a given difference is statistically significant if the *p*-value < 0.001. Moreover, we assumed that a given functional/structural characteristic is significantly different between the hubs (hub interactors) that have higher than expected disorder content in their hub PPIs compared to the remaining hubs (hub interactors) if the *p*-value < 0.001 and |*Rdiff*| is larger than 20%.

4. Conclusions

In this study, we analyzed the distribution of intrinsic disorder in the context of hub PPIs for both hubs and their interactors. This type of analysis is different from that of previous studies which were focused solely on the commonness of disorder in the hubs [10,47–49,59]. We focused our analysis on hub PPIs pertaining to proteins that have a higher than expected amount of disorder. We quantified and compared a wide range of relevant characteristics of these disorder-enriched hubs and their interactors. The characteristics include structural stability, evolutionary conservation, presence of twenty one types of PTMs, and several categories of functional sites associated with protein binding, catalysis, non-catalytic ligand binding, degradation, proteolytic processing, and subcellular targeting. Our analysis revealed that the disorder-enriched hub PPIs (including both hubs and their interactors) are significantly enriched in the DPB regions and long IDPRs. It also include high levels of catalytic sites, targeting sites, and PTM sites, in particular hydroxylation, phosphorylation, glycosylation, carboxylation, sulfation, and methylation. We also investigated whether disease-associated proteins are common among the proteins involved in the disorder-enriched hub PPIs. Our empirical analysis showed that the hubs registered significant enrichment for 11 out of the considered 18 classes of diseases. They include cancers and diseases of the stomatognathic, endocrine, digestive, respiratory, female urogenital, nervous, and musculoskeletal systems.

Besides the above analyses that functionally and structurally characterize the overall set of disorder-enriched human hubs PPIs, we also illustrated how specific highly disordered and highly ordered human hubs utilize intrinsic disorder for their promiscuous interactivity. We provided the description of functional and structural characteristics of the five most disordered human hubs, five highly disordered human hubs with the largest number of interactors; and five ordered hubs with the most number of interactors. In agreement with previous studies [61], these analyses revealed that mechanisms by which intrinsic disorder contributes to the binding promiscuity of hub proteins can be grouped in three major categories. The first is where intrinsically disordered hubs interact with multiple partners, both ordered and intrinsically disordered. The second is where the ordered hubs interact with disordered partners that might wrap around hubs. The final category involves hybrid hubs that contain functional domains connected by intrinsically disordered or flexible linkers that facilitate binding diversity.

Supplementary Materials: Supplementary materials can be found at www.mdpi.com/1422-0067/18/12/2761/s1.

Acknowledgments: This research was supported in part by the Qimonda Endowment to and the National Science Foundation grant 1617369 to Lukasz Kurgan.

Author Contributions: Lukasz Kurgan and Vladimir N. Uversky conceived the project; Lukasz Kurgan, Gang Hu and Zhonghua Wu designed the experiments; Gang Hu and Zhonghua Wu performed the experiments; Gang Hu, Zhonghua Wu, Lukasz Kurgan and Vladimir N. Uversky analyzed the data; Lukasz Kurgan and Vladimir N. Uversky wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. De Las Rivas, J.; Fontanillo, C. Protein-protein interactions essentials: Key concepts to building and analyzing interactome networks. *PLoS Comput. Biol.* **2010**, *6*, e1000807. [[CrossRef](#)] [[PubMed](#)]
2. Cusick, M.E.; Klitgord, N.; Vidal, M.; Hill, D.E. Interactome: Gateway into systems biology. *Hum. Mol. Genet.* **2005**, *14*, R171–R181. [[CrossRef](#)] [[PubMed](#)]
3. Erdős, P.; Rényi, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **1960**, *5*, 17–61.
4. Barabasi, A.L.; Bonabeau, E. Scale-free networks. *Sci. Am.* **2003**, *288*, 60–69. [[CrossRef](#)] [[PubMed](#)]
5. Goh, K.I.; Oh, E.; Jeong, H.; Kahng, B.; Kim, D. Classification of scale-free networks. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12583–12588. [[CrossRef](#)] [[PubMed](#)]
6. Barabasi, A.L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512. [[PubMed](#)]
7. Watts, D.J.; Strogatz, S.H. Collective dynamics of 'small-world' networks. *Nature* **1998**, *393*, 440–442. [[CrossRef](#)] [[PubMed](#)]

8. Wuchty, S. Scale-free behavior in protein domain networks. *Mol. Biol. Evol.* **2001**, *18*, 1694–1702. [[CrossRef](#)] [[PubMed](#)]
9. Han, J.D.; Bertin, N.; Hao, T.; Goldberg, D.S.; Berriz, G.F.; Zhang, L.V.; Dupuy, D.; Walhout, A.J.; Cusick, M.E.; Roth, F.P.; et al. Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* **2004**, *430*, 88–93. [[CrossRef](#)] [[PubMed](#)]
10. Haynes, C.; Oldfield, C.J.; Ji, F.; Klitgord, N.; Cusick, M.E.; Radivojac, P.; Uversky, V.N.; Vidal, M.; Iakoucheva, L.M. Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS Comput. Biol.* **2006**, *2*, 890–901. [[CrossRef](#)] [[PubMed](#)]
11. Batada, N.N.; Reguly, T.; Breitkreutz, A.; Boucher, L.; Breitkreutz, B.J.; Hurst, L.D.; Tyers, M. Stratus not altocumulus: A new view of the yeast protein interaction network. *PLoS Biol.* **2006**, *4*, e317. [[CrossRef](#)] [[PubMed](#)]
12. Aragues, R.; Sali, A.; Bonet, J.; Marti-Renom, M.A.; Oliva, B. Characterization of protein hubs by inferring interacting motifs from protein interactions. *PLoS Comput. Biol.* **2007**, *3*, 1761–1771. [[CrossRef](#)] [[PubMed](#)]
13. Jin, G.; Zhang, S.; Zhang, X.S.; Chen, L. Hubs with network motifs organize modularity dynamically in the protein-protein interaction network of yeast. *PLoS ONE* **2007**, *2*, e1207. [[CrossRef](#)] [[PubMed](#)]
14. Hartwell, L.H.; Hopfield, J.J.; Leibler, S.; Murray, A.W. From molecular to modular cell biology. *Nature* **1999**, *402*, C47–C52. [[CrossRef](#)] [[PubMed](#)]
15. Dunker, A.K.; Garner, E.; Guillot, S.; Romero, P.; Albrecht, K.; Hart, J.; Obradovic, Z.; Kissinger, C.; Villafranca, J.E. Protein disorder and the evolution of molecular recognition: Theory, predictions and observations. *Pac. Symp. Biocomput.* **1998**, *3*, 473–484.
16. Wright, P.E.; Dyson, H.J. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **1999**, *293*, 321–331. [[CrossRef](#)] [[PubMed](#)]
17. Uversky, V.N.; Gillespie, J.R.; Fink, A.L. Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins* **2000**, *41*, 415–427. [[CrossRef](#)]
18. Dunker, A.K.; Lawson, J.D.; Brown, C.J.; Williams, R.M.; Romero, P.; Oh, J.S.; Oldfield, C.J.; Campen, A.M.; Ratliff, C.M.; Hipps, K.W.; et al. Intrinsically disordered protein. *J. Mol. Graph. Model.* **2001**, *19*, 26–59. [[CrossRef](#)]
19. Tompa, P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* **2002**, *27*, 527–533. [[CrossRef](#)]
20. Daughdrill, G.W.; Pielak, G.J.; Uversky, V.N.; Cortese, M.S.; Dunker, A.K. Natively disordered proteins. In *Handbook of Protein Folding*; Buchner, J., Kiefhaber, T., Eds.; Wiley-VCH: Weinheim, Germany; Verlag GmbH & Co. KGaA: Weinheim, Germany, 2005; pp. 271–353.
21. Uversky, V.N.; Dunker, A.K. Understanding protein non-folding. *Biochim. Biophys. Acta* **2010**, *1804*, 1231–1264. [[CrossRef](#)] [[PubMed](#)]
22. Dunker, A.K.; Obradovic, Z.; Romero, P.; Garner, E.C.; Brown, C.J. Intrinsic protein disorder in complete genomes. *Genome Inform.* **2000**, *11*, 161–171.
23. Uversky, V.N. The mysterious unfoldome: Structureless, underappreciated, yet vital part of any given proteome. *J. Biomed. Biotechnol.* **2010**, *2010*, 568068. [[CrossRef](#)] [[PubMed](#)]
24. Ward, J.J.; Sodhi, J.S.; McGuffin, L.J.; Buxton, B.F.; Jones, D.T. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **2004**, *337*, 635–645. [[CrossRef](#)] [[PubMed](#)]
25. Peng, Z.; Yan, J.; Fan, X.; Mizianty, M.J.; Xue, B.; Wang, K.; Hu, G.; Uversky, V.N.; Kurgan, L. Exceptionally abundant exceptions: Comprehensive characterization of intrinsic disorder in all domains of life. *Cell. Mol. Life Sci.* **2015**, *72*, 137–151. [[CrossRef](#)] [[PubMed](#)]
26. Peng, Z.; Mizianty, M.J.; Kurgan, L. Genome-scale prediction of proteins with long intrinsically disordered regions. *Proteins* **2014**, *82*, 145–158. [[CrossRef](#)] [[PubMed](#)]
27. Radivojac, P.; Iakoucheva, L.M.; Oldfield, C.J.; Obradovic, Z.; Uversky, V.N.; Dunker, A.K. Intrinsic disorder and functional proteomics. *Biophys. J.* **2007**, *92*, 1439–1456. [[CrossRef](#)] [[PubMed](#)]
28. Galea, C.A.; Wang, Y.; Sivakolundu, S.G.; Kriwacki, R.W. Regulation of cell division by intrinsically unstructured proteins: Intrinsic flexibility, modularity, and signaling conduits. *Biochemistry* **2008**, *47*, 7598–7609. [[CrossRef](#)] [[PubMed](#)]
29. Buljan, M.; Chalancon, G.; Dunker, A.K.; Bateman, A.; Balaji, S.; Fuxreiter, M.; Babu, M.M. Alternative splicing of intrinsically disordered regions and rewiring of protein interactions. *Curr. Opin. Struct. Biol.* **2013**, *23*, 443–450. [[CrossRef](#)] [[PubMed](#)]
30. Uversky, V.N. Intrinsic disorder in proteins associated with neurodegenerative diseases. *Front. Biosci. (Landmark Ed.)* **2009**, *14*, 5188–5238. [[CrossRef](#)] [[PubMed](#)]

31. Uversky, V.N.; Oldfield, C.J.; Midic, U.; Xie, H.; Xue, B.; Vucetic, S.; Iakoucheva, L.M.; Obradovic, Z.; Dunker, A.K. Unfoldomics of human diseases: Linking protein intrinsic disorder with diseases. *BMC Genom.* **2009**, *10* (Suppl. 1), S7. [[CrossRef](#)] [[PubMed](#)]
32. Uversky, V.N. Targeting intrinsically disordered proteins in neurodegenerative and protein dysfunction diseases: Another illustration of the d(2) concept. *Expert Rev. Proteom.* **2010**, *7*, 543–564. [[CrossRef](#)] [[PubMed](#)]
33. Uversky, V.N. Intrinsically disordered proteins and novel strategies for drug discovery. *Expert Opin. Drug Discov.* **2012**, *7*, 475–488. [[CrossRef](#)] [[PubMed](#)]
34. Uversky, V.N. Wrecked regulation of intrinsically disordered proteins in diseases: Pathogenicity of deregulated regulators. *Front. Mol. Biosci.* **2014**, *1*, 6. [[CrossRef](#)] [[PubMed](#)]
35. Uversky, V.N.; Oldfield, C.J.; Dunker, A.K. Intrinsically disordered proteins in human diseases: Introducing the d2 concept. *Annu. Rev. Biophys.* **2008**, *37*, 215–246. [[CrossRef](#)] [[PubMed](#)]
36. Uversky, V.N.; Dave, V.; Iakoucheva, L.M.; Malaney, P.; Metallo, S.J.; Pathak, R.R.; Joerger, A.C. Pathological unfoldomics of uncontrolled chaos: Intrinsically disordered proteins and human diseases. *Chem. Rev.* **2014**, *114*, 6844–6879. [[CrossRef](#)] [[PubMed](#)]
37. Uversky, V.N.; Oldfield, C.J.; Dunker, A.K. Showing your id: Intrinsic disorder as an id for recognition, regulation and cell signaling. *J. Mol. Recognit.* **2005**, *18*, 343–384. [[CrossRef](#)] [[PubMed](#)]
38. Hasty, J.; Collins, J.J. Protein interactions. Unspinning the web. *Nature* **2001**, *411*, 30–31. [[CrossRef](#)] [[PubMed](#)]
39. Fischer, E. Einfluss der configuration auf die wirkung derenzyme. *Ber. Dtsch. Chem. Ges.* **1894**, *27*, 2985–2993. [[CrossRef](#)]
40. Koshland, D.E., Jr.; Ray, W.J., Jr.; Erwin, M.J. Protein structure and enzyme action. *Fed. Proc.* **1958**, *17*, 1145–1150. [[PubMed](#)]
41. Landsteiner, K. *The Specificity of Serological Reactions*; Courier Dover Publications: Mineola, NY, USA, 1936.
42. Pauling, L. A theory of the structure and process of formation of antibodies. *J. Am. Chem. Soc.* **1940**, *62*, 2643–2657. [[CrossRef](#)]
43. Karush, F. Heterogeneity of the binding sites of bovine serum albumin. *J. Am. Chem. Soc.* **1950**, *72*, 2705–2713. [[CrossRef](#)]
44. Meador, W.E.; Means, A.R.; Quijcho, F.A. Modulation of calmodulin plasticity in molecular recognition on the basis of X-ray structures. *Science* **1993**, *262*, 1718–1721. [[CrossRef](#)] [[PubMed](#)]
45. Kriwacki, R.W.; Hengst, L.; Tennant, L.; Reed, S.I.; Wright, P.E. Structural studies of p21^{Waf1/Cip1/Sdi1} in the free and cdk2-bound state: Conformational disorder mediates binding diversity. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 11504–11509. [[CrossRef](#)] [[PubMed](#)]
46. Uversky, V.N. A protein-chameleon: Conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders. *J. Biomol. Struct. Dyn.* **2003**, *21*, 211–234. [[CrossRef](#)] [[PubMed](#)]
47. 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*, 2041–2045. [[CrossRef](#)] [[PubMed](#)]
48. Dosztanyi, Z.; Chen, J.; Dunker, A.K.; Simon, I.; Tompa, P. Disorder and sequence repeats in hub proteins and their implications for network evolution. *J. Proteome Res.* **2006**, *5*, 2985–2995. [[CrossRef](#)] [[PubMed](#)]
49. Singh, G.P.; Ganapathi, M.; Sandhu, K.S.; Dash, D. Intrinsic unstructuredness and abundance of pest motifs in eukaryotic proteomes. *Proteins* **2006**, *62*, 309–315. [[CrossRef](#)] [[PubMed](#)]
50. Meng, F.; Na, I.; Kurgan, L.; Uversky, V.N. Compartmentalization and functionality of nuclear disorder: Intrinsic disorder and protein-protein interactions in intra-nuclear compartments. *Int. J. Mol. Sci.* **2016**, *17*, 24. [[CrossRef](#)] [[PubMed](#)]
51. Marinissen, M.J.; Gutkind, J.S. Scaffold proteins dictate rho gtpase-signaling specificity. *Trends Biochem. Sci.* **2005**, *30*, 423–426. [[CrossRef](#)] [[PubMed](#)]
52. Jaffe, A.B.; Aspenstrom, P.; Hall, A. Human cnk1 acts as a scaffold protein, linking rho and ras signal transduction pathways. *Mol. Cell. Biol.* **2004**, *24*, 1736–1746. [[CrossRef](#)] [[PubMed](#)]
53. Jaffe, A.B.; Hall, A. Rho gtpases: Biochemistry and biology. *Annu. Rev. Cell Dev. Biol.* **2005**, *21*, 247–269. [[CrossRef](#)] [[PubMed](#)]
54. Hohenstein, P.; Giles, R.H. Brca1: A scaffold for p53 response? *Trends Genet.* **2003**, *19*, 489–494. [[CrossRef](#)]
55. Rui, Y.; Xu, Z.; Lin, S.; Li, Q.; Rui, H.; Luo, W.; Zhou, H.M.; Cheung, P.Y.; Wu, Z.; Ye, Z.; et al. Axin stimulates p53 functions by activation of hipk2 kinase through multimeric complex formation. *EMBO J.* **2004**, *23*, 4583–4594. [[CrossRef](#)] [[PubMed](#)]

56. Salahshor, S.; Woodgett, J.R. The links between axin and carcinogenesis. *J. Clin. Pathol.* **2005**, *58*, 225–236. [[CrossRef](#)] [[PubMed](#)]
57. Wong, W.; Scott, J.D. Akap signalling complexes: Focal points in space and time. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 959–970. [[CrossRef](#)] [[PubMed](#)]
58. Carpousis, A.J. The RNA degradosome of *Escherichia coli*: A multiprotein mRNA-degrading machine assembled on RNase E. *Annu. Rev. Microbiol.* **2007**, *61*, 71–87. [[CrossRef](#)] [[PubMed](#)]
59. Ekman, D.; Light, S.; Bjorklund, A.K.; Elofsson, A. What properties characterize the hub proteins of the protein-protein interaction network of *saccharomyces cerevisiae*? *Genome Biol.* **2006**, *7*, R45. [[CrossRef](#)] [[PubMed](#)]
60. Oldfield, C.J.; Meng, J.; Yang, J.Y.; Yang, M.Q.; Uversky, V.N.; Dunker, A.K. Flexible nets: Disorder and induced fit in the associations of p53 and 14-3-3 with their partners. *BMC Genom.* **2008**, *9* (Suppl. 1), S1. [[CrossRef](#)] [[PubMed](#)]
61. Dunker, A.K.; Cortese, M.S.; Romero, P.; Iakoucheva, L.M.; Uversky, V.N. Flexible nets: The roles of intrinsic disorder in protein interaction networks. *FEBS J.* **2005**, *272*, 5129–5148. [[CrossRef](#)] [[PubMed](#)]
62. Luo, W.; Lin, S.C. Axin: A master scaffold for multiple signaling pathways. *Neurosignals* **2004**, *13*, 99–113. [[CrossRef](#)] [[PubMed](#)]
63. Calderone, A.; Castagnoli, L.; Cesareni, G. Mentha: A resource for browsing integrated protein-interaction networks. *Nat. Methods* **2013**, *10*, 690–691. [[CrossRef](#)] [[PubMed](#)]
64. Ota, M.; Gonja, H.; Koike, R.; Fukuchi, S. Multiple-localization and hub proteins. *PLoS ONE* **2016**, *11*, e0156455. [[CrossRef](#)] [[PubMed](#)]
65. Bosnjak, I.; Bojovic, V.; Segvic-Bubic, T.; Bielen, A. Occurrence of protein disulfide bonds in different domains of life: A comparison of proteins from the protein data bank. *Protein Eng. Des. Sel.* **2014**, *27*, 65–72. [[CrossRef](#)] [[PubMed](#)]
66. Wang, C.; Uversky, V.N.; Kurgan, L. Disordered nucleome: Abundance of intrinsic disorder in the DNA- and RNA-binding proteins in 1121 species from Eukaryota, Bacteria and Archaea. *Proteomics* **2016**, *16*, 1486–1498. [[CrossRef](#)] [[PubMed](#)]
67. Nguyen Ba, A.N.; Yeh, B.J.; van Dyk, D.; Davidson, A.R.; Andrews, B.J.; Weiss, E.L.; Moses, A.M. Proteome-wide discovery of evolutionary conserved sequences in disordered regions. *Sci. Signal.* **2012**, *5*, rs1. [[CrossRef](#)] [[PubMed](#)]
68. Walsh, C.T.; Garneau-Tsodikova, S.; Gatto, G.J. Protein posttranslational modifications: The chemistry of proteome diversifications. *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 7342–7372. [[CrossRef](#)] [[PubMed](#)]
69. Deribe, Y.L.; Pawson, T.; Dikic, I. Post-translational modifications in signal integration. *Nat. Struct. Mol. Biol.* **2010**, *17*, 666–672. [[CrossRef](#)] [[PubMed](#)]
70. Mann, M.; Jensen, O.N. Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* **2003**, *21*, 255–261. [[CrossRef](#)] [[PubMed](#)]
71. Witze, E.S.; Old, W.M.; Resing, K.A.; Ahn, N.G. Mapping protein post-translational modifications with mass spectrometry. *Nat. Methods* **2007**, *4*, 798–806. [[CrossRef](#)] [[PubMed](#)]
72. Xie, H.; Vucetic, S.; Iakoucheva, L.M.; Oldfield, C.J.; Dunker, A.K.; Obradovic, Z.; Uversky, V.N. Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. *J. Proteome Res.* **2007**, *6*, 1917–1932. [[CrossRef](#)] [[PubMed](#)]
73. Dunker, A.K.; Brown, C.J.; Obradovic, Z. Identification and functions of usefully disordered proteins. *Adv. Protein Chem.* **2002**, *62*, 25–49. [[PubMed](#)]
74. Dunker, A.K.; Brown, C.J.; Lawson, J.D.; Iakoucheva, L.M.; Obradovic, Z. Intrinsic disorder and protein function. *Biochemistry* **2002**, *41*, 6573–6582. [[CrossRef](#)] [[PubMed](#)]
75. Iakoucheva, L.M.; Radivojac, P.; Brown, C.J.; O'Connor, T.R.; Sikes, J.G.; Obradovic, Z.; Dunker, A.K. The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **2004**, *32*, 1037–1049. [[CrossRef](#)] [[PubMed](#)]
76. Radivojac, P.; Vacic, V.; Haynes, C.; Cocklin, R.R.; Mohan, A.; Heyen, J.W.; Goebel, M.G.; Iakoucheva, L.M. Identification, analysis, and prediction of protein ubiquitination sites. *Proteins* **2010**, *78*, 365–380. [[CrossRef](#)] [[PubMed](#)]
77. Pejaver, V.; Hsu, W.L.; Xin, F.; Dunker, A.K.; Uversky, V.N.; Radivojac, P. The structural and functional signatures of proteins that undergo multiple events of post-translational modification. *Protein Sci.* **2014**, *23*, 1077–1093. [[CrossRef](#)] [[PubMed](#)]

78. Midic, U.; Oldfield, C.J.; Dunker, A.K.; Obradovic, Z.; Uversky, V.N. Protein disorder in the human diseaseome: Unfoldomics of human genetic diseases. *BMC Genom.* **2009**, *10* (Suppl. 1), S12. [[CrossRef](#)] [[PubMed](#)]
79. Midic, U.; Oldfield, C.J.; Dunker, A.K.; Obradovic, Z.; Uversky, V.N. Unfoldomics of human genetic diseases: Illustrative examples of ordered and intrinsically disordered members of the human diseaseome. *Protein Pept. Lett.* **2009**, *16*, 1533–1547. [[CrossRef](#)] [[PubMed](#)]
80. Xie, H.; Vucetic, S.; Iakoucheva, L.M.; Oldfield, C.J.; Dunker, A.K.; Uversky, V.N.; Obradovic, Z. Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. *J. Proteome Res.* **2007**, *6*, 1882–1898. [[CrossRef](#)] [[PubMed](#)]
81. Vucetic, S.; Xie, H.; Iakoucheva, L.M.; Oldfield, C.J.; Dunker, A.K.; Obradovic, Z.; Uversky, V.N. Functional anthology of intrinsic disorder. 2. Cellular components, domains, technical terms, developmental processes, and coding sequence diversities correlated with long disordered regions. *J. Proteome Res.* **2007**, *6*, 1899–1916. [[CrossRef](#)] [[PubMed](#)]
82. Iakoucheva, L.M.; Brown, C.J.; Lawson, J.D.; Obradovic, Z.; Dunker, A.K. Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* **2002**, *323*, 573–584. [[CrossRef](#)]
83. Baker, J.M.; Hudson, R.P.; Kanelis, V.; Choy, W.Y.; Thibodeau, P.H.; Thomas, P.J.; Forman-Kay, J.D. Cfr regulatory region interacts with nbd1 predominantly via multiple transient helices. *Nat. Struct. Mol. Biol.* **2007**, *14*, 738–745. [[CrossRef](#)] [[PubMed](#)]
84. Wells, M.; Tidow, H.; Rutherford, T.J.; Markwick, P.; Jensen, M.R.; Mylonas, E.; Svergun, D.I.; Blackledge, M.; Fersht, A.R. Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5762–5767. [[CrossRef](#)] [[PubMed](#)]
85. Mark, W.Y.; Liao, J.C.; Lu, Y.; Ayed, A.; Laister, R.; Szymczynna, B.; Chakrabartty, A.; Arrowsmith, C.H. Characterization of segments from the central region of brca1: An intrinsically disordered scaffold for multiple protein-protein and protein-DNA interactions? *J. Mol. Biol.* **2005**, *345*, 275–287. [[CrossRef](#)] [[PubMed](#)]
86. Ng, K.P.; Potikyan, G.; Savene, R.O.; Denny, C.T.; Uversky, V.N.; Lee, K.A. Multiple aromatic side chains within a disordered structure are critical for transcription and transforming activity of EWS family oncoproteins. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 479–484. [[CrossRef](#)] [[PubMed](#)]
87. Uversky, V.N.; Roman, A.; Oldfield, C.J.; Dunker, A.K. Protein intrinsic disorder and human papillomaviruses: Increased amount of disorder in E6 and E7 oncoproteins from high risk HPVs. *J. Proteome Res.* **2006**, *5*, 1829–1842. [[CrossRef](#)] [[PubMed](#)]
88. Malaney, P.; Pathak, R.R.; Xue, B.; Uversky, V.N.; Dave, V. Intrinsic disorder in pten and its interactome confers structural plasticity and functional versatility. *Sci. Rep.* **2013**, *3*, 2035. [[CrossRef](#)] [[PubMed](#)]
89. Rajagopalan, K.; Mooney, S.M.; Parekh, N.; Getzenberg, R.H.; Kulkarni, P. A majority of the cancer/testis antigens are intrinsically disordered proteins. *J. Cell. Biochem.* **2011**, *112*, 3256–3267. [[CrossRef](#)] [[PubMed](#)]
90. Santamaria, N.; Alhothali, M.; Alfonso, M.H.; Breydo, L.; Uversky, V.N. Intrinsic disorder in proteins involved in amyotrophic lateral sclerosis. *Cell. Mol. Life Sci.* **2017**, *74*, 1297–1318. [[CrossRef](#)] [[PubMed](#)]
91. Ferreon, A.C.; Ferreon, J.C.; Wright, P.E.; Deniz, A.A. Modulation of allostery by protein intrinsic disorder. *Nature* **2013**, *498*, 390–394. [[CrossRef](#)] [[PubMed](#)]
92. Cozzetto, D.; Jones, D.T. The contribution of intrinsic disorder prediction to the elucidation of protein function. *Curr. Opin. Struct. Biol.* **2013**, *23*, 467–472. [[CrossRef](#)] [[PubMed](#)]
93. Dyson, H.J.; Wright, P.E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197–208. [[CrossRef](#)] [[PubMed](#)]
94. Collins, M.O.; Yu, L.; Campuzano, I.; Grant, S.G.; Choudhary, J.S. Phosphoproteomic analysis of the mouse brain cytosol reveals a predominance of protein phosphorylation in regions of intrinsic sequence disorder. *Mol. Cell. Proteom. MCP* **2008**, *7*, 1331–1348. [[CrossRef](#)] [[PubMed](#)]
95. Kurotani, A.; Tokmakov, A.A.; Kuroda, Y.; Fukami, Y.; Shinozaki, K.; Sakurai, T. Correlations between predicted protein disorder and post-translational modifications in plants. *Bioinformatics* **2014**, *30*, 1095–1103. [[CrossRef](#)] [[PubMed](#)]
96. Romero, P.R.; Zaidi, S.; Fang, Y.Y.; Uversky, V.N.; Radivojac, P.; Oldfield, C.J.; Cortese, M.S.; Sickmeier, M.; LeGall, T.; Obradovic, Z.; et al. Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8390–8395. [[CrossRef](#)] [[PubMed](#)]

97. Buljan, M.; Chalancon, G.; Eustermann, S.; Wagner, G.P.; Fuxreiter, M.; Bateman, A.; Babu, M.M. Tissue-specific splicing of disordered segments that embed binding motifs rewires protein interaction networks. *Mol. Cell* **2012**, *46*, 871–883. [[CrossRef](#)] [[PubMed](#)]
98. Uversky, V.N. Intrinsic disorder-based protein interactions and their modulators. *Curr. Pharm. Des.* **2013**, *19*, 4191–4213. [[CrossRef](#)] [[PubMed](#)]
99. Colak, R.; Kim, T.; Michaut, M.; Sun, M.; Irimia, M.; Bellay, J.; Myers, C.L.; Blencowe, B.J.; Kim, P.M. Distinct types of disorder in the human proteome: Functional implications for alternative splicing. *PLoS Comput. Biol.* **2013**, *9*, e1003030. [[CrossRef](#)] [[PubMed](#)]
100. Mitrea, D.M.; Kriwacki, R.W. Regulated unfolding of proteins in signaling. *FEBS Lett.* **2013**, *587*, 1081–1088. [[CrossRef](#)] [[PubMed](#)]
101. Follis, A.V.; Llambi, F.; Ou, L.; Baran, K.; Green, D.R.; Kriwacki, R.W. The DNA-binding domain mediates both nuclear and cytosolic functions of p53. *Nat. Struct. Mol. Biol.* **2014**, *21*, 535–543. [[CrossRef](#)] [[PubMed](#)]
102. Wang, Y.; Fisher, J.C.; Mathew, R.; Ou, L.; Otieno, S.; Sublet, J.; Xiao, L.; Chen, J.; Roussel, M.F.; Kriwacki, R.W. Intrinsic disorder mediates the diverse regulatory functions of the cdk inhibitor p21. *Nat. Chem. Biol.* **2011**, *7*, 214–221. [[CrossRef](#)] [[PubMed](#)]
103. Ou, L.; Waddell, M.B.; Kriwacki, R.W. Mechanism of cell cycle entry mediated by the intrinsically disordered protein p27^{kip1}. *ACS Chem. Biol.* **2012**, *7*, 678–682. [[CrossRef](#)] [[PubMed](#)]
104. Follis, A.V.; Galea, C.A.; Kriwacki, R.W. Intrinsic protein flexibility in regulation of cell proliferation: Advantages for signaling and opportunities for novel therapeutics. *Adv. Exp. Med. Biol.* **2012**, *725*, 27–49. [[PubMed](#)]
105. Yoon, M.K.; Mitrea, D.M.; Ou, L.; Kriwacki, R.W. Cell cycle regulation by the intrinsically disordered proteins p21 and p27. *Biochem. Soc. Trans.* **2012**, *40*, 981–988. [[CrossRef](#)] [[PubMed](#)]
106. Mitrea, D.M.; Yoon, M.K.; Ou, L.; Kriwacki, R.W. Disorder-function relationships for the cell cycle regulatory proteins p21 and p27. *Biol. Chem.* **2012**, *393*, 259–274. [[CrossRef](#)] [[PubMed](#)]
107. Moldoveanu, T.; Grace, C.R.; Llambi, F.; Nourse, A.; Fitzgerald, P.; Gehring, K.; Kriwacki, R.W.; Green, D.R. Bid-induced structural changes in bak promote apoptosis. *Nat. Struct. Mol. Biol.* **2013**, *20*, 589–597. [[CrossRef](#)] [[PubMed](#)]
108. Frye, J.J.; Brown, N.G.; Petzold, G.; Watson, E.R.; Grace, C.R.; Nourse, A.; Jarvis, M.A.; Kriwacki, R.W.; Peters, J.M.; Stark, H.; et al. Electron microscopy structure of human APC/C(CDH1)-EMI1 reveals multimodal mechanism of E3 ligase shutdown. *Nat. Struct. Mol. Biol.* **2013**, *20*, 827–835. [[CrossRef](#)] [[PubMed](#)]
109. Mei, Y.; Su, M.; Soni, G.; Salem, S.; Colbert, C.L.; Sinha, S.C. Intrinsically disordered regions in autophagy proteins. *Proteins* **2014**, *82*, 565–578. [[CrossRef](#)] [[PubMed](#)]
110. Chakrabortee, S.; Tripathi, R.; Watson, M.; Schierle, G.S.; Kurniawan, D.P.; Kaminski, C.F.; Wise, M.J.; Tunnacliffe, A. Intrinsically disordered proteins as molecular shields. *Mol. Biosyst.* **2012**, *8*, 210–219. [[CrossRef](#)] [[PubMed](#)]
111. De Jonge, N.; Garcia-Pino, A.; Buts, L.; Haesaerts, S.; Charlier, D.; Zangger, K.; Wyns, L.; De Greve, H.; Loris, R. Rejuvenation of cddb-poisoned gyrase by an intrinsically disordered protein domain. *Mol. Cell* **2009**, *35*, 154–163. [[CrossRef](#)] [[PubMed](#)]
112. Norholm, A.B.; Hendus-Altenburger, R.; Bjerre, G.; Kjaergaard, M.; Pedersen, S.F.; Kragelund, B.B. The intracellular distal tail of the Na⁺/H⁺ exchanger NHE1 is intrinsically disordered: Implications for NHE1 trafficking. *Biochemistry* **2011**, *50*, 3469–3480. [[CrossRef](#)] [[PubMed](#)]
113. Follis, A.V.; Chipuk, J.E.; Fisher, J.C.; Yun, M.K.; Grace, C.R.; Nourse, A.; Baran, K.; Ou, L.; Min, L.; White, S.W.; et al. Puma binding induces partial unfolding within bcl-xl to disrupt p53 binding and promote apoptosis. *Nat. Chem. Biol.* **2013**, *9*, 163–168. [[CrossRef](#)] [[PubMed](#)]
114. Rudiger, S.; Freund, S.M.; Veprintsev, D.B.; Fersht, A.R. Crinept-trosy nmr reveals p53 core domain bound in an unfolded form to the chaperone hsp90. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11085–11090. [[CrossRef](#)] [[PubMed](#)]
115. Martinez-Yamout, M.A.; Venkitakrishnan, R.P.; Preece, N.E.; Kroon, G.; Wright, P.E.; Dyson, H.J. Localization of sites of interaction between p23 and hsp90 in solution. *J. Biol. Chem.* **2006**, *281*, 14457–14464. [[CrossRef](#)] [[PubMed](#)]
116. Rodriguez, F.; Arsene-Ploetze, F.; Rist, W.; Rudiger, S.; Schneider-Mergener, J.; Mayer, M.P.; Bukau, B. Molecular basis for regulation of the heat shock transcription factor sigma32 by the dnak and dnaj chaperones. *Mol. Cell* **2008**, *32*, 347–358. [[CrossRef](#)] [[PubMed](#)]

117. Didenko, T.; Duarte, A.M.; Karagoz, G.E.; Rudiger, S.G. Hsp90 structure and function studied by NMR spectroscopy. *Biochim. Biophys. Acta* **2012**, *1823*, 636–647. [[CrossRef](#)] [[PubMed](#)]
118. Karagoz, G.E.; Duarte, A.M.; Akoury, E.; Ippel, H.; Biernat, J.; Moran Luengo, T.; Radli, M.; Didenko, T.; Nordhues, B.A.; Veprintsev, D.B.; et al. Hsp90-Tau complex reveals molecular basis for specificity in chaperone action. *Cell* **2014**, *156*, 963–974. [[CrossRef](#)] [[PubMed](#)]
119. Tsvetkov, P.; Reuven, N.; Shaul, Y. The nanny model for IDPs. *Nat. Chem. Biol.* **2009**, *5*, 778–781. [[CrossRef](#)] [[PubMed](#)]
120. Demarest, S.J.; Martinez-Yamout, M.; Chung, J.; Chen, H.; Xu, W.; Dyson, H.J.; Evans, R.M.; Wright, P.E. Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **2002**, *415*, 549–553. [[CrossRef](#)] [[PubMed](#)]
121. Ebert, M.O.; Bae, S.H.; Dyson, H.J.; Wright, P.E. NMR relaxation study of the complex formed between CBP and the activation domain of the nuclear hormone receptor coactivator ACTR. *Biochemistry* **2008**, *47*, 1299–1308. [[CrossRef](#)] [[PubMed](#)]
122. Chipuk, J.E.; Fisher, J.C.; Dillon, C.P.; Kriwacki, R.W.; Kuwana, T.; Green, D.R. Mechanism of apoptosis induction by inhibition of the anti-apoptotic BCL-2 proteins. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 20327–20332. [[CrossRef](#)] [[PubMed](#)]
123. Ferreon, J.C.; Lee, C.W.; Arai, M.; Martinez-Yamout, M.A.; Dyson, H.J.; Wright, P.E. Cooperative regulation of p53 by modulation of ternary complex formation with CBP/p300 and hdm2. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 6591–6596. [[CrossRef](#)] [[PubMed](#)]
124. Wojciak, J.M.; Martinez-Yamout, M.A.; Dyson, H.J.; Wright, P.E. Structural basis for recruitment of CBP/p300 coactivators by stat1 and stat2 transactivation domains. *EMBO J.* **2009**, *28*, 948–958. [[CrossRef](#)] [[PubMed](#)]
125. Ferreon, J.C.; Martinez-Yamout, M.A.; Dyson, H.J.; Wright, P.E. Structural basis for subversion of cellular control mechanisms by the adenoviral e1a oncoprotein. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13260–13265. [[CrossRef](#)] [[PubMed](#)]
126. Lee, C.W.; Martinez-Yamout, M.A.; Dyson, H.J.; Wright, P.E. Structure of the p53 transactivation domain in complex with the nuclear receptor coactivator binding domain of creb binding protein. *Biochemistry* **2010**, *49*, 9964–9971. [[CrossRef](#)] [[PubMed](#)]
127. Kostic, M.; Matt, T.; Martinez-Yamout, M.A.; Dyson, H.J.; Wright, P.E. Solution structure of the HDM2 C2H2C4 RING, a domain critical for ubiquitination of p53. *J. Mol. Biol.* **2006**, *363*, 433–450. [[CrossRef](#)] [[PubMed](#)]
128. Grimmler, M.; Wang, Y.; Mund, T.; Cilensek, Z.; Keidel, E.M.; Waddell, M.B.; Jakel, H.; Kullmann, M.; Kriwacki, R.W.; Hengst, L. Cdk-inhibitory activity and stability of p27^{Kip1} are directly regulated by oncogenic tyrosine kinases. *Cell* **2007**, *128*, 269–280. [[CrossRef](#)] [[PubMed](#)]
129. Zhan, J.; Easton, J.B.; Huang, S.; Mishra, A.; Xiao, L.; Lacy, E.R.; Kriwacki, R.W.; Houghton, P.J. Negative regulation of ASK1 by p21^{Cip1} involves a small domain that includes serine 98 that is phosphorylated by ASK1 in vivo. *Mol. Cell. Biol.* **2007**, *27*, 3530–3541. [[CrossRef](#)] [[PubMed](#)]
130. Galea, C.A.; Nourse, A.; Wang, Y.; Sivakolundu, S.G.; Heller, W.T.; Kriwacki, R.W. Role of intrinsic flexibility in signal transduction mediated by the cell cycle regulator, p27^{Kip1}. *J. Mol. Biol.* **2008**, *376*, 827–838. [[CrossRef](#)] [[PubMed](#)]
131. Mitrea, D.M.; Grace, C.R.; Buljan, M.; Yun, M.K.; Pytel, N.J.; Satumba, J.; Nourse, A.; Park, C.G.; Madan Babu, M.; White, S.W.; et al. Structural polymorphism in the N-terminal oligomerization domain of NPM1. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 4466–4471. [[CrossRef](#)] [[PubMed](#)]
132. Asher, G.; Tsvetkov, P.; Kahana, C.; Shaul, Y. A mechanism of ubiquitin-independent proteasomal degradation of the tumor suppressors p53 and p73. *Genes Dev.* **2005**, *19*, 316–321. [[CrossRef](#)] [[PubMed](#)]
133. Tsvetkov, P.; Asher, G.; Paz, A.; Reuven, N.; Sussman, J.L.; Silman, I.; Shaul, Y. Operational definition of intrinsically unstructured protein sequences based on susceptibility to the 20s proteasome. *Proteins* **2008**, *70*, 1357–1366. [[CrossRef](#)] [[PubMed](#)]
134. Tsvetkov, P.; Reuven, N.; Prives, C.; Shaul, Y. Susceptibility of p53 unstructured n terminus to 20 s proteasomal degradation programs the stress response. *J. Biol. Chem.* **2009**, *284*, 26234–26242. [[CrossRef](#)] [[PubMed](#)]
135. Wiggins, C.M.; Tsvetkov, P.; Johnson, M.; Joyce, C.L.; Lamb, C.A.; Bryant, N.J.; Komander, D.; Shaul, Y.; Cook, S.J. BIM(El), an intrinsically disordered protein, is degraded by 20S proteasomes in the absence of poly-ubiquitylation. *J. Cell Sci.* **2011**, *124*, 969–977. [[CrossRef](#)] [[PubMed](#)]

136. Suskiewicz, M.J.; Sussman, J.L.; Silman, I.; Shaul, Y. Context-dependent resistance to proteolysis of intrinsically disordered proteins. *Protein Sci.* **2011**, *20*, 1285–1297. [[CrossRef](#)] [[PubMed](#)]
137. Tsvetkov, P.; Myers, N.; Moscovitz, O.; Sharon, M.; Prilusky, J.; Shaul, Y. Thermo-resistant intrinsically disordered proteins are efficient 20s proteasome substrates. *Mol. Biosyst.* **2012**, *8*, 368–373. [[CrossRef](#)] [[PubMed](#)]
138. Peng, Z.; Kurgan, L. High-throughput prediction of RNA, DNA and protein binding regions mediated by intrinsic disorder. *Nucleic Acids Res.* **2015**, *43*, e121. [[CrossRef](#)] [[PubMed](#)]
139. Peng, Z.; Wang, C.; Uversky, V.N.; Kurgan, L. Prediction of disordered RNA, DNA, and protein binding regions using disordpbind. *Methods Mol. Biol.* **2017**, *1484*, 187–203. [[PubMed](#)]
140. Dosztanyi, Z.; Meszaros, B.; Simon, I. Anchor: Web server for predicting protein binding regions in disordered proteins. *Bioinformatics* **2009**, *25*, 2745–2746. [[CrossRef](#)] [[PubMed](#)]
141. Di Domenico, T.; Walsh, I.; Martin, A.J.M.; Tosatto, S.C.E. Mobidb: A comprehensive database of intrinsic protein disorder annotations. *Bioinformatics* **2012**, *28*, 2080–2081. [[CrossRef](#)] [[PubMed](#)]
142. Potenza, E.; Di Domenico, T.; Walsh, I.; Tosatto, S.C. Mobidb 2.0: An improved database of intrinsically disordered and mobile proteins. *Nucleic Acids Res.* **2015**, *43*, D315–D320. [[CrossRef](#)] [[PubMed](#)]
143. Necci, M.; Piovesan, D.; Dosztanyi, Z.; Tosatto, S.C.E. Mobidb-lite: Fast and highly specific consensus prediction of intrinsic disorder in proteins. *Bioinformatics* **2017**, *33*, 1402–1404. [[CrossRef](#)] [[PubMed](#)]
144. Dosztanyi, Z.; Csizmok, V.; Tompa, P.; Simon, I. The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *J. Mol. Biol.* **2005**, *347*, 827–839. [[CrossRef](#)] [[PubMed](#)]
145. Walsh, I.; Martin, A.J.; Di Domenico, T.; Tosatto, S.C. Espritz: Accurate and fast prediction of protein disorder. *Bioinformatics* **2012**, *28*, 503–509. [[CrossRef](#)] [[PubMed](#)]
146. Lee, K.M.; Hsu Ia, W.; Tarn, W.Y. Trap150 activates pre-mRNA splicing and promotes nuclear mRNA degradation. *Nucleic Acids Res.* **2010**, *38*, 3340–3350. [[CrossRef](#)] [[PubMed](#)]
147. Heyd, F.; Lynch, K.W. Phosphorylation-dependent regulation of PSF by GSK3 controls CD45 alternative splicing. *Mol. Cell* **2010**, *40*, 126–137. [[CrossRef](#)] [[PubMed](#)]
148. Beli, P.; Lukashchuk, N.; Wagner, S.A.; Weinert, B.T.; Olsen, J.V.; Baskcomb, L.; Mann, M.; Jackson, S.P.; Choudhary, C. Proteomic investigations reveal a role for RNA processing factor THRAP3 in the DNA damage response. *Mol. Cell* **2012**, *46*, 212–225. [[CrossRef](#)] [[PubMed](#)]
149. Ito, M.; Okano, H.J.; Darnell, R.B.; Roeder, R.G. The TRAP100 component of the TRAP/Mediator complex is essential in broad transcriptional events and development. *EMBO J.* **2002**, *21*, 3464–3475. [[CrossRef](#)] [[PubMed](#)]
150. Katano-Toki, A.; Satoh, T.; Tomaru, T.; Yoshino, S.; Ishizuka, T.; Ishii, S.; Ozawa, A.; Shibusawa, N.; Tsuchiya, T.; Saito, T.; et al. THRAP3 interacts with HELZ2 and plays a novel role in adipocyte differentiation. *Mol. Endocrinol.* **2013**, *27*, 769–780. [[CrossRef](#)] [[PubMed](#)]
151. Lande-Diner, L.; Boyault, C.; Kim, J.Y.; Weitz, C.J. A positive feedback loop links circadian clock factor clock-bmal1 to the basic transcriptional machinery. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 16021–16026. [[CrossRef](#)] [[PubMed](#)]
152. Horiuchi, K.; Kawamura, T.; Iwanari, H.; Ohashi, R.; Naito, M.; Kodama, T.; Hamakubo, T. Identification of wilms' tumor 1-associating protein complex and its role in alternative splicing and the cell cycle. *J. Biol. Chem.* **2013**, *288*, 33292–33302. [[CrossRef](#)] [[PubMed](#)]
153. Ino, Y.; Arakawa, N.; Ishiguro, H.; Uemura, H.; Kubota, Y.; Hirano, H.; Toda, T. Phosphoproteome analysis demonstrates the potential role of THRAP3 phosphorylation in androgen-independent prostate cancer cell growth. *Proteomics* **2016**, *16*, 1069–1078. [[CrossRef](#)] [[PubMed](#)]
154. Strasser, K.; Masuda, S.; Mason, P.; Pfannstiel, J.; Oppizzi, M.; Rodriguez-Navarro, S.; Rondon, A.G.; Aguilera, A.; Struhl, K.; Reed, R.; et al. Trex is a conserved complex coupling transcription with messenger RNA export. *Nature* **2002**, *417*, 304–308. [[CrossRef](#)] [[PubMed](#)]
155. Chi, B.; Wang, K.; Du, Y.; Gui, B.; Chang, X.; Wang, L.; Fan, J.; Chen, S.; Wu, X.; Li, G.; et al. A Sub-Element in PRE enhances nuclear export of intronless mRNAs by recruiting the TREX complex via ZC3H18. *Nucleic Acids Res.* **2014**, *42*, 7305–7318. [[CrossRef](#)] [[PubMed](#)]
156. Garee, J.P.; Oesterreich, S. Safb1's multiple functions in biological control-lots still to be done! *J. Cell. Biochem.* **2010**, *109*, 312–319. [[CrossRef](#)] [[PubMed](#)]

157. Nayler, O.; Stratling, W.; Bourquin, J.P.; Stajlar, I.; Lindemann, L.; Jasper, H.; Hartmann, A.M.; Fackelmayer, F.O.; Ullrich, A.; Stamm, S. SAF-B protein couples transcription and pre-mRNA splicing to SAR/MAR elements. *Nucleic Acids Res.* **1998**, *26*, 3542–3549. [[CrossRef](#)] [[PubMed](#)]
158. Aravind, L.; Koonin, E.V. Sap—A putative DNA-binding motif involved in chromosomal organization. *Trends Biochem. Sci.* **2000**, *25*, 112–114. [[CrossRef](#)]
159. Nery, F.C.; Rui, E.; Kuniyoshi, T.M.; Kobarg, J. Evidence for the interaction of the regulatory protein ki-1/57 with p53 and its interacting proteins. *Biochem. Biophys. Res. Commun.* **2006**, *341*, 847–855. [[CrossRef](#)] [[PubMed](#)]
160. Nery, F.C.; Passos, D.O.; Garcia, V.S.; Kobarg, J. Ki-1/57 interacts with RACK1 and is a substrate for the phosphorylation by phorbol 12-myristate 13-acetate-activated protein kinase C. *J. Biol. Chem.* **2004**, *279*, 11444–11455. [[CrossRef](#)] [[PubMed](#)]
161. Jobert, L.; Argentini, M.; Tora, L. PRMT1 mediated methylation of TAF15 is required for its positive gene regulatory function. *Exp. Cell Res.* **2009**, *315*, 1273–1286. [[CrossRef](#)] [[PubMed](#)]
162. Bertolotti, A.; Lutz, Y.; Heard, D.J.; Chambon, P.; Tora, L. Htaf(II)68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. *EMBO J.* **1996**, *15*, 5022–5031. [[PubMed](#)]
163. Law, W.J.; Cann, K.L.; Hicks, G.G. TLS, EWS and TAF15: A model for transcriptional integration of gene expression. *Brief. Funct. Genom. Proteom.* **2006**, *5*, 8–14. [[CrossRef](#)] [[PubMed](#)]
164. Arvand, A.; Denny, C.T. Biology of EWS/ETS fusions in Ewing's family tumors. *Oncogene* **2001**, *20*, 5747–5754. [[CrossRef](#)] [[PubMed](#)]
165. Bertolotti, A.; Bell, B.; Tora, L. The N-terminal domain of human TAFII68 displays transactivation and oncogenic properties. *Oncogene* **1999**, *18*, 8000–8010. [[CrossRef](#)] [[PubMed](#)]
166. Takayama, S.; Xie, Z.; Reed, J.C. An evolutionarily conserved family of HSP70/HSC70 molecular chaperone regulators. *J. Biol. Chem.* **1999**, *274*, 781–786. [[CrossRef](#)] [[PubMed](#)]
167. Rauch, J.N.; Zuiderweg, E.R.; Gestwicki, J.E. Non-canonical interactions between heat shock cognate protein 70 (HSC70) and BCL2-associated anthanogene (BAG) co-chaperones are important for client release. *J. Biol. Chem.* **2016**, *291*, 19848–19857. [[CrossRef](#)] [[PubMed](#)]
168. Ingham, R.J.; Colwill, K.; Howard, C.; Dettwiler, S.; Lim, C.S.; Yu, J.; Hersi, K.; Raaijmakers, J.; Gish, G.; Mbamalu, G.; et al. Ww domains provide a platform for the assembly of multiprotein networks. *Mol. Cell. Biol.* **2005**, *25*, 7092–7106. [[CrossRef](#)] [[PubMed](#)]
169. Chan, H.M.; La Thangue, N.B. P300/CBP proteins: Hats for transcriptional bridges and scaffolds. *J. Cell Sci.* **2001**, *114*, 2363–2373. [[PubMed](#)]
170. Cazzalini, O.; Sommatis, S.; Tillhon, M.; Dutto, I.; Bachi, A.; Rapp, A.; Nardo, T.; Scovassi, A.I.; Necchi, D.; Cardoso, M.C.; et al. CBP and p300 acetylate pcna to link its degradation with nucleotide excision repair synthesis. *Nucleic Acids Res.* **2014**, *42*, 8433–8448. [[CrossRef](#)] [[PubMed](#)]
171. Imhof, A.; Yang, X.J.; Ogryzko, V.V.; Nakatani, Y.; Wolffe, A.P.; Ge, H. Acetylation of general transcription factors by histone acetyltransferases. *Curr. Biol.* **1997**, *7*, 689–692. [[CrossRef](#)]
172. Zhang, W.; Bieker, J.J. Acetylation and modulation of erythroid kruppel-like factor (EKLF) activity by interaction with histone acetyltransferases. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 9855–9860. [[CrossRef](#)] [[PubMed](#)]
173. Matsuzaki, H.; Daitoku, H.; Hatta, M.; Aoyama, H.; Yoshimochi, K.; Fukamizu, A. Acetylation of FOXO1 alters its DNA-binding ability and sensitivity to phosphorylation. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 11278–11283. [[CrossRef](#)] [[PubMed](#)]
174. Snowden, A.W.; Perkins, N.D. Cell cycle regulation of the transcriptional coactivators p300 and CREB binding protein. *Biochem. Pharmacol.* **1998**, *55*, 1947–1954. [[CrossRef](#)]
175. Andrisani, O.M. CREB-mediated transcriptional control. *Crit. Rev. Eukaryot. Gene Expr.* **1999**, *9*, 19–32. [[PubMed](#)]
176. Dyson, H.J.; Wright, P.E. Role of intrinsic protein disorder in the function and interactions of the transcriptional coactivators CREB-binding protein (CBP) and p300. *J. Biol. Chem.* **2016**, *291*, 6714–6722. [[CrossRef](#)] [[PubMed](#)]
177. Bedford, D.C.; Kasper, L.H.; Fukuyama, T.; Brindle, P.K. Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases. *Epigenetics* **2010**, *5*, 9–15. [[CrossRef](#)] [[PubMed](#)]
178. Zakaryan, R.P.; Gehring, H. Identification and characterization of the nuclear localization/retention signal in the EWS proto-oncoprotein. *J. Mol. Biol.* **2006**, *363*, 27–38. [[CrossRef](#)] [[PubMed](#)]

179. Benjamin, L.E.; Fredericks, W.J.; Barr, F.G.; Rauscher, F.J., 3rd. Fusion of the EWS1 and WT1 genes as a result of the t(11;22)(p13;q12) translocation in desmoplastic small round cell tumors. *Med. Pediatr. Oncol.* **1996**, *27*, 434–439. [[CrossRef](#)]
180. May, W.A.; Denny, C.T. Biology of EWS/FLI and related fusion genes in Ewing's sarcoma and primitive neuroectodermal tumor. *Curr. Top. Microbiol. Immunol.* **1997**, *220*, 143–150. [[PubMed](#)]
181. Erkizan, H.V.; Uversky, V.N.; Toretzky, J.A. Oncogenic partnerships: EWS-FLI1 protein interactions initiate key pathways of Ewing's sarcoma. *Clin. Cancer Res.* **2010**, *16*, 4077–4083. [[CrossRef](#)] [[PubMed](#)]
182. Hegyi, H.; Buday, L.; Tompa, P. Intrinsic structural disorder confers cellular viability on oncogenic fusion proteins. *PLoS Comput. Biol.* **2009**, *5*, e1000552. [[CrossRef](#)] [[PubMed](#)]
183. Kriwacki, R.W.; Wu, J.; Tennant, L.; Wright, P.E.; Siuzdak, G. Probing protein structure using biochemical and biophysical methods. Proteolysis, matrix-assisted laser desorption/ionization mass spectrometry, high-performance liquid chromatography and size-exclusion chromatography of p21^{Waf1/Cip1/Sdi1}. *J. Chromatogr. A* **1997**, *777*, 23–30. [[CrossRef](#)]
184. LaBaer, J.; Garrett, M.D.; Stevenson, L.F.; Slingerland, J.M.; Sandhu, C.; Chou, H.S.; Fattaey, A.; Harlow, E. New functional activities for the p21 family of CDK inhibitors. *Genes Dev.* **1997**, *11*, 847–862. [[CrossRef](#)] [[PubMed](#)]
185. Karimian, A.; Ahmadi, Y.; Yousefi, B. Multiple functions of p21 in cell cycle, apoptosis and transcriptional regulation after DNA damage. *DNA Repair* **2016**, *42*, 63–71. [[CrossRef](#)] [[PubMed](#)]
186. Kroker, A.J.; Bruning, J.B. P21 exploits residue Tyr151 as a tether for high-affinity PCNA binding. *Biochemistry* **2015**, *54*, 3483–3493. [[CrossRef](#)] [[PubMed](#)]
187. Stucki, M.; Jackson, S.P. MDC1/NFBD1: A key regulator of the DNA damage response in higher eukaryotes. *DNA Repair* **2004**, *3*, 953–957. [[CrossRef](#)] [[PubMed](#)]
188. Lou, Z.; Chini, C.C.; Minter-Dykhouse, K.; Chen, J. Mediator of DNA damage checkpoint protein 1 regulates brca1 localization and phosphorylation in DNA damage checkpoint control. *J. Biol. Chem.* **2003**, *278*, 13599–13602. [[CrossRef](#)] [[PubMed](#)]
189. Stewart, G.S.; Wang, B.; Bignell, C.R.; Taylor, A.M.; Elledge, S.J. MDC1 is a mediator of the mammalian DNA damage checkpoint. *Nature* **2003**, *421*, 961–966. [[CrossRef](#)] [[PubMed](#)]
190. Lou, Z.; Minter-Dykhouse, K.; Wu, X.; Chen, J. MDC1 is coupled to activated CHK2 in mammalian DNA damage response pathways. *Nature* **2003**, *421*, 957–961. [[CrossRef](#)] [[PubMed](#)]
191. Goldberg, M.; Stucki, M.; Falck, J.; D'Amours, D.; Rahman, D.; Pappin, D.; Bartek, J.; Jackson, S.P. MDC1 is required for the intra-S-phase DNA damage checkpoint. *Nature* **2003**, *421*, 952–956. [[CrossRef](#)] [[PubMed](#)]
192. Peng, A.; Chen, P.L. NFBD1, like 53BP1, is an early and redundant transducer mediating CHK2 phosphorylation in response to DNA damage. *J. Biol. Chem.* **2003**, *278*, 8873–8876. [[CrossRef](#)] [[PubMed](#)]
193. Xu, X.; Stern, D.F. NFBD1/KIAA0170 is a chromatin-associated protein involved in DNA damage signaling pathways. *J. Biol. Chem.* **2003**, *278*, 8795–8803. [[CrossRef](#)] [[PubMed](#)]
194. Shang, Y.L.; Boderio, A.J.; Chen, P.L. NFBD1, a novel nuclear protein with signature motifs of FHA and BRCT, and an internal 41-amino acid repeat sequence, is an early participant in DNA damage response. *J. Biol. Chem.* **2003**, *278*, 6323–6329. [[CrossRef](#)] [[PubMed](#)]
195. Lehmann, G.; Udasin, R.G.; Livneh, I.; Ciechanover, A. Identification of UBact, a ubiquitin-like protein, along with other homologous components of a conjugation system and the proteasome in different gram-negative Bacteria. *Biochem. Biophys. Res. Commun.* **2017**, *483*, 946–950. [[CrossRef](#)] [[PubMed](#)]
196. Delley, C.L.; Muller, A.U.; Ziemski, M.; Weber-Ban, E. Prokaryotic ubiquitin-like protein and its ligase/delignase enzymes. *J. Mol. Biol.* **2017**, *429*, 3486–3499. [[CrossRef](#)] [[PubMed](#)]
197. Hicke, L. Protein regulation by monoubiquitin. *Nat. Rev. Mol. Cell Biol.* **2001**, *2*, 195–201. [[CrossRef](#)] [[PubMed](#)]
198. Pickart, C.M. Ubiquitin enters the new millennium. *Mol. Cell* **2001**, *8*, 499–504. [[CrossRef](#)]
199. Pornillos, O.; Garrus, J.E.; Sundquist, W.I. Mechanisms of enveloped RNA virus budding. *Trends Cell Biol.* **2002**, *12*, 569–579. [[CrossRef](#)]
200. Muratani, M.; Tansey, W.P. How the ubiquitin-proteasome system controls transcription. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 192–201. [[CrossRef](#)] [[PubMed](#)]
201. Terrell, J.; Shih, S.; Dunn, R.; Hicke, L. A function for monoubiquitination in the internalization of a G protein-coupled receptor. *Mol. Cell* **1998**, *1*, 193–202. [[CrossRef](#)]
202. Carmichael, R.E.; Henley, J.M. Transcriptional and post-translational regulation of Arc in synaptic plasticity. *Semin. Cell Dev. Biol.* **2017**. [[CrossRef](#)] [[PubMed](#)]

203. Torrecilla, I.; Oehler, J.; Ramadan, K. The role of ubiquitin-dependent segregase p97 (VCP or Cdc48) in chromatin dynamics after DNA double strand breaks. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2017**, *372*. [[CrossRef](#)] [[PubMed](#)]
204. Rome, S.; Meugnier, E.; Vidal, H. The ubiquitin-proteasome pathway is a new partner for the control of insulin signaling. *Curr. Opin. Clin. Nutr. Metab. Care* **2004**, *7*, 249–254. [[CrossRef](#)] [[PubMed](#)]
205. Izzì, L.; Attisano, L. Regulation of the TGFβ signaling pathway by ubiquitin-mediated degradation. *Oncogene* **2004**, *23*, 2071–2078. [[CrossRef](#)] [[PubMed](#)]
206. Zhou, Y.; He, C.; Wang, L.; Ge, B. Post-translational regulation of antiviral innate signaling. *Eur. J. Immunol.* **2017**, *47*, 1414–1426. [[CrossRef](#)] [[PubMed](#)]
207. Lindorff-Larsen, K.; Best, R.B.; Depristo, M.A.; Dobson, C.M.; Vendruscolo, M. Simultaneous determination of protein structure and dynamics. *Nature* **2005**, *433*, 128–132. [[CrossRef](#)] [[PubMed](#)]
208. Chen, X.; Solomon, W.C.; Kang, Y.; Cerda-Maira, F.; Darwin, K.H.; Walters, K.J. Prokaryotic ubiquitin-like protein pup is intrinsically disordered. *J. Mol. Biol.* **2009**, *392*, 208–217. [[CrossRef](#)] [[PubMed](#)]
209. Liao, S.; Shang, Q.; Zhang, X.; Zhang, J.; Xu, C.; Tu, X. Pup, a prokaryotic ubiquitin-like protein, is an intrinsically disordered protein. *Biochem. J.* **2009**, *422*, 207–215. [[CrossRef](#)] [[PubMed](#)]
210. Vidal, M.; Goudreau, N.; Cornille, F.; Cussac, D.; Gincel, E.; Garbay, C. Molecular and cellular analysis of GRB2 SH3 domain mutants: Interaction with Sos and dynamin. *J. Mol. Biol.* **1999**, *290*, 717–730. [[CrossRef](#)] [[PubMed](#)]
211. Etmayer, P.; France, D.; Gounarides, J.; Jarosinski, M.; Martin, M.S.; Rondeau, J.M.; Sabio, M.; Topiol, S.; Weidmann, B.; Zurini, M.; et al. Structural and conformational requirements for high-affinity binding to the SH2 domain of GRB2(1). *J. Med. Chem.* **1999**, *42*, 971–980. [[CrossRef](#)] [[PubMed](#)]
212. Rice, C.; Shastrula, P.K.; Kossenkov, A.V.; Hills, R.; Baird, D.M.; Showe, L.C.; Doukov, T.; Janicki, S.; Skordalakes, E. Structural and functional analysis of the human POT1-TPP1 telomeric complex. *Nat. Commun.* **2017**, *8*, 14928. [[CrossRef](#)] [[PubMed](#)]
213. Bono, F.; Ebert, J.; Lorentzen, E.; Conti, E. The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. *Cell* **2006**, *126*, 713–725. [[CrossRef](#)] [[PubMed](#)]
214. Tanase, C.A. Histidine domain-protein tyrosine phosphatase interacts with GRB2 and GRPL. *PLoS ONE* **2010**, *5*, e14339. [[CrossRef](#)] [[PubMed](#)]
215. Belov, A.A.; Mohammadi, M. GRB2, a double-edged sword of receptor tyrosine kinase signaling. *Sci. Signal.* **2012**, *5*, pe49. [[CrossRef](#)] [[PubMed](#)]
216. Jang, I.K.; Zhang, J.; Gu, H. GRB2, a simple adapter with complex roles in lymphocyte development, function, and signaling. *Immunol. Rev.* **2009**, *232*, 150–159. [[CrossRef](#)] [[PubMed](#)]
217. Lemmon, M.A.; Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **2010**, *141*, 1117–1134. [[CrossRef](#)] [[PubMed](#)]
218. Giubellino, A.; Burke, T.R., Jr.; Bottaro, D.P. GRB2 signaling in cell motility and cancer. *Expert Opin. Ther. Targets* **2008**, *12*, 1021–1033. [[CrossRef](#)] [[PubMed](#)]
219. Pandey, P.; Kharbanda, S.; Kufe, D. Association of the DF3/MUC1 breast cancer antigen with GRB2 and the SOS/RAS exchange protein. *Cancer Res.* **1995**, *55*, 4000–4003. [[PubMed](#)]
220. Chardin, P.; Camonis, J.H.; Gale, N.W.; van Aelst, L.; Schlessinger, J.; Wigler, M.H.; Bar-Sagi, D. Human SOS1: A guanine nucleotide exchange factor for RAS that binds to GRB2. *Science* **1993**, *260*, 1338–1343. [[CrossRef](#)] [[PubMed](#)]
221. Tarcic, G.; Boguslavsky, S.K.; Wakim, J.; Kiuchi, T.; Liu, A.; Reinitz, F.; Nathanson, D.; Takahashi, T.; Mischel, P.S.; Ng, T.; et al. An unbiased screen identifies dep-1 tumor suppressor as a phosphatase controlling egfr endocytosis. *Curr. Biol.* **2009**, *19*, 1788–1798. [[CrossRef](#)] [[PubMed](#)]
222. Braverman, L.E.; Quilliam, L.A. Identification of GRB4/Nckβ, a src homology 2 and 3 domain-containing adapter protein having similar binding and biological properties to Nck. *J. Biol. Chem.* **1999**, *274*, 5542–5549. [[CrossRef](#)] [[PubMed](#)]
223. Pfrepper, K.I.; Marie-Cardine, A.; Simeoni, L.; Kuramitsu, Y.; Leo, A.; Spicka, J.; Hilgert, I.; Scherer, J.; Schraven, B. Structural and functional dissection of the cytoplasmic domain of the transmembrane adaptor protein SIT (SHP2-interacting transmembrane adaptor protein). *Eur. J. Immunol.* **2001**, *31*, 1825–1836. [[CrossRef](#)]
224. Fantin, V.R.; Sparling, J.D.; Slot, J.W.; Keller, S.R.; Lienhard, G.E.; Lavan, B.E. Characterization of insulin receptor substrate 4 in human embryonic kidney 293 cells. *J. Biol. Chem.* **1998**, *273*, 10726–10732. [[CrossRef](#)] [[PubMed](#)]

225. Tobe, K.; Matuoka, K.; Tamemoto, H.; Ueki, K.; Kaburagi, Y.; Asai, S.; Noguchi, T.; Matsuda, M.; Tanaka, S.; Hattori, S.; et al. Insulin stimulates association of insulin receptor substrate-1 with the protein abundant src homology/growth factor receptor-bound protein 2. *J. Biol. Chem.* **1993**, *268*, 11167–11171. [[PubMed](#)]
226. Skolnik, E.Y.; Lee, C.H.; Batzer, A.; Vicentini, L.M.; Zhou, M.; Daly, R.; Myers, M.J., Jr.; Backer, J.M.; Ullrich, A.; White, M.F.; et al. The SH2/SH3 domain-containing protein GRB2 interacts with tyrosine-phosphorylated IRS1 and SHC: Implications for insulin control of RAS signalling. *EMBO J.* **1993**, *12*, 1929–1936. [[PubMed](#)]
227. Yokouchi, M.; Suzuki, R.; Masuhara, M.; Komiyama, S.; Inoue, A.; Yoshimura, A. Cloning and characterization of APS, an adaptor molecule containing PH and SH2 domains that is tyrosine phosphorylated upon B-cell receptor stimulation. *Oncogene* **1997**, *15*, 7–15. [[CrossRef](#)] [[PubMed](#)]
228. Janssen, E.; Zhu, M.; Zhang, W.; Koonpaew, S.; Zhang, W. Lab: A new membrane-associated adaptor molecule in B cell activation. *Nat. Immunol.* **2003**, *4*, 117–123. [[CrossRef](#)] [[PubMed](#)]
229. Brdicka, T.; Imrich, M.; Angelisova, P.; Brdickova, N.; Horvath, O.; Spicka, J.; Hilgert, I.; Luskova, P.; Draber, P.; Novak, P.; et al. Non-T cell activation linker (NTAL): A transmembrane adaptor protein involved in immunoreceptor signaling. *J. Exp. Med.* **2002**, *196*, 1617–1626. [[CrossRef](#)] [[PubMed](#)]
230. Zhu, M.; Janssen, E.; Leung, K.; Zhang, W. Molecular cloning of a novel gene encoding a membrane-associated adaptor protein (LAX) in lymphocyte signaling. *J. Biol. Chem.* **2002**, *277*, 46151–46158. [[CrossRef](#)] [[PubMed](#)]
231. Zhang, W.; Sloan-Lancaster, J.; Kitchen, J.; Tribble, R.P.; Samelson, L.E. LAT: The ZAP-70 tyrosine kinase substrate that links T cell receptor to cellular activation. *Cell* **1998**, *92*, 83–92. [[CrossRef](#)]
232. Kharitonov, A.; Chen, Z.; Sures, I.; Wang, H.; Schilling, J.; Ullrich, A. A family of proteins that inhibit signalling through tyrosine kinase receptors. *Nature* **1997**, *386*, 181–186. [[CrossRef](#)] [[PubMed](#)]
233. Sano, H.; Liu, S.C.; Lane, W.S.; Piletz, J.E.; Lienhard, G.E. Insulin receptor substrate 4 associates with the protein IRAS. *J. Biol. Chem.* **2002**, *277*, 19439–19447. [[CrossRef](#)] [[PubMed](#)]
234. Ikeda, M.; Ishida, O.; Hinoi, T.; Kishida, S.; Kikuchi, A. Identification and characterization of a novel protein interacting with Ral-binding protein 1, a putative effector protein of Ral. *J. Biol. Chem.* **1998**, *273*, 814–821. [[CrossRef](#)] [[PubMed](#)]
235. Wheeler, M.; Domin, J. Recruitment of the class ii phosphoinositide 3-kinase c2beta to the epidermal growth factor receptor: Role of GRB2. *Mol. Cell. Biol.* **2001**, *21*, 6660–6667. [[CrossRef](#)] [[PubMed](#)]
236. Pollard, T.D.; Blanchoin, L.; Mullins, R.D. Molecular mechanisms controlling actin filament dynamics in nonmuscle cells. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 545–576. [[CrossRef](#)] [[PubMed](#)]
237. Povarova, O.I.; Uversky, V.N.; Kuznetsova, I.M.; Turoverov, K.K. Actin: enigma or enigmatic actin: Folding, structure, and functions of the most abundant eukaryotic protein. *Intrinsically Disord. Proteins* **2014**, *2*, e34500. [[CrossRef](#)] [[PubMed](#)]
238. Kuznetsova, I.M.; Povarova, O.I.; Uversky, V.N.; Turoverov, K.K. Native globular actin has a thermodynamically unstable quasi-stationary structure with elements of intrinsic disorder. *FEBS J.* **2016**, *283*, 438–445. [[CrossRef](#)] [[PubMed](#)]
239. Barua, B.; Winkelmann, D.A.; White, H.D.; Hitchcock-DeGregori, S.E. Regulation of actin-myosin interaction by conserved periodic sites of tropomyosin. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 18425–18430. [[CrossRef](#)] [[PubMed](#)]
240. Carlier, M.F.; Valentin-Ranc, C.; Combeau, C.; Fievez, S.; Pantaloni, D. Actin polymerization: Regulation by divalent metal ion and nucleotide binding, atp hydrolysis and binding of myosin. *Adv. Exp. Med. Biol.* **1994**, *358*, 71–81. [[PubMed](#)]
241. Guharoy, M.; Szabo, B.; Contreras Martos, S.; Kosol, S.; Tompa, P. Intrinsic structural disorder in cytoskeletal proteins. *Cytoskeleton* **2013**, *70*, 550–571. [[CrossRef](#)] [[PubMed](#)]
242. Dos Remedios, C.G.; Chhabra, D.; Kekic, M.; Dedova, I.V.; Tsubakihara, M.; Berry, D.A.; Nosworthy, N.J. Actin binding proteins: Regulation of cytoskeletal microfilaments. *Physiol. Rev.* **2003**, *83*, 433–473. [[CrossRef](#)] [[PubMed](#)]
243. De Lange, T. Shelterin: The protein complex that shapes and safeguards human telomeres. *Genes Dev.* **2005**, *19*, 2100–2110. [[CrossRef](#)] [[PubMed](#)]
244. Liu, D.; O'Connor, M.S.; Qin, J.; Songyang, Z. Telosome, a mammalian telomere-associated complex formed by multiple telomeric proteins. *J. Biol. Chem.* **2004**, *279*, 51338–51342. [[CrossRef](#)] [[PubMed](#)]
245. Venteicher, A.S.; Abreu, E.B.; Meng, Z.; McCann, K.E.; Terns, R.M.; Veenstra, T.D.; Terns, M.P.; Artandi, S.E. A human telomerase holoenzyme protein required for cajal body localization and telomere synthesis. *Science* **2009**, *323*, 644–648. [[CrossRef](#)] [[PubMed](#)]

246. Schmidt, U.; Im, K.B.; Benzing, C.; Janjetovic, S.; Rippe, K.; Lichter, P.; Wachsmuth, M. Assembly and mobility of exon-exon junction complexes in living cells. *RNA* **2009**, *15*, 862–876. [[CrossRef](#)] [[PubMed](#)]
247. Jurica, M.S.; Licklider, L.J.; Gygi, S.R.; Grigorieff, N.; Moore, M.J. Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. *RNA* **2002**, *8*, 426–439. [[CrossRef](#)] [[PubMed](#)]
248. Michelle, L.; Cloutier, A.; Toutant, J.; Shkreta, L.; Thibault, P.; Durand, M.; Garneau, D.; Gendron, D.; Lapointe, E.; Couture, S.; et al. Proteins associated with the exon junction complex also control the alternative splicing of apoptotic regulators. *Mol. Cell. Biol.* **2012**, *32*, 954–967. [[CrossRef](#)] [[PubMed](#)]
249. Tange, T.O.; Nott, A.; Moore, M.J. The ever-increasing complexities of the exon junction complex. *Curr. Opin. Cell Biol.* **2004**, *16*, 279–284. [[CrossRef](#)] [[PubMed](#)]
250. Licata, L.; Briganti, L.; Peluso, D.; Perfetto, L.; Iannuccelli, M.; Galeota, E.; Sacco, F.; Palma, A.; Nardoza, A.P.; Santonico, E.; et al. Mint, the molecular interaction database: 2012 update. *Nucleic Acids Res.* **2012**, *40*, D857–D861. [[CrossRef](#)] [[PubMed](#)]
251. Licata, L.; Orchard, S. The mintact project and molecular interaction databases. *Methods Mol. Biol.* **2016**, *1415*, 55–69. [[PubMed](#)]
252. Orchard, S.; Ammari, M.; Aranda, B.; Breuza, L.; Briganti, L.; Broackes-Carter, F.; Campbell, N.H.; Chavali, G.; Chen, C.; del-Toro, N.; et al. The MIntAct project—IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.* **2014**, *42*, D358–D363. [[CrossRef](#)] [[PubMed](#)]
253. Salwinski, L.; Miller, C.S.; Smith, A.J.; Pettit, F.K.; Bowie, J.U.; Eisenberg, D. The database of interacting proteins: 2004 update. *Nucleic Acids Res.* **2004**, *32*, D449–D451. [[CrossRef](#)] [[PubMed](#)]
254. Launay, G.; Salza, R.; Multedo, D.; Thierry-Mieg, N.; Ricard-Blum, S. Matrixdb, the extracellular matrix interaction database: Updated content, a new navigator and expanded functionalities. *Nucleic Acids Res.* **2015**, *43*, D321–D327. [[CrossRef](#)] [[PubMed](#)]
255. Chatr-Aryamontri, A.; Oughtred, R.; Boucher, L.; Rust, J.; Chang, C.; Kolas, N.K.; O'Donnell, L.; Oster, S.; Theesfeld, C.; Sellam, A.; et al. The biogrid interaction database: 2017 update. *Nucleic Acids Res.* **2017**, *45*, D369–D379. [[CrossRef](#)] [[PubMed](#)]
256. The UniProt Consortium. Uniprot: The universal protein knowledgebase. *Nucleic Acids Res.* **2017**, *45*, D158–D169.
257. Kiran, M.; Nagarajaram, H.A. Global versus local hubs in human protein-protein interaction network. *J. Proteome Res.* **2013**, *12*, 5436–5446. [[CrossRef](#)] [[PubMed](#)]
258. Bertin, N.; Simonis, N.; Dupuy, D.; Cusick, M.E.; Han, J.D.; Fraser, H.B.; Roth, F.P.; Vidal, M. Confirmation of organized modularity in the yeast interactome. *PLoS Biol.* **2007**, *5*, e153. [[CrossRef](#)] [[PubMed](#)]
259. Yu, H.; Kim, P.M.; Sprecher, E.; Trifonov, V.; Gerstein, M. The importance of bottlenecks in protein networks: Correlation with gene essentiality and expression dynamics. *PLoS Comput. Biol.* **2007**, *3*, e59. [[CrossRef](#)] [[PubMed](#)]
260. Batada, N.N.; Reguly, T.; Breitkreutz, A.; Boucher, L.; Breitkreutz, B.J.; Hurst, L.D.; Tyers, M. Still stratus not altocumulus: Further evidence against the date/party hub distinction. *PLoS Biol.* **2007**, *5*, 1202–1206. [[CrossRef](#)] [[PubMed](#)]
261. Agarwal, S.; Deane, C.M.; Porter, M.A.; Jones, N.S. Revisiting date and party hubs: Novel approaches to role assignment in protein interaction networks. *PLoS Comput. Biol.* **2010**, *6*, e1000817. [[CrossRef](#)] [[PubMed](#)]
262. Peng, Z.L.; Kurgan, L. Comprehensive comparative assessment of in-silico predictors of disordered regions. *Curr. Protein Pept. Sci.* **2012**, *13*, 6–18. [[CrossRef](#)] [[PubMed](#)]
263. Walsh, I.; Giollo, M.; Di Domenico, T.; Ferrari, C.; Zimmermann, O.; Tosatto, S.C. Comprehensive large-scale assessment of intrinsic protein disorder. *Bioinformatics* **2015**, *31*, 201–208. [[CrossRef](#)] [[PubMed](#)]
264. Meng, F.; Uversky, V.N.; Kurgan, L. Comprehensive review of methods for prediction of intrinsic disorder and its molecular functions. *Cell. Mol. Life Sci.* **2017**, *74*, 3069–3090. [[CrossRef](#)] [[PubMed](#)]
265. Piovesan, D.; Tabaro, F.; Micetic, I.; Necci, M.; Quaglia, F.; Oldfield, C.J.; Aspromonte, M.C.; Davey, N.E.; Davidovic, R.; Dosztanyi, Z.; et al. Disprot 7.0: A major update of the database of disordered proteins. *Nucleic Acids Res.* **2016**, *D1*, D219–D227. [[CrossRef](#)] [[PubMed](#)]
266. Sickmeier, M.; Hamilton, J.A.; LeGall, T.; Vacic, V.; Cortese, M.S.; Tantos, A.; Szabo, B.; Tompa, P.; Chen, J.; Uversky, V.N.; et al. Disprot: The database of disordered proteins. *Nucleic Acids Res.* **2007**, *35*, D786–D793. [[CrossRef](#)] [[PubMed](#)]
267. Peng, Z.; Kurgan, L. On the complementarity of the consensus-based disorder prediction. *Pac. Symp. Biocomput.* **2012**, *8*, 176–187.
268. Fan, X.; Kurgan, L. Accurate prediction of disorder in protein chains with a comprehensive and empirically designed consensus. *J. Biomol. Struct. Dyn.* **2014**, *32*, 448–464. [[CrossRef](#)] [[PubMed](#)]

269. Na, I.; Meng, F.; Kurgan, L.; Uversky, V.N. Autophagy-related intrinsically disordered proteins in intra-nuclear compartments. *Mol. Biosyst.* **2016**, *12*, 2798–2817. [[CrossRef](#)] [[PubMed](#)]
270. Peng, Z.; Oldfield, C.J.; Xue, B.; Mizianty, M.J.; Dunker, A.K.; Kurgan, L.; Uversky, V.N. A creature with a hundred waggly tails: Intrinsically disordered proteins in the ribosome. *Cell. Mol. Life Sci.* **2014**, *71*, 1477–1504. [[CrossRef](#)] [[PubMed](#)]
271. Hu, G.; Wu, Z.; Wang, K.; Uversky, V.N.; Kurgan, L. Untapped potential of disordered proteins in current druggable human proteome. *Curr. Drug Targets* **2016**, *17*, 1198–1205. [[CrossRef](#)] [[PubMed](#)]
272. Peng, Z.; Uversky, V.N.; Kurgan, L. Genes encoding intrinsic disorder in Eukaryota have high gc content. *Intrinsically Disord. Proteins* **2016**, *4*, e1262225. [[CrossRef](#)] [[PubMed](#)]
273. Oates, M.E.; Romero, P.; Ishida, T.; Ghalwash, M.; Mizianty, M.J.; Xue, B.; Dosztanyi, Z.; Uversky, V.N.; Obradovic, Z.; Kurgan, L.; et al. D(2)p(2): Database of disordered protein predictions. *Nucleic Acids Res.* **2013**, *41*, D508–D516. [[CrossRef](#)] [[PubMed](#)]
274. Monastyrskyy, B.; Kryshtafovych, A.; Moulton, J.; Tramontano, A.; Fidelis, K. Assessment of protein disorder region predictions in casp10. *Proteins* **2014**, *82* (Suppl. 2), 127–137. [[CrossRef](#)] [[PubMed](#)]
275. Pentony, M.M.; Jones, D.T. Modularity of intrinsic disorder in the human proteome. *Proteins* **2010**, *78*, 212–221. [[CrossRef](#)] [[PubMed](#)]
276. Tompa, P.; Fuxreiter, M.; Oldfield, C.J.; Simon, I.; Dunker, A.K.; Uversky, V.N. Close encounters of the third kind: Disordered domains and the interactions of proteins. *Bioessays* **2009**, *31*, 328–335. [[CrossRef](#)] [[PubMed](#)]
277. Disfani, F.M.; Hsu, W.L.; Mizianty, M.J.; Oldfield, C.J.; Xue, B.; Dunker, A.K.; Uversky, V.N.; Kurgan, L. Morfpred, a computational tool for sequence-based prediction and characterization of short disorder-to-order transitioning binding regions in proteins. *Bioinformatics* **2012**, *28*, i75–i83. [[CrossRef](#)] [[PubMed](#)]
278. Jones, D.T.; Cozzetto, D. Disopred3: Precise disordered region predictions with annotated protein-binding activity. *Bioinformatics* **2015**, *31*, 857–863. [[CrossRef](#)] [[PubMed](#)]
279. Wang, K.; Samudrala, R. Incorporating background frequency improves entropy-based residue conservation measures. *BMC Bioinform.* **2006**, *7*, 385.
280. Altschul, S.F.; Madden, T.L.; Schaffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped blast and psi-blast: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [[CrossRef](#)] [[PubMed](#)]
281. Van Roey, K.; Uyar, B.; Weatheritt, R.J.; Dinkel, H.; Seiler, M.; Budd, A.; Gibson, T.J.; Davey, N.E. Short linear motifs: Ubiquitous and functionally diverse protein interaction modules directing cell regulation. *Chem. Rev.* **2014**, *114*, 6733–6778. [[CrossRef](#)] [[PubMed](#)]
282. Dinkel, H.; Van Roey, K.; Michael, S.; Kumar, M.; Uyar, B.; Altenberg, B.; Milchevskaya, V.; Schneider, M.; Kuhn, H.; Behrendt, A.; et al. Elm 2016—Data update and new functionality of the eukaryotic linear motif resource. *Nucleic Acids Res.* **2016**, *44*, D294–D300. [[CrossRef](#)] [[PubMed](#)]
283. Van Roey, K.; Dinkel, H.; Weatheritt, R.J.; Gibson, T.J.; Davey, N.E. The switches.Elm resource: A compendium of conditional regulatory interaction interfaces. *Sci. Signal.* **2013**, *6*, rs7. [[CrossRef](#)] [[PubMed](#)]
284. Davey, N.E.; Van Roey, K.; Weatheritt, R.J.; Toedt, G.; Uyar, B.; Altenberg, B.; Budd, A.; Diella, F.; Dinkel, H.; Gibson, T.J. Attributes of short linear motifs. *Mol. Biosyst.* **2012**, *8*, 268–281. [[CrossRef](#)] [[PubMed](#)]
285. Fuxreiter, M.; Tompa, P.; Simon, I. Local structural disorder imparts plasticity on linear motifs. *Bioinformatics* **2007**, *23*, 950–956. [[CrossRef](#)] [[PubMed](#)]
286. Dinkel, H.; Van Roey, K.; Michael, S.; Davey, N.E.; Weatheritt, R.J.; Born, D.; Speck, T.; Kruger, D.; Grebnev, G.; Kuban, M.; et al. The eukaryotic linear motif resource elm: 10 years and counting. *Nucleic Acids Res.* **2014**, *42*, D259–D266. [[CrossRef](#)] [[PubMed](#)]
287. Gould, C.M.; Diella, F.; Via, A.; Puntervoll, P.; Gemund, C.; Chabanis-Davidson, S.; Michael, S.; Sayadi, A.; Bryne, J.C.; Chica, C.; et al. Elm: The status of the 2010 eukaryotic linear motif resource. *Nucleic Acids Res.* **2010**, *38*, D167–D180. [[CrossRef](#)] [[PubMed](#)]

