# Wrecked regulation of intrinsically disordered proteins in diseases: pathogenicity of deregulated regulators

## Vladimir N. Uversky [1,2,3] *

[1] Department of Molecular Medicine and USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL, USA

[2] Biology Department, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia

[3] Laboratory of New Methods in Biology, Institute for Biological Instrumentation, Russian Academy of Sciences, Moscow, Russia

Biologically active proteins without stable tertiary structure are common in all known proteomes. Functions of these intrinsically disordered proteins (IDPs) are typically related to regulation, signaling, and control. Cellular levels of these important regulators are tightly regulated by a variety mechanisms ranging from firmly controlled expression to precisely targeted degradation. Functions of IDPs are controlled by binding to specific partners, alternative splicing, and posttranslational modifications among other means. In the norm, right amounts of precisely activated IDPs have to be present in right time at right places. Wrecked regulation brings havoc to the ordered world of disordered proteins, leading to protein misfolding, misidentification, and missignaling that give rise to numerous human diseases, such as cancer, cardiovascular disease, neurodegenerative diseases, and diabetes. Among factors inducing pathogenic transformations of IDPs are various cellular mechanisms, such as chromosomal translocations, damaged splicing, altered expression, frustrated posttranslational modifications, aberrant proteolytic degradation, and defective trafficking. This review presents some of the aspects of deregulated regulation of IDPs leading to human diseases.

**Keywords: intrinsically disordered proteins, conformational diseases, posttranslational modification, alternative splicing, transcriptional activation, expression, proteolytic degradation, trafficking**

**Abbreviations:** Aβ, amyloid-β; ACE, angiotensin I-converting enzyme; ACE2, angiotensin-converting enzyme 2; AChE, acetylcholinesterase; AD, Alzheimer's disease; ALCL, anaplastic large cell lymphoma; α-MHC, α-myosin heavy chain; ALK, anaplastic lymphoma kinase; ALL, acute lymphoblastic leukemia; AML, acute myelogenous leukemia; APC, Adenomatous polyposis coli protein; ApoB, apolipoprotein B; AQP2, aquaporin-2 channel; AR, androgen receptor; AS, alternative splicing; BBB, blood-brain barrier; BDNF, brain-derived neurotrophic factor; BRCA1, breast cancer type 1 susceptibility protein; Brunol, Bruno-like; CARS, cysteinyl-tRNA synthetase; CELF, CUG-BP, Elav-like family; CCRK, cell-cycle-regulated kinase; CFTR, cystic fibrosis transmembrane conductance regulator; CIMF, chronic idiopathic myelofibrosis; CLTC, clathrin heavy chain; CLU, clusterin; CML, chronic myelogenous leukemia; CMML, chronic myelomonocytic leukemia; CMTX, Charcot-Marie-Tooth disease; CRC, colorectal cancer; CRMP, collapsin response mediator protein; CTA, cancer/testis antigens; CVD, cardiovascular disease; Cx, connexin; DCM, dilated cardiomyopathy; DCTN1, dynactin; DM1, myotonic dystrophy type 1; DS, Down syndrome (trisomy 21); DYRK1A, dual-specificity tyrosine phosphorylation-regulated kinase 1A; ECM, extracellular matrix; ER, endoplasmic reticulum; ETV6, ETS translocation variant 6 (also known as TEL); EWS, Ewing's sarcoma; FGFR1, fibroblast growth factor receptor 1; FGFR1OP, fibroblast growth factor receptor 1 oncogene partner; FIR, FUSE-binding protein-interacting repressor (also known as PUF60); FTDP-17, frontotemporal dementia with Parkinsonism; Gli, Glioma-associated oncogene; HAT, histone acetyltransferase; HCC, hepatocellular carcinoma; HD, Huntington's disease; HDAC, histone deacetylase; Hh, Hedgehog; HLH, helix-loop-helix; hnRNPM, heterogeneous nuclear ribonucleoprotein M; Htt, Huntingtin protein; IDP, intrinsically disordered protein; IDPR, intrinsically disordered protein region; IGF, insulin-like growth factor; IGFBPs, insulin-like growth factor binding protein; IL1RL1, interleukin-1 receptor-like 1 (also known as ST2 protein); KLF5, Kruppel-like factor 5; JAK2, Janus tyrosine kinase 2; KIF5B, kinesin 1 heavy chain; KLC1, kinesin light chain 1; KLF6, Kruppel-like Zn-finger transcription factor 6; MSN, moesin; LDLR, low-density lipoprotein receptor; LFSNHI, low frequency sensorineural hearing impairment; LTCC, L-type calcium channel; MDS, myelodysplastic syndromes; Mfn2, mitochondrial fusion protein mitofusin 2; MI, myocardial infarction; MMP, metalloproteinase; MYH9, myosin II heavy chain type A; NaV1.5, voltage-gated Na channel isoform 1.5; NBD, neurodegenerative brain disease; NDI, nephrogenic diabetes insipidus; NGS, next-generation sequencing; NO, nitric oxide; NOS, nitric oxide synthase; NPM, nucleolar phosphoprotein nucleophosmin; NRSF, neuron-restrictive silencer factor; NSCLC, non-small-cell lung carcinoma; O-GlcNAc, O-linked N-acetylglucosamine; OSCC, oral squamous cell carcinoma; PAI-1, plasma plasminogen activator inhibitor-1; PD, Parkinson's disease; PDGFRB, platelet-derived growth factor receptor beta; PG, proteoglycan; PGC-1α, peroxisome proliferator-activated receptor gamma coactivator 1-α; PGC-1β, peroxisome proliferator-activated receptor gamma coactivator 1-β; PJS, Peutz-Jeghers Syndrome; PMD, potentially malignant disorder; PMF, primary myelofibrosis; PPAR, peroxisome proliferator-activated receptor; PPFIBP1, protein tyrosine phosphatase receptor type F polypeptide-interacting protein-binding protein 1 (also known as liprin beta 1); PrP, prion protein; PrP^C, soluble form of PrP; PRPF6, pre-mRNA processing factor 6; PrP^Sc, abnormal (infectious) scrapie form of PrP; PTEN, phosphatase and tensin homolog, deleted on chromosome TEN; PTM, posttranslational modification; PUF60, Poly(U)-binding-splicing factor (also known as FIR); SCN5A, sodium channel 5A; SERCA, sarcoplasmic reticulum Ca²⁺ ATPase; SEPT, septin; SMA, spinal muscular atrophy; SMN1, survival motor neuron-1; snRNP, small nuclear ribonucleic particle; TEL, ETS translocation variant 6 (also known as ETV6); TDP-43, TAR DNA binding protein 43; TFG, TRK-fused gene; TGFBR2, transforming growth factor beta receptor 2; TK, tyrosine kinase; TMPRSS2, transmembrane protease serine 2; TPM4, tropomyosin 4; tri-snRNP, tri-small ribonucleoprotein; RANBP2, RAN binding protein 2; RAS, renin-angiotensin system; REST, RE-1 silencing transcription factor; RET, REarranged during Transfection; RNF213, ring finger protein 213; RyR2, ryanodine receptor 2, cardiac; SQSTM1, sequestosome 1; TSE, spongiform encephalopathy; VCL, vinculin; VLDL, very low density lipoprotein; VSMC, vascular smooth muscle cell; ZO, zona occludens protein.

## INTRODUCTION

For more than a 100 years, it was believed that unique protein function is defined by the unique 3-D structure of a protein, which, in its turn, is defined by a unique amino acid sequence (Fischer, 1894), and the validity of this "one sequence—one structure—one function" concept was unquestioned for a long time, especially after the crystal structures of proteins started to be solved by X-ray diffraction. However, research, especially during the past 15 years, clearly indicated that the lack of stable tertiary and/or secondary structure does not preclude proteins from being biologically active (Wright and Dyson, 1999; Uversky et al., 2000; Dunker et al., 2001; Tompa, 2002; Uversky and Dunker, 2010). In fact, many proteins with crucial biological functions exist as collapsed or extended dynamically mobile conformational ensembles and do not have unique well-defined 3D structures as a whole or in their noticeable parts (at least in *in vitro* experiments). These proteins are known as intrinsically disordered proteins (IDPs) or hybrid proteins possessing both structured domains d biologically important intrinsically disordered protein regions (IDPRs). In addition to foldable (ordered) and non-foldable (intrinsically disordered) proteins, many proteins are known to misfold. Often such misfolding is accompanied by protein aggregation causing several human diseases that originate from the deposition of protein aggregates formed from specific proteins or protein fragments, which accumulate in a variety of organs and tissues (Kelly, 1998; Bellotti et al., 1999; Dobson, 1999; Uversky et al., 1999a,b; Rochet and Lansbury, 2000; Uversky and Fink, 2004; Gasperini et al., 2012; Moreau and King, 2012; Safar, 2012; Walker and LeVine, 2012; Cuanalo-Contreras et al., 2013; Mulligan and Chakrabartty, 2013; Hipp et al., 2014). More than 20 different proteins are known so far to be involved in these diseases referred to as amyloidoses. Furthermore, many other diseases [such as cancer or cardiovascular disease (CVD)] are caused by the misfolded and therefore dysfunctional proteins (Iakoucheva et al., 2002; Cheng et al., 2006; Uversky, 2008a, 2009, 2014; Uversky et al., 2008, 2009; Uversky and Dunker, 2010). Therefore, natural proteins can be found in one of three major protein forms, functional and folded, non-functional and misfolded, and functional and intrinsically disordered.

The structural plasticity and conformational adaptability of IDPs/IDPRs, their ability to react easily and quickly in response to changes in their environment, and their binding promiscuity and unique capability to fold differently while interacting with different binding partners (Dyson and Wright, 2005; Oldfield et al., 2008) define a wide set of functional advantages of IDPs/IDPRs over the ordered proteins (Uversky and Dunker, 2010; Cozzetto and Jones, 2013; Ferreon et al., 2013). These factors determine the abundant involvement of IDPs/IDPRs in various signaling, regulation, and recognition processes. They also allow these flexible proteins to play diverse roles in modulation and control of functions of their binding partners and in promotion of the assembly of supra-molecular complexes. In fact, IDPs are promiscuous binders and can form highly stable complexes, or be involved in signaling interactions where they undergo constant "bound-unbound" transitions, thus acting as dynamic and sensitive "on-off" switches. The ability of these proteins 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 unique physiological properties of IDPs which define the ability of these proteins to exert different functions in different cellular contexts according to a specific conformational state (Uversky and Dunker, 2010).

Furthermore, biological activities of IDPs/IDPRs are known to be precisely and tightly controlled and regulated by extensive posttranslational modifications (PTMs), such as phosphorylation, acetylation, glycosylation, etc. (Collins et al., 2008; Uversky and Dunker, 2010; Kurotani et al., 2014; Pejaver et al., 2014), and by alternative splicing (AS) (Romero et al., 2006; Buljan et al., 2012, 2013). The ability of AS to generate extended sets of protein isoforms with highly diverse regulatory elements (Romero et al., 2006) 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 multifunctionality of these proteins (Uversky, 2013). Clearly, AS-driven removal of pieces of the IDP/IDPRs sequence containing different functional elements could dramatically reshuffle such multifunctionality (Buljan et al., 2012, 2013; Colak et al., 2013).

Careful examination of different proteomes and various large protein datasets (e.g., UniProt database, http://web.expasy.org/docs/swiss-prot_guideline.html) by various disorder predictors revealed that IDPs are not some obscure and rare exceptions from the general "one sequence—one unique structure—one unique function" paradigm (Wright and Dyson, 1999; Uversky et al., 2000; Dunker et al., 2001; Tompa, 2002), but, in fact, are highly abundant in nature (Dunker et al., 2000; Ward et al., 2004; Tompa et al., 2006; Krasowski et al., 2008; Shimizu and Toh, 2009; Tokuriki et al., 2009; Pentony and Jones, 2010; Tompa and Kalmar, 2010; Uversky, 2010a; Xue et al., 2010a, 2012a; Dyson, 2011; Schad et al., 2011; Hegyi and Tompa, 2012; Korneta and Bujnicki, 2012; Midic and Obradovic, 2012; Pancsa and Tompa, 2012; Di Domenico et al., 2013; Kahali and Ghosh, 2013). These natural abundance of IDPs and IDPRs are now documented in multiple entries populating various disorder-related databases, such as DisProt (Vucetic et al., 2005; Sickmeier et al., 2007), $D^2P^2$ (Oates et al., 2013), Ideal (Fukuchi et al., 2012, 2014), MobiDB (Di Domenico et al., 2012), ComSin (Lobanov et al., 2010), and pE-DB (Varadi et al., 2014). The overall amount of IDPs and IDPRs in various proteomes increases with the increase in the organism's complexity, with over half of all eukaryotic proteins predicted to contain long IDPRs (Dunker et al., 2000; Ward et al., 2004; Oldfield et al., 2005; Uversky, 2010a; Xue et al., 2012a). It is believed that the potential explanation for this trend has its roots in a change in the cellular requirements for certain protein functions, particularly for proteins involved in cellular signaling. In agreement with this hypothesis, a computational analysis revealed that the majority of known eukaryotic signaling proteins contain significant regions of disorder (Dunker et al., 2002; Iakoucheva et al., 2002).

Many individual IDPs and hybrid proteins (e.g., α-synuclein, p53, PTEN, etc.) are known to interact with large number of unrelated partners, thereby serving as hubs in cellular protein-protein interaction networks (Dunker et al., 2005; Uversky et al., 2005). Also, the binding regions of partner proteins interacting

with structured hubs, such as 14-3-3 and calmodulin, are often intrinsically disordered (Bustos and Iglesias, 2006; Radivojac et al., 2006). Therefore, there are two general mechanisms by which intrinsic disorder is exploited in protein-protein interaction networks, where one IDP/IDPR binds to many partners and many IDPs/IDPRs interact with one partner (Dunker et al., 2005; Uversky et al., 2005). Subsequent comprehensive bioinformatics studies supported the hypothesis that hub proteins commonly use disordered regions to bind to multiple partners (Dosztanyi et al., 2006; Ekman et al., 2006; Haynes et al., 2006; Patil and Nakamura, 2006; Singh and Dash, 2007; Singh et al., 2007).

## IDPs IN HUMAN DISEASES

### A GENERAL OVERVIEW OF THE IDP INVOLVEMENT IN PATHOGENESIS OF VARIOUS MALADIES

Since IDPs/IDPRs are very common in all proteomes studied to date, possess numerous crucial functions, are promiscuous binders, are abundantly involved in signaling, control, and regulation of important biological processes, where it is crucial for a given protein to be available in appropriate amounts and not to be present longer than needed, IDPs and hybrid proteins have to be tightly regulated and controlled themselves. To check this hypothesis, careful analysis of the IDP regulation inside the cell at different stages of protein synthesis and degradation was recently conducted using the corresponding data available for the *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Homo sapiens* proteomes (Gsponer et al., 2008). This analysis revealed that IDPs were less abundant than ordered proteins in these tree proteomes because of the increased decay rates of mRNAs encoding IDPs, lower rates of IDP protein synthesis, and shorter half-lives of IDPs (Gsponer et al., 2008). Also, the majority of IDP-targeting kinases were either regulated in a cell-cycle dependent manner, or were activated upon exposure to specific stimuli or stress (Gsponer et al., 2008), thereby adding another level of the IDP control. It is important to remember though that although the abundance of many IDPs is tightly regulated, some disordered, and hybrid proteins are present in cells in large amounts [for example, α-synuclein constitutes ∼1% of the total soluble protein in the brain (Iwai et al., 1995)] or/and for long periods of time due to either specific PTMs, or via interactions with other factors, or due to the localization is specific compartments (Uversky and Dunker, 2010). These events could promote changes in cellular localization of IDPs or protect them from degradation (Dunker et al., 2001; Iakoucheva et al., 2004; Tompa, 2005; Grimmler et al., 2007). Therefore, accumulated data clearly show that the chaos seemingly associated with highly flexible and promiscuous IDPs/IDPRs is under tight control (Uversky and Dunker, 2008).

In addition to mentioned bioinformatics analyses numerous experimental studies not only emphasize the important roles of disordered regulators in signaling (Mitrea and Kriwacki, 2013; Follis et al., 2014), regulation (Galea et al., 2008a; Wang et al., 2011; Follis et al., 2012; Mitrea et al., 2012; Ou et al., 2012; Yoon et al., 2012; Frye et al., 2013; Moldoveanu et al., 2013), cell protection (Mei et al., 2014), protein protection (De Jonge et al., 2009; Chakrabortee et al., 2012), and cellular homeostasis (Norholm et al., 2011; Follis et al., 2013), but also show that IDPs/IDPRs are

concisely controlled by themselves via multiple mechanisms, such as interaction with chaperones (Rudiger et al., 2002; Martinez-Yamout et al., 2006; Rodriguez et al., 2008; Didenko et al., 2012; Karagoz et al., 2014) or nanny proteins (Tsvetkov et al., 2009a), partner binding (Demarest et al., 2002; Chipuk et al., 2008; Ebert et al., 2008; Ferreon et al., 2009a,b; Wojciak et al., 2009; Lee et al., 2010), various PTMs (Kostic et al., 2006; Grimmler et al., 2007; Zhan et al., 2007; Galea et al., 2008b; Mitrea et al., 2014), and regulated degradation (Asher et al., 2005; Tsvetkov et al., 2008, 2009b, 2012; Suskiewicz et al., 2011; Wiggins et al., 2011).

Obviously, when tightly controlled process is suddenly coming out of control, consequences could be disastrous. In agreement with this statement, numerous cases are known where the malfunction of a protein (which could be ordered or intrinsically disordered) is associated with the development of particular pathological conditions, and a broad range of human diseases is linked to the failure of a specific peptide or protein to adopt its functional conformational state. Each of these conformational diseases originates from the dysfunction of a particular protein. The failure of such a protein to adopt, possess, or keep functional state is commonly associated with protein misfolding, loss of normal function, gain of toxic function, and/or protein aggregation (Uversky et al., 2009; Uversky, 2010b). Some disease-related proteins have an intrinsic propensity to form pathologic conformation(s). For other proteins, interactions or impaired interactions with chaperones, intracellular or extracellular matrices, other proteins, small molecules, and other endogenous factors can induce conformational changes and increase propensity to misfold. Often, misfolding, and dysfunction originate from point mutation(s) or result from a protein exposure to internal or external toxins. Furthermore, such conformational diseases can also be caused by impaired PTMs, such as phosphorylation, advanced glycation, deamidation, racemization, etc., an increased probability of degradation, impaired trafficking, loss of binding partners, or oxidative damage. All these factors can act independently, additively, or synergistically (Uversky, 2012).

It is recognized now that IDPs and hybrid proteins with long IDPRs are commonly involved in human diseases. For example, some well-known cancer-related proteins with experimentally confirmed IDPRs include p53 (Wells et al., 2008), BRCA1 (Mark et al., 2005), EWS (Ng et al., 2007), HPV protein (Uversky et al., 2006), PTEN (Malaney et al., 2013a), axin (AXis Inhibition) (Noutsou et al., 2011; Xue et al., 2012b, 2013), adenomatous polyposis coli (APC) protein (Xue et al., 2012b; Minde et al., 2013), apoptosis-stimulating proteins of p53 (ASPPs) (Rotem et al., 2007; Ahn et al., 2009), BH3-only proteins (Kvansakul and Hinds, 2014), sirtuins (McBurney et al., 2013; Sharma et al., 2013), CBP/?p300 [CREB-binding protein (CBP) and its paralog, E1A-binding protein p300] (Wojciak et al., 2009), cell cycle regulatory proteins p21 and p27 (Mitrea et al., 2012), and many others. Recent computational analysis revealed that a majority of the cancer/testis antigens (CTAs) are typical IDPs (Rajagopalan et al., 2011). CTAs constitute a heterogeneous protein group, members of which are typically expressed in the normal testis and aberrantly expressed in several types of cancer (Rajagopalan et al., 2011). Among the neurodegeneration-related proteins is such well-characterized IDPs as a protein-chameleon

α-synuclein that can adopt a variety of different conformations, starting from random coil and ending with a more compact molten globular state, or even with poly-(L-proline) II-like conformations, depending on the cellular environment (Uversky, 2003), and aggregates of the α-synuclein are accumulated in Parkinson's disease, dementia with Lewy bodies, Alzheimer's disease (AD), Down's syndrome, and several other syncleinopathies (Uversky, 2007, 2008b; Uversky and Eliezer, 2009; Breydo et al., 2012; Alderson and Markley, 2013). Other IDPs implicated in neurodegenerative diseases include amyloid β and tau proteins (AD), prions (Creutzfeldt-Jakob disease, scrapie, bovine spongiform encephalopathy), and ataxin (spinocerebellar ataxia) (Uversky, 2009, 2014). Other human diseases with well-established pathogenic IDPs are CVDs (hirudin and thrombin) (Cheng et al., 2006), type II diabetes (amylin) (Uversky et al., 2008), AIDS (HIV Rev protein) (Casu et al., 2013), and cystic fibrosis (CFTR) (Baker et al., 2007).

## COMPUTATIONAL APPROACHES FOR ESTIMATING THE IDP ABUNDANCE IN DIFFERENT DISEASES

The fact that IDPs and hybrid proteins with long IDPRs are commonly associated with the pathogenesis of various human disorders gave rise to the "disorder in disorders" or $D^2$ concept (Uversky et al., 2008). Here, the dysfunction, misidentification, misregulation, misfolding, and missignaling of causative IDPs/IDPRs are all can be considered as causative factors of the conformational diseases (Uversky et al., 2008, 2009, 2014; Uversky, 2009, 2010b, 2012; Midic et al., 2009a,b). Although the fact of the involvement of individual IDPs and hybrid proteins with long IDPRs in the pathogenesis of some human maladies is well-established, the application of specially designed computational and bioinformatics protocols opens a unique opportunity for the accurate evaluation of the abundance of IDPs in various pathological conditions.

The first of these computational approaches is based on the assembly of specific datasets of proteins associated with a given disease and the computational analysis of these datasets using a number of disorder predictors (Iakoucheva et al., 2002; Cheng et al., 2006; Uversky et al., 2006; Mohan et al., 2008; Uversky, 2008a, 2009). This approach represents an extension of the analysis of individual proteins to a set of independent proteins. Such analysis revealed that 79% of cancer-associated and 66% of cell-signaling proteins contain predicted regions of disorder of 30 residues or longer (Iakoucheva et al., 2002). Similar analyses revealed that the percentage of proteins with 30 or more consecutive disordered residues was 61% for proteins associated with CVD (Cheng et al., 2006). Many CVD-related proteins were predicted to be entirely disordered, with 101 proteins from the CVD dataset predicted to have a total of almost 200 specific disorder-based binding motifs (thus about 2 binding sites per protein) (Cheng et al., 2006). Finally, the dataset analysis revealed that intrinsic disorder is commonly found in neurodegenerative diseases and diabetes (Uversky et al., 2008, 2009).

In a second approach, the abundance of intrinsic disorder was analyzed in the human diseasome (Midic et al., 2009a), which is a complex network that systematically links the human disease phenome with the human disease genome (Goh et al., 2007). These analyses showed that many human genetic diseases are caused by alterations of IDPs, that different disease classes vary in the disorder contents of their associated proteins, and that many IDPs involved in some diseases are enriched in disorder-based protein interaction sites (Midic et al., 2009a,b).

Finally, a third approach is based on the evaluation of the association between the level of intrinsic disorder and a particular protein function (including the disease-specific functional keywords) in a set of proteins known to carry out this function (Vucetic et al., 2007; Xie et al., 2007a,b). In this approach, it is hypothesized that if intrinsic disorder is important for a function described by a given keyword, then, a greater level of predicted disorder would be found in a protein associated with this keyword than the level of disorder predicted in a protein randomly chosen from the Swiss-Prot (Vucetic et al., 2007; Xie et al., 2007a,b). To test this hypothesis, functional keywords associated with at least 20 Swiss-Prot proteins were found and corresponding keyword-associated datasets of proteins were assembled. For each keyword-associated set, 1000 length-matching and number-matching sets of random proteins were drawn from Swiss Prot. Order-disorder predictions were carried out for the keyword-associated sets and for the matching random sets. If a function described by a given keyword were carried out by a long region of disordered protein, one would expect the keyword-associated set to have a greater amount of predicted disorder compared to the matching random sets. The keyword-associated set would be expected to have less prediction of disorder compared to the random sets if the keyword-associated function were carried out by structured protein. Given the predictions for the function-associated and matching random sets, it is possible to calculate the $p$-values, where a $p$-value $> 0.95$ suggests a disorder-associated function, a $p$-value $< 0.05$ suggests an order-associated function, and intermediate $p$-values are ambiguous (Vucetic et al., 2007; Xie et al., 2007a,b). This analysis revealed that out of 710 Swiss-Prot keywords, 310 functional keywords were associated with ordered proteins, 238 functional keywords were attributed to disordered proteins, and the remainder 162 keywords yield ambiguity in the likely function-structure associations (Vucetic et al., 2007; Xie et al., 2007a,b). Furthermore, many diseases were found to be strongly correlated with proteins predicted to be disordered (Vucetic et al., 2007; Xie et al., 2007a,b). Contrary to this, almost no disease-associated proteins were found to be ordered (Xie et al., 2007a).

## WRECKED REGULATION OF IDPs AND DISEASE

The physiological protein function and the ability to be converted from a normal protein to a pathological form depend on multiple factors which can be grouped into two major classes, genetic and non-genetic. Genetic factors include pathological mutations, aberrant splicing, chromosomal translocation, alternative transcription, and altered AS. Non-genetic factors are related to the peculiarities and levels of protein expression, protein availability, regulation, interaction patterns, cleavage propensity, and PTMs. Some illustrative examples of these transforming factors leading to the appearance of pathological proteins are given below. Since the association between the pathological mutations and IDPs/IDPRs was a subject of several research papers and

reviews, this mechanism of affecting the physiological protein function will not be discussed in this review, and the interested readers are encouraged to look for the corresponding information in the original works (e.g., Joerger and Fersht, 2007, 2008; Uversky, 2008b; Midic et al., 2009a,b; Vacic and Iakoucheva, 2012; Vacic et al., 2012; Yates and Sternberg, 2013). **Figure 1** is a scheme illustrating different ways by which aberrant regulation of IDPs/IDPRs at different levels can result in the development of pathological conditions.
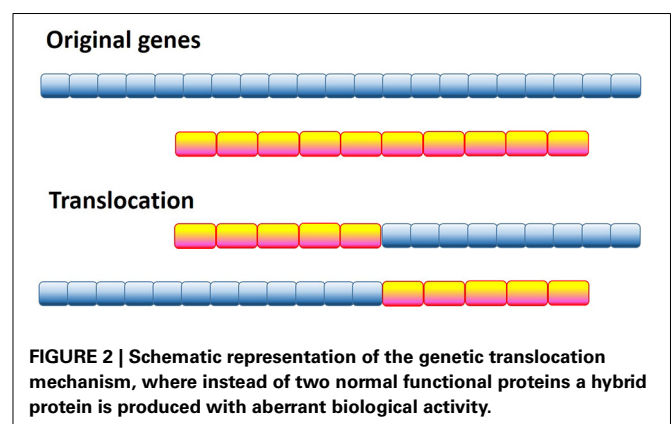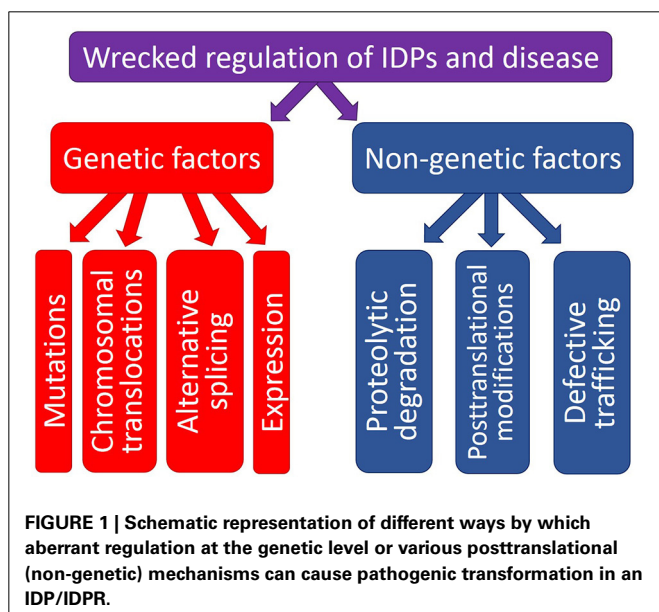
### GENETIC FACTORS: CHROMOSOMAL TRANSLOCATIONS

One of the most radical and obvious ways to generate a pathological protein is chromosomal translocation, which generates chimeric proteins by fusing segments of two otherwise separated genes (see **Figure 2**). Several forms of cancer, such as acute myelogenous leukemia (AML), acute lymphoblastic leukemia (ALL), chronic myelogenous leukemia (CML), chronic myelomonocytic leukemia (CMML), primary myelofibrosis (PMF), anaplastic large cell lymphoma, non-Hodgkin's lymphoma, Ewing's sarcoma (EWS), colorectal cancer (CRC), non-small cell lung cancer, lung adenocarcinoma, and sporadic and radiation-associated papillary thyroid carcinomas are caused by chromosomal translocation. Computational analysis of the 406 translocation-related human proteins revealed that these oncoproteins are significantly enriched in intrinsic disorder, with the translocation breakpoints being mostly located outside the functional domains (Hegyi et al., 2009). Furthermore, the vicinities of the breakpoint were shown to be even more disordered than the rest of these already highly disordered fusion proteins. These observations suggest that high levels of intrinsic disorder represents an important factor that helps fusion proteins to escape detection by cellular surveillance mechanisms that eliminate misfolded proteins and to live long enough to manifest their altered function(s) (Hegyi et al., 2009).

The authors found that these translocation-generated fusions enable the long-range structural communication of remote

binding and/or catalytic domains in the chimeric proteins and thereby define the acquired oncogenic functions. One of the illustrative examples of such acquired oncogenicity is the acquired intramolecular phosphorylation of the BCR-ABL fusion protein related to CML and ALL. Here, chromosomal translocation results in fusion of a Tyr-kinase phosphorylation motif in BCR with the Tyr-kinase domain within ABL, with disorder of the intervening region enabling intramolecular phosphorylation (Hegyi et al., 2009). Another recent example in this category is the BCR-RET chimeric protein originated from the fusion of the catalytic domain of the tyrosine kinase (TK) receptor RET (REarranged during Transfection) with BCR (Ballerini et al., 2012).

The second mechanism of generation of oncoproteins by chromosomal translocation is related to fusion of the dimerization domain with the kinase domain, generating a dimeric hybrid protein. Monomers comprising this protein are engaged in multiple mutual intermolecular phosphorylation reactions that promote auto-activation and generate novel binding sites for signaling proteins. Examples of this mechanism include various chimeric proteins, such as constitutively activated TEL-JAK2 fusion (ETS translocation variant 6 – Janus tyrosine kinase 2 fusion) with kinase activity in human leukemia (Lacronique et al., 1997); TFG-ALK (TRK-fused gene - anaplastic lymphoma kinase fusion) related to anaplastic large cell lymphomas (Roccato et al., 2003); NPM-ALK (nucleolar phosphoprotein nucleophosmin – anaplastic lymphoma kinase fusion), the chimeric protein that is created by translocation in non-Hodgkin's lymphoma and that requires the activation of its ALK kinase function as a result of oligomerization mediated by the NPM segment (Bischof et al., 1997); EML4-ALK (echinoderm microtubule-associated protein-like 4 − anaplastic lymphoma kinase fusion) found in non-small lung carcinoma (Soda et al., 2007); *TPM3-ALK* (Lamant et al., 1999), *TFG-ALK* (Hernandez et al., 1999), and *ATIC-ALK* (Trinei et al., 2000) fusions in anaplastic large cell lymphoma (ALCL) (Marino-Enriquez and Dal Cin, 2013). Other ALK fusion partners include tropomyosin 4 (TPM4); clathrin heavy chain (CLTC); moesin (MSN); the cysteinyl-tRNA synthetase (CARS); RNF213, a ring finger protein; RAN binding protein 2 (RANBP2); myosin II heavy chain type A (MYH9); vinculin (VCL); dynactin (DCTN1); liprin beta 1 (PPFIBP1); KIF5B, a kinasin family member; KLC1,



**FIGURE 1 | Schematic representation of different ways by which aberrant regulation at the genetic level or various posttranslational (non-genetic) mechanisms can cause pathogenic transformation in an IDP/IDPR.**



**FIGURE 2 | Schematic representation of the genetic translocation mechanism, where instead of two normal functional proteins a hybrid protein is produced with aberrant biological activity.**

kinesin light chain 1; and sequestosome 1 (SQSTM1) (Marino-Enriquez and Dal Cin, 2013). In the vast majority of the ALK-based fusions, the chimeric proteins include all 563 cytoplasmic amino acids from ALK (residues 1058–1620), which is the constitutively active C-terminal tyrosine kinase domain (Marino-Enriquez and Dal Cin, 2013), whereas the ALK fusion partner determines the subcellular localization of the fusion oncoprotein via its peptide localization signals and its homotypic or heterotypic oligomerization domains, which in turn will condition the protein-protein interactions and modulate the oncogenic signaling and molecular consequences of the ALK oncoprotein action (Marino-Enriquez and Dal Cin, 2013). Another highly promiscuous chromosomal translocator is RET, the catalytic domain of which, in addition to the aforementioned BCR-RET fusion, is fused with the heterologous oligomerization domains encoded by different genes, such as *H4* (for *RET-PTC1*), *RIα* (*RET-PTC2*), *ELE1* (*RET-PTC3* and *RET-PTC4*), *RFG5* (*RET-PTC5*), *hTIF* (*RET-PTC6*), *RFG7* (*RET-PTC7*), *kinectin* (*RET-PTC8*), *RFG9* (*RET-PTC9*), and *ELKS* (*ELKS-RET*) (Takahashi, 2001), with *KIF5B* (*KIF5B-RET*), or fusion with the fibroblast growth factor receptor 1 (FGFR1) oncogene partner (FGFR1OP-RET) (Bossi et al., 2014). In the case of FGFR1OP-RET fusion, the chimeric protein found in papillary thyroid carcinomas and other myeloproliferative disorders, such as CML, chronic eosinophilic leukemia, chronic neutrophilic leukemia, polycytemia vera, PMF, essential thrombocytosis, myelodysplastic syndromes (MDS), and CMML was shown to display constitutive tyrosine kinase and transforming activity (Bossi et al., 2014).

Chromosomal translocation can affect transcription factors, as illustrated by the EWS-ATF or EWS-Fli1 hybrids, where the DNA-binding elements of transcription factors ATF1 or Fli1 are fused to the disordered transactivation domain of the EWS oncogene to generate an aberrant transcription factor related to Ewing sarcoma (Ng et al., 2007). Another very interesting example of the transcription factor whose function is affected by chromosomal translocation is the ETV6 protein containing two major domains, the HLH (helix-loop-helix, residues 40–124) domain and the ETS domain (residues 340–420) connected by the internal domain (residues 125–339) (De Braekeleer et al., 2012), whose gene spans a region of less than 250 kb at band 12p13.1 and consists of 8 exons (Baens et al., 1996). The importance of this protein is determined by the fact that the *ETV6* chromosomal translocations are among the most commonly observed chromosomal abnormalities in human leukemia and myelodysplastic syndrome (Baens et al., 1996), where there are 48 chromosomal bands involved in *ETV6* translocations, insertions, or inversions, with at least 28 translocations being characterized at the molecular level and 30 *ETV6* partner genes being identified (De Braekeleer et al., 2012). In fact, the *ETV6* gene is known to be fused with a wide array of genes encoding proteins with different functionalities, such as receptor tyrosine kinase genes (e.g., *ETV6-PDGFRB*, *ETV6-PDGFRA*, *ETV6-NTRK3*, and *ETV6-FLT3*), non-receptor tyrosine kinase genes (such as *ETV6-ABL1*, *ETV6-ABL2*, *ETV6-JAK2*, *ETV6-FGFR3*, *ETV6-SYK*, *ETV6-FRK*, and *ETV6-LYN*), transcription factor genes (such as *ETV6-RUNX1*, *MN1-ETV6*, *ETV6-ARNT*, *ETV6-PER1*, and *ETV6-EVI1*), homeobox genes (such as *ETV6-CDX2*, *PAX5-ETV6*, and *MNX1-ETV6*) and many

other genes (e.g., *CHIC2-ETV6*, *ETV6-MDS2*, *TTL-ETV6*, *ETV6-STL*, *ETV6-PTPRR*, *ETV6-NCOA2*, *ETV6-BAZ2A*, *ETV6-GOT1*, *ETV6-FCHO2*, *ETV6-IGH*, and *ETV6-ACSL6*) (De Braekeleer et al., 2012). It is important to emphasize here that the mentioned fusions do not include the full-length ETV6 protein. Furthermore, even fusions of the *ETV6* gene with the same target gene not always will have the same parts of the ETV6 protein. For example, upon fusion to the *PDGFRB* gene encoding a cell surface tyrosine kinase receptor for members of the platelet-derived growth factor family, the resulting *ETV6-PDGFRB* fusion gene in myeloproliferative disorders will have exon 4 of the *ETV6* gene fused in-frame to exon 11 of the *PDGFRB* gene, with the resulting chimeric protein containing the HLH domain of ETV6 [ETS translocation variant 6 (also known as TEL), N-terminal part] and the tyrosine kinase domain of PDGFRB (platelet-derived growth factor receptor beta, C-terminal part), whereas in chronic idiopathic myelofibrosis (CIMF) the fusion of the same two genes will produce in-frame fusion between full-length *ETV6* exon 7, 34 bp derived from *ETV6* intron 7 and a truncated *PDGFRB* exon 12, giving rise to the chimeric protein retaining the internal domain of ETV6 which has the ability to bind to corepressors and induce the transcription-repressive activity of ETV6 (De Braekeleer et al., 2012).

To bring the aforementioned data on the chromosomal translocation to the frames of this review, **Figure 3** represents disorder propensities of some of the most promiscuous chromosomal translocators, ALK, RET, and ETV6. **Figure 3A** shows that the C-terminal domain of ALK, which is included in the vast majority of ALK-based chimeric proteins, is predicted to have highly disordered N-terminal and C-terminal tails. **Figure 3B** illustrates that the same statement is also applicable to RET, catalytic domain of which (residues 708–114) is fused to many target proteins and is predicted to be surrounded by long IDPRs. Finally, **Figure 3C** shows that the transcription factor ETV6 is predicted to have two ordered domains (residues 43–145 and 310–420), which coincide with known functional domains of this protein, PNT domain (residues 40–124) and DNA-binding ETS domain (residues 339–420).

## GENETIC FACTORS: ABERRANT RNA SPLICING
### Intrinsic disorder and alternative splicing
Alternative splicing (AS) of pre-mRNAs, which generates two or more protein isoforms from a single gene, is believed to be responsible for tissue specificity of many of the abundant proteins. Estimates indicate that between 35 and 60% of human genes yield protein isoforms by means of AS mRNA (Stamm et al., 2005). Regions of mRNA affected by AS were shown to correspond to protein regions enriched in intrinsic disorder (Romero et al., 2006). The finding that the alternatively spliced regions of mRNA encode IDPRs with greater frequencies than structured regions suggests a link between AS and signaling by IDPRs. This connection constitutes a plausible mechanism that could underlie and support cell differentiation, which ultimately gave rise to the multicellular eukaryotic organisms (Romero et al., 2006). Furthermore, associating AS with protein disorder enables time- and tissue-specific modulations of protein functions. Since disorder is frequently utilized in protein binding regions, having
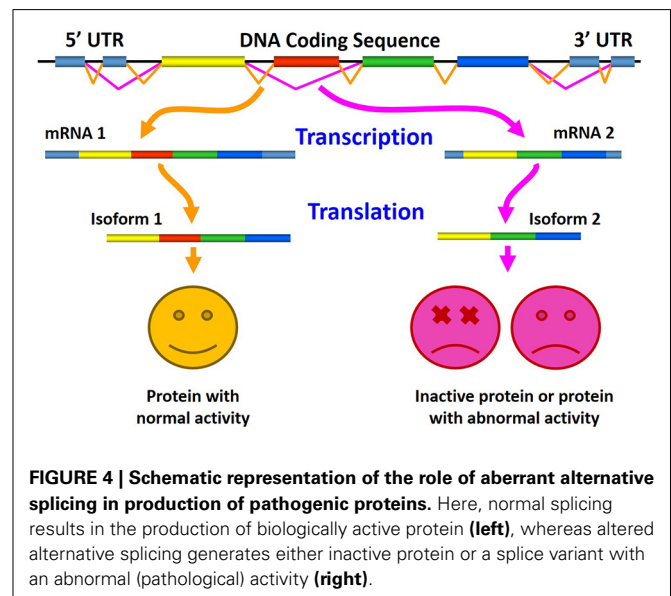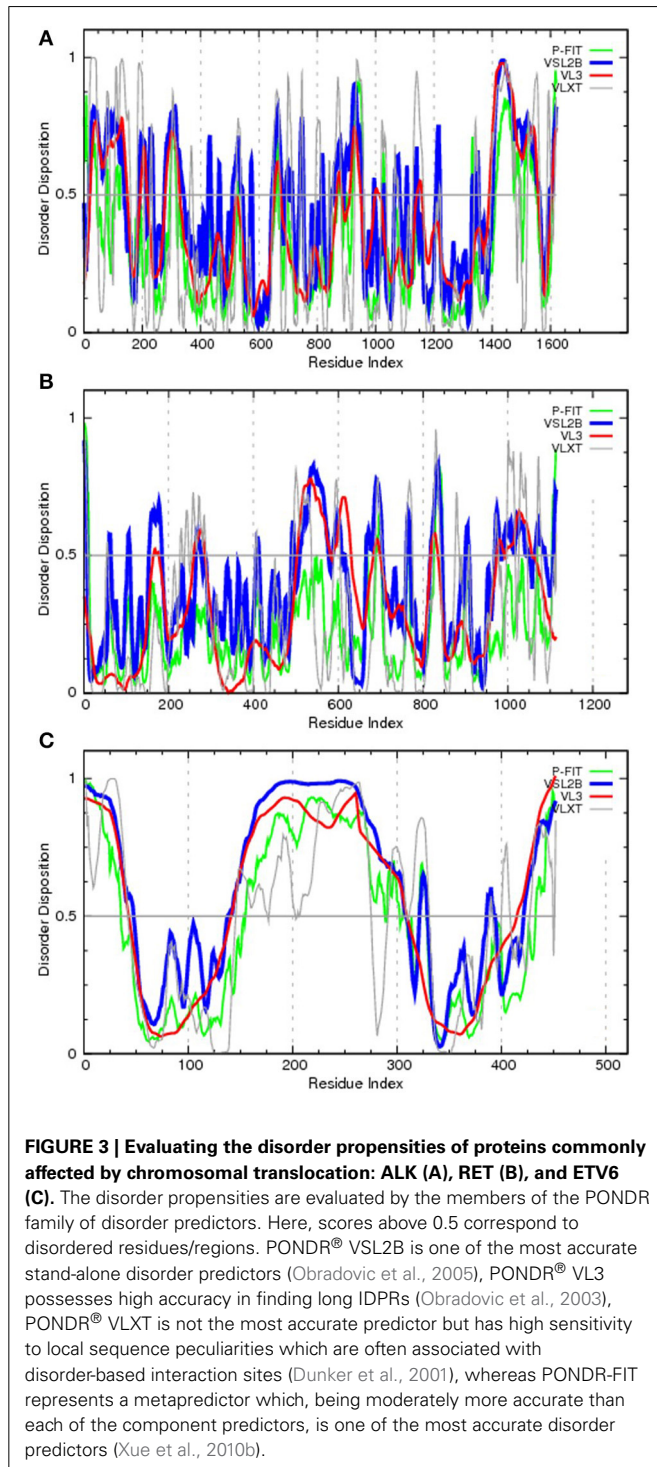
AS of pre-mRNA coupled to regions of protein disorder can lead to tissue-specific signaling and regulatory diversity (Romero et al., 2006; Weatheritt and Gibson, 2012). In agreement with this hypothesis, recent bioinformatics analysis clearly showed that tissue-specific AS of IDPRs with embedded binding motifs is responsible for rewiring of protein interaction networks and signaling pathways (Buljan et al., 2012, 2013). Also, in the human proteome, two types of conserved IDPRs, conserved disorder (i.e., disordered regions where the amino acid sequence is well conserved) and flexible disorder (i.e., disordered regions where the amino acid sequence has diverged) were shown to be highly abundant in proteins with regulatory functions and were highly enriched in regions of proteins that undergo tissue-specific AS (Colak et al., 2013).

### Altered alternative splicing and diseases

The role of aberrant AS in generating pathogenic proteins is shown in **Figure 4**, where the normal splicing results in the production of normally active protein (left), and where altered AS generates either inactive protein or a splice variant with pathological or abnormal activity (right).

***Aberrant alternative splicing in cancer.*** Although the flexibility of AS constitutes an evolutionary advantage for higher eukaryotes, it also represents a risk. In fact, strong evidence indicates that defective AS regulation correlates with the onset and progression of human cancers (Lee et al., 2006; David and Manley, 2010; Pal et al., 2012; Chen and Miller, 2013; Shkreta et al., 2013; Biamonti et al., 2014; Chen and Weiss, 2014), and many cancer-associated genes are regulated through AS suggesting a significant role of this post-transcriptional regulatory mechanism in the production of oncogenes and tumor suppressors (Bonomi et al., 2013). Therefore, not surprisingly, aberrant AS is considered now as an important hallmark of cancer (Ladomery, 2013). The phenomenon of cancer-associated (or cancer-promoting) aberrant splicing is widespread. For example, such aberrant AS events have been found in tumor suppressor LKB1, reduced levels of which are found in patients with Peutz-Jeghers Syndrome (PJS), an autosomal dominant disorder associated with gastrointestinal polyposis and an increased cancer risk, oncogene KIT associated with gastrointestinal stromal tumors, cell-cell adhesion protein CDH17 overexpressed in hepatocellular carcinomas (HCCs) and in gastric and pancreatic cancer, and a Kruppel-like Zn-finger



**FIGURE 3 | Evaluating the disorder propensities of proteins commonly affected by chromosomal translocation: ALK (A), RET (B), and ETV6 (C).** The disorder propensities are evaluated by the members of the PONDR family of disorder predictors. Here, scores above 0.5 correspond to disordered residues/regions. PONDR® VSL2B is one of the most accurate stand-alone disorder predictors (Obradovic et al., 2005), PONDR® VL3 possesses high accuracy in finding long IDPRs (Obradovic et al., 2003), PONDR® VLXT is not the most accurate predictor but has high sensitivity to local sequence peculiarities which are often associated with disorder-based interaction sites (Dunker et al., 2001), whereas PONDR-FIT represents a metapredictor which, being moderately more accurate than each of the component predictors, is one of the most accurate disorder predictors (Xue et al., 2010b).



**FIGURE 4 | Schematic representation of the role of aberrant alternative splicing in production of pathogenic proteins.** Here, normal splicing results in the production of biologically active protein **(left)**, whereas altered alternative splicing generates either inactive protein or a splice variant with an abnormal (pathological) activity **(right)**.

transcription factor KLF6 that functions as a tumor suppressor (Srebrow and Kornblihtt, 2006), as well as in genes implicated in tumor progression (for example, CD44, MDM2, and FHIT) and in susceptibility to cancer [for example, APC (Adenomatous polyposis coli protein)] (Kalnina et al., 2005).

Also, ∼29% of genome-wide expressed genes were shown to be differentially and recurrently spliced in AML patients compared to healthy individuals (Adamia et al., 2014). Among these differentially spliced genes were genes encoding several oncogenes, tumor suppressor proteins, splicing factors and heterogeneous nuclear ribonucleoproteins, proteins involved in apoptosis, cell proliferation, and spliceosome assembly (Adamia et al., 2014). Many of these targets of the aberrant splicing in AML, are proteins with known enrichment in intrinsic disorder, such as proteins related to apoptosis (Peng et al., 2013) as well as proteins involved in spliceosome assembly (Korneta and Bujnicki, 2012; Coelho Ribeiro Mde et al., 2013). Some of the crucial proteins affected by AS in prostate cancer [e.g., the androgen receptor (AR), zinc finger transcription factor Kruppel-like factor 6, Bcl-x, and cyclin D1] (Sette, 2013) are known to contain IDPRs (McEwan, 2012; Peng et al., 2013).

Three major players in the prevention of breast cancer are p53, BRCA1, and PTEN. Mutations in the p53, BRCA1, and PTEN genes account for about 10% of familial breast and ovarian cancer cases overall (Muggia et al., 2011; Okumura et al., 2011). Furthermore, genes of these proteins undergo extensive AS events and some splicing variants of these tumor suppressor genes are associated with cancer (Okumura et al., 2011). These three proteins are known to possess functionally important IDRPs. Since they are among the well-characterized protein with disorder, they are considered in more details below.

Although p53 is best known for its role as a tumor suppressor and guardian of the genome (Lane, 1992), besides induction of cell-cycle arrest, apoptosis, or DNA repair, it is also involved in many other cellular processes, including senescence and differentiation (Vousden and Prives, 2009). p53 gene is expressed as multiple isoforms (with sometimes antagonistic functions) due to AS, alternative promoter usage and alternative initiation of translation (Bourdon, 2007). The human p53 gene transcribes multiple splice variants, which are differentially expressed in human breast tumors compared with normal breast tissue (Okumura et al., 2011) and also are differentially expressed from tumor to tumor (Bourdon et al., 2005). There are four domains in p53: the unfolded N-terminal translational activation domain, the structured central DNA binding domain, and the unstructured C-terminal tetramerization and regulatory domain (Oldfield et al., 2008; Uversky and Dunker, 2010). At the transactivation region, p53 interacts with TFIID, TFIIH, Mdm2, RPA, CBP/p300, and CSN5/Jab1 among many other proteins (Anderson and Appella, 2004). At the C-terminal domain, it interacts with GSK3β, PARP-1, TAF1, TRRAP, hGcn5, TAF, 14-3-3, S100B(ββ) and many other proteins (Anderson and Appella, 2004). p53 serves as a unique example illustrating how intrinsic disorder brings about binding promiscuity, since the same short segment near the p53 C-terminus, being disordered in its unbound form, was shown to bind to four different partners and to adopt different bound conformations. In fact, this disordered segment adopts an α-helix, a β-strand, and two irregular structures with very different morphology upon binding to its four different partners (Oldfield et al., 2008; Uversky and Dunker, 2010).

Another important player in the controlling the breast cancer development is BRCA1 (breast cancer type 1 susceptibility protein), which is the 1863 amino acids protein with an amino terminal zinc ring finger motif, two nuclear localization signals, and two C-terminally located BRCT domains. Comprehensive structural and functional analysis revealed that the long central region of BRCA1 (residues 170–1645) represents an intrinsically disordered scaffold for multiple protein-protein and protein-DNA interactions (Mark et al., 2005). Dysregulation of AS of BRCA1 has also been indicated to be involved in the formation of the breast and ovary tumors (Orban and Olah, 2001; Okumura et al., 2011).

Finally, a few words about PTEN (phosphatase and tensin homolog, deleted on chromosome TEN), which is a tumor suppressor on 10q23.3 mutated in many types of cancers (Blumenthal and Dennis, 2008; Carracedo et al., 2011; Hollander et al., 2011). In fact, PTEN is the second most frequently mutated tumor suppressor gene after p53 (Uversky et al., 2014). Also, differential expression of PTEN and its splice variants can play a role in the pathogenesis of sporadic breast cancers (Okumura et al., 2011). For example, the PTEN splice variants retained intron 3 regions (SV3) and intron 5 regions (SV5) have been found in breast cancers (Okumura et al., 2011). Structurally, PTEN consists of the intrinsically disordered N-terminal phosphatidyl inositol (Dunker et al., 2001; Tompa, 2002)-bisphosphate (PIP2) binding module (PBM), which allows PTEN to anchor to the membrane (its site of action) via PIP2 molecules; the dual-specificity lipid and protein phosphatase domain (PD), which contains a conserved catalytic motif HCKAGKGR; the C2 domain, consisting of positively charged residues that help PTEN to associate with phosphatidylserine residues in the plasma membrane; and the C-terminal intrinsically disordered region (C-tail), which regulates membrane association and stability through several phosphorylation modifications (Malaney et al., 2013a). A longer variant of PTEN (PTEN-long) contains additional 173 amino acids at its N-terminus (the N-173 region). PTEN-long is a product of alternative translation initiation at a non-canonical start site (CTG), which is 519 bps upstream of the canonical ATG (Hopkins et al., 2013). The N-173 region is similar to the viral cell-penetrating protein Tat and allows PTEN-long to be secreted from and taken up by cells at distant locations in the body (Hopkins et al., 2013). A poly-alanine stretch within the N-173 region is essential to its secretion, while a poly-arginine stretch is critical for uptake (Hopkins et al., 2013). This N-173 region is largely disordered (Malaney et al., 2013b).

In CRCs, the AS takes place that affects a c-myc transcriptional suppressor, the far-upstream element-binding protein-interacting repressor [FIR, also known as Poly(U)-binding-splicing factor PUF60], that lack the transcriptional repression domain within exon 2 (FIRΔexon2) in CRCs (Rahmutulla et al., 2013). FIR is predicted to possess two long IDPRs (residues 1–126 and 300–500), with the N-terminal disordered domain being missing in AS isoforms. The high prevalence

of AS in transforming growth factor beta receptor 2 (TGFBR2) has been recently reported in oral squamous cell carcinoma (OSCC) and in potentially malignant oral disorders (PMDs), with almost no such event detected in normal tissues (Sivadas et al., 2014). The AS of TGFBR2 occurred at the non-canonical splice sites and resulted in a selective exclusion of TGFBR2 mRNA sequences that code for TGFBR2 regions involved in tumor suppressive interactions such as ATM, CLK2, and CDC2, whereas the TGFBR2 sites related to the tumor promoting interactions such as interactions with MAPK1, PDK1, and PRKCZ were maintained in these aberrant transcripts (Sivadas et al., 2014).

As far as potential molecular mechanisms of the aberrant AS in cancer are concerned, recent study of human primary and metastatic colon cancer samples clearly indicated that the cancer proliferation can be driven by the overexpression or amplification of a component of the tri-small ribonucleoprotein (trisnRNP) spliceosome complex, the pre-mRNA processing factor 6 (PRPF6), which commonly occurs in colon tumorigenesis (Adler et al., 2014; Lokody, 2014). Another recent study indicated that the switch of AS that occurs during epithelial-mesenchymal transition and is correlated with the aggressiveness of breast cancer correlate and breast cancer metastasis is promoted by the heterogeneous nuclear ribonucleoprotein M (hnRNPM), which is an RNA-binding protein that competes with an epithelial splicing regulator ESRP1 for binding to the same cis-regulatory RNA elements (Xu et al., 2014). Also, the activity of the spliceosome (a large macromolecular complex that mediates splicing and AS events) and accessory splicing factors is modulated by their reversible phosphorylation. The kinases and phosphatases involved in these PTMs significantly contribute to AS regulation, to its integration in the complex regulative network that controls gene expression in eukaryotic cells and its aberrations that characterize neoplastic transformation (Naro and Sette, 2013).

Finally, a case of extensive AS of the *TMPRSS2-ERG* gene fusion represents an important illustration of the combined effects of chromosomal translocations and AS (Hu et al., 2008; Sette, 2013). Here, a member of the ETS transcription factor family, ERG that is typically expressed at very low levels in benign prostate epithelial cells is fused with the androgen-responsive TMPRSS2 (transmembrane protease serine 2) to generate a prostate cancer oncogene. The resulting TMPRSS2-ERG hybrid causes abnormally high expression levels of the transcription factor in neoplastic cells. Furthermore, this fusion-derived gene was shown to undergo AS and generated multiple mRNA variants encoding both full-length ERG proteins and isoforms lacking the ETS domain. Notably, an increase in the abundance of transcripts encoding full-length ERG was shown to correlate with less favorable outcomes in prostate cancer patients (Hu et al., 2008).

***Aberrant alternative splicing in neurodegeneration.*** Although numerous studies confirmed the existence of specific differences in AS profiles between normal and cancer tissues (Omenn et al., 2010, 2014; Shkreta et al., 2013), cancers are not the only set of diseases affected/promoted/caused by altered AS. In fact, brain has the greatest amount of AS of all human tissues, and therefore it is quite clear that the aberrant AS of brain proteins is associated with numerous neurodegenerative diseases. For example, in an autosomal dominant neurodegenerative disease called frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17), patients possess a twofold increase in the 4R:3R ratio of tau isoforms [i.e., isoforms containing four (4R) or three (3R) microtubule binding domains, respectively] leading to enhanced aggregation causing the disease (Philips and Cooper, 2000). In spinal muscular atrophy (SMA), constitutive AS of the survival of motor neuron gene (*SMN*) generates the SMNΔ7 isoform lacking the region encoded by exon 7. This SMNΔ7 isoform displays decreased self-oligomerization and is unable to participate in the assembly of small nuclear ribonucleic particles (snRNPs), thereby affecting the biogenesis and localization of spliceosomal snRNPs in the cell and dramatically reducing the ability of cells to produce functional mRNAs (Philips and Cooper, 2000). In Parkinson's disease, six genes, including *PARK2*, *SNCAIP*, *LRRK2*, *SNCA*, *SRRM2*, and *MAPT*, are affected by aberrant AS (Fu et al., 2013). One of these genes, *SNCA*, encodes the most studied neurodegeneration-related IDPs, α-synuclein, aggregation propensity of which is modulated by AS (Beyer, 2006; Beyer and Ariza, 2013). Furthermore, a link between the α-synuclein AS and risk of sporadic Parkinson's disease has been established (Pihlstrom and Toft, 2011). Regulated AS of another classical neurodegeneration-related IDP, tau protein, produces multiple isoforms controlling function of this protein in normal brain by influencing tau localization, conformation, post-translational modifications, availability, and affinity for microtubules and other ligands, whereas aberrations in tau splicing directly cause several neurodegenerative diseases, sporadic, and familial "tauopathies," exemplified by AD, frontotemporal dementia with Parkinsonism (FTDP-17), Down syndrome (DS; trisomy 21), and myotonic dystrophy type 1 (DM1) (Andreadis, 2012). In general, the global transcriptome analysis by exon arrays and next-generation sequencing (NGS) techniques (e.g., RNA-Seq) revealed that many of the AS events are associated with neurodegenerative brain diseases (NBDs), such as AD and Parkinson's disease (PD) (Mills and Janitz, 2012).

In acetyltransferase p300 (which is a new component of cellular aggregates in α-synuclein positive Lewy bodies of patients affected by PD), a highly disordered regions was found that displays similarities with prion-like domains, is encoded as an alternative spliced variant independently of the acetyltransferase domain, and provides an interaction interface for various misfolded proteins, promoting their aggregation (Kirilyuk et al., 2012). AS of acetylcholinesterase (AChE) produces two neuronal tissue-specific splice variants, the synaptic or tailed form (AChE-T) with the C-terminal extension of the common core by a peptide containing a cysteine, which favors dimerization, and the read-through form (AChE-R), whose C-terminal extension over the common core is shorter than that of AChE-T and lacks cysteine (Zimmermann, 2013). AChE-T and its C-terminal peptides T14 and T30 were suggested to play a pivotal, non-hydrolytic role in neurodegeneration (Greenfield et al., 2008; Halliday and Greenfield, 2012).

The high degree of AS in the brain is associated with expression of a larger number of splicing regulators than most other tissues (de la Grange et al., 2010). Among these brain-specific splicing regulators are the members of the CUG-BP, Elav-like

family (CELF), also known as Bruno-like (Brunol) proteins. Six CELF proteins (CELF1-6) and their AS isoforms are responsible for the regulation of AS of transcripts involved in neuronal function, and the dysregulation of CELF-mediated AS in the brain has been implicated in the pathogenesis of some neurological disorders (Ladd, 2013).

***Aberrant alternative splicing in cardiovascular diseases.*** Multiple CVDs, such as cardiomyopathies, arrhythmias, and a number of inherited heart-related conditions are developed due to the altered AS of sarcomeric genes, ion channels, and cell signaling proteins (Lara-Pezzi et al., 2013). For example, clusterin (CLU) or apolipoprotein J is a chaperone for misfolded proteins that might contribute to survival by reducing oxidative stress. CLU maintains partially unfolded proteins in a state appropriate for subsequent refolding by other chaperones, such as HSPA8/HSC70. Several AS isoforms of CLU are encoded from a single gene located on chromosome 8 in humans. These isoforms are ubiquitously expressed in the tissues, and have been implicated in aging, neurodegenerative disorders, cancer progression, and metabolic/ CVDs including dyslipidemia, diabetes, atherosclerosis, and myocardial infarction (MI) (Park et al., 2014). CLU is predicted to have several long IDPRs (residues 1–110, 280–350, and 400–449). The highly disordered N-terminal domain is the CLU region most affected by AS. AS of the *IL1RL1* gene encoding interleukin-1 receptor-like 1 or ST2 protein produces two forms of the protein, a transmembrane (ST2L) form and a soluble form (sST2), the levels of which are elevated in serum of CVD patients with heart failure or MI (Willems et al., 2012). Aberrant AS of the L-type calcium channel (LTCC, a membrane ion channel most responsible for calcium entry and excitation-contraction coupling) is among genetic factors favoring Timothy syndrome, which is a malady characterized by multiorgan dysfunction including lethal arrhythmias, webbing of fingers and toes, congenital heart disease, immune deficiency, intermittent hypoglycemia, cognitive abnormalities, and autism (Splawski et al., 2004).

At the level of individual genes, the AS-produced isoforms were shown to play diverse roles during heart failure (Lara-Pezzi et al., 2012). Some of the illustrative examples include: the AS-driven generation of two non-functional variants of the sodium channel SCN5A (which, in the norm, controls cardioprotection by ischemic preconditioning) during heart failure (Shang et al., 2007); downregulation of the AS variant of the cell-cycle-regulated kinase (CCRK, which promotes cardiomyocyte growth and survival) in heart failure (Qiu et al., 2008); production of the AS variants of the troponin I gene that reduce contraction efficiency (Feng and Jin, 2010); and appearance of the AS-generated titin isoforms affecting the cardiac stiffness in individuals with dilated cardiomyopathy (DCM) (Makarenko et al., 2004); and the improvement of cardiac function after infarction by the CnAβ1 AS variant of the phosphatase calcineurin (Felkin et al., 2011). Curiously, troponin I (Hoffman and Sykes, 2008), titin (Ma et al., 2006), and calcineurin (Dunker et al., 2001) are all known to possess functionally important IDPRs.

***Altered alternative splicing in diabetes.*** Wolframin is an 890 amino-acid glycoprotein predominantly localized in the endoplasmic reticulum (ER) encoded by the *WFS1* gene. Various alterations in *WFS1* gene including mutations and AS are associated with autosomal recessive Wolfram syndrome, autosomal dominant low frequency sensorineural hearing impairment (LFSNHI) DFNA6/14, some psychiatric diseases, and diabetes mellitus (Cryns et al., 2003). In type 2 diabetes mellitus, the resistance to insulin is assumed to be caused by changes in the abundance of the AS isoforms of insulin receptor (Sesti et al., 2001). Normally, the insulin receptor exists in two isoforms, which are characterized by the absence (Ex11$^-$) or presence (Ex11$^+$) of a 12 amino acid sequence at the C-terminal tail of α-subunit caused by the AS of exon 11. The Ex11$^-$ form binds insulin with two-fold higher affinity than the Ex11$^+$, a form which is more abundantly expressed in target tissues from the type 2 diabetic patients. This alteration in the abundance of two forms might be related to the insulin resistance and the decreased sensitivity for metabolic actions of insulin in type 2 diabetes (Imai et al., 1997; Sesti, 2000; Sesti et al., 2001).

## GENETIC FACTORS: ALTERED EXPRESSION OF IDPs AND DISEASE

As it was already pointed out, cells have evolved multiple complex mechanisms during transcription and translation to regulate the availability of IDPs (Gsponer et al., 2008). Since IDPs are important players in various signaling and regulatory networks, their tightly controlled availability represents a very important factor for the normal functioning of a healthy cell. It was also proposed that this tight control of the availability of IDPs might provide fidelity in signaling, regulation, and recognition by minimizing the likelihood of unwanted, non-functional interactions, and inappropriate sequestering of proteins into non-physiological protein complexes (Gsponer et al., 2008). In agreement with this hypothesis, a careful analysis of dosage-sensitive genes (i.e., genes which are harmful when over-expressed) revealed that the proteins encoded by these genes are often intrinsically disordered and that these genes are tightly regulated at both mRNA and protein levels, suggesting that this tight control prevents a potentially deleterious increase in protein concentration under physiological conditions (Vavouri et al., 2009; Babu et al., 2011).

### Aberrant expression of IDPs and cancer

In solid tumors there is a markedly altered pH profile compared to normal tissues. One of the important players in corresponding pH regulation is the electroneutral Na$^+$, HCO$_3^-$ cotransporter (SLC4A7, NBCn1), which is significantly over-expressed in human mammary carcinomas (Gorbatenko et al., 2014). Curiously, the N-terminal half of this 1214 residue-long protein is predicted to be mostly disordered. Altered levels of the collapsin response mediator proteins (CRMPs, C-terminal tails of which are predicted to be disordered), were reported for several malignant tumors, such as lung, breast, colorectal, prostate, pancreatic, and neuroendocrine lung cancer (Tan et al., 2014).

Recently, it was pointed out that several N-glycoproteins are differentially expressed in some cancerous diseases (Tuccillo et al., 2014). The illustrative examples of this phenomenon include upregulation of alpha-1-antichymotrypsin in HCC (Ishihara et al.,

2011) and non-small-cell lung carcinoma (NSCLC) (Zeng et al., 2010); down-regulation of insulin-like growth factor binding protein 3 in HCC (Chen et al., 2011) and NSCLC (Zeng et al., 2010); up-regulation of galectin-3-binding protein (Gal3BP or Mac-2 BP) in HCC (Chen et al., 2011) and ovarian cancer (Tian et al., 2011a); and increased expression of periostin in most ovarian cancer subtypes (Tian et al., 2011a) and aggressive prostate cancer (Tian et al., 2011b).

In cancers, high levels of the Sry-containing protein Sox2 are correlated with poor prognosis and increased proliferation of cancer stem cells (Liu et al., 2013). Sox2, a member of the family of high-mobility group transcription factors, is crucial for early development and maintenance of undifferentiated ESCs (embryonic stem cells). Together with other reprogramming transcription factors, Sox2 was predicted to be highly disordered (Xue et al., 2012c). Recently is has been shown that overexpression and gene amplification of Sox2 is associated with the development of squamous cell carcinoma in multiple tissues such as the lung and esophagus (Liu et al., 2013).

Kruppel-like factor 5 (KLF5) is a basic transcription factor binding to GC boxes at a number of gene promoters and regulating their transcription. This transcription factor has essential roles in cell cycle regulation, apoptosis, migration, and differentiation. Expression of KLF5 was found to be abnormal in many cancer types (Dong and Chen, 2009). As many other transcription factors (Liu et al., 2006), KLF5 is predicted to be mostly disordered.

In uterine leiomyoma, the most common tumors found in the women of the reproductive age which is characterized by the excessive extracellular matrix (ECM), levels of integrin-β1 are significantly up-regulated (Chen et al., 2013). Second half integrin-β1 is predicted to be mostly disordered and this prediction is supported by the structural analysis of this protein, since structure is known for residues 21–465 (e.g., PDB ID: 3VI3). The cellular levels of MDMX, which is a heterodimeric partner of MDM2 and a critical regulator of p53, are noticeably elevated in tumors with wild-type p53 (Li et al., 2012). Such elevated levels of MDMX contribute to p53 inactivation. Crystal structures are known for the N-terminal SWIB domain (residues 25–110, PMID: 2VYR) and the C-terminal region necessary for interaction with UBP2 (residues 429–490, PMID: 2VJE), whereas the reminder of MDMX is disordered.

Aberrant expression of various members of the S100 protein family (which is the largest subfamily of calcium binding proteins of EF-hand type containing at least 25 different members) is associated with pathogenesis of various cancers (Chen et al., 2014). Curiously, since these proteins display a unique pattern of tissue/cell type specific expression in the norm, different cancer types were shown to be characterized by opposite alterations in the levels of different members of this interesting family. For example, over-expression of S100A2, S100A3, S100A6, S100A8/A9, and S100A11 is related to several types of cancer, whereas other types of cancer are characterized by the under-expression of these same proteins (Chen et al., 2014). This remarkable variability in the outputs of the deregulation of levels of these proteins can be understood by taking into account their wide functional spectrum, which spreads from to proliferation,

to apoptosis, to metastasis, to tumor microenvironment, and to cancer stem cells (Chen et al., 2014). Another important factor defining this variability is intrinsically disordered nature of S100 proteins (Permyakov et al., 2011). In fact, comprehensive bioinformatics analysis revealed that these proteins are enriched in intrinsic disorder, with 62% of them being predicted to be disordered by at least one of the predictors: 31% are recognized as "molten globules" and 15% are shown to be in extended disordered form (Permyakov et al., 2011).

Various cancers are associated with the altered expression of many subunits of the Mediator complex (Schiano et al., 2014). Mediator is an evolutionarily conserved large proteinaceous machine, needed for growth and survival of all cells. A high prevalence of IDPRs was reported for various subunits of Mediator from both *Saccharomyces cerevisiae* and *Homo sapiens*, especially in the Tail and the Middle modules (Toth-Petroczy et al., 2008). The IDPRs were shown to contribute to function of this important complex in at least three different ways, individually serving as target sites for multiple partners having distinctive structures, acting as malleable linkers between globular domains, and facilitating assembly and disassembly of complexes in response to regulatory signals (Fuxreiter et al., 2008; Toth-Petroczy et al., 2008). Over-expression of five out of 13 subunits of an important translation-related machinery, human translation initiation factor eIF3 (which interacts with the 40S ribosomal subunit, and promotes binding of tRNA$_i^{Met}$ and mRNA), is known to cause malignant transformation (Zhang et al., 2008; Le Quesne et al., 2010).

### Altered expression levels of IDPs in neurodegenerative diseases

An illustrative example of the role of over-expression of a particular protein in the pathogenesis of a given neurodegenerative disease is given by the dual-specificity tyrosine phosphorylation-regulated kinase 1A (DYRK1A) in Down syndrome (Wegiel et al., 2011). In fact, DYRK1A over-expression deregulate multiple pathways in the developing and aging Down syndrome brain, affecting hundreds of proteins, including cytosolic, cytoskeletal, and nuclear proteins, transcription factors (Wegiel et al., 2011). Over-expressed DYRK1A directly leads to the hyperphosphorylation of tau and phosphorylation of AS factor controlling balance between the tau AS isoforms (Wegiel et al., 2011). Although DYRK1A is an enzyme and therefore is expected to be mostly ordered, computational analysis revealed that this important kinase contains two very long disordered tails decorating catalytic domain (residues 1–150 and 500–763). Alterations in expression of highly conserved filamentous proteins septins is associated with variety of neurological conditions, such as Alzheimer's and Parkinson's disease (Peterson and Petty, 2010). Dozens of different human septins are encoded by 14 loci (SEPT1-SEPT14), and all septin proteins contain highly conserved GTPase and polybasic domain regions and highly diverged N- and/or C-termini (Peterson and Petty, 2010), which are predicted to be highly disordered.

An altered expression of brain-derived neurotrophic factor (BDNF, N-terminal half of which is predicted to be disordered) was found in postmortem brains and serum from patients with schizophrenia, Alzheimer's disease and mood disorders

(Carlino et al., 2013). Also, in addition to the abnormal hyper-phosphorylation reported below (see "Aberrant PTMs and neurodegenerative diseases" section below), aberrant expression of the intrinsically disordered tau protein is known to be associated with AD and Down syndrome (Cardenas et al., 2012). The expression levels of glutamate receptor are dramatically altered in regions of the AD brain possessing the greatest pathological changes (Proctor et al., 2011). Another illustration of the role of protein overexpression in neurodegeneration is given by a classical pathology-related IDP, α-synuclein, where multiplications of the *SNCA* gene that encodes this IDP are associated with autosomal dominant PD (Eriksen et al., 2005; Elia et al., 2013). One of the causative agents of SMA, an autosomal recessive neurodegenerative disorder which is a leading genetic cause of infantile mortality, is deletion of a highly disordered protein, survival motor neuron-1 (SMN1) (Lorson et al., 2010).

### Alterations of IDP expression in cardiovascular diseases

In CVDs, changes in expression levels of connexins (Cxs, which are intercellular channels forming low-resistance pathways permitting ions and metabolites up to 1 kDa in molecular mass to flow from cell to cell), such as Cx43, are known to modulate cell-cell coupling and the path of excitation spreading throughout the heart (Zhang and Shaw, 2013). C-tail of Cx43 (residues 230–382) is predicted to be highly disordered. Other proteins with altered expression and activity in major heart diseases, such as ischemic heart disease, cardiomyopathies, and congestive heart failure, are sarcoplasmic reticulum $Ca^{2+}$ ATPases (SERCA) 2a and 2b, which are crucial for recycling cytosolic $Ca^{2+}$ into the lumen of the sarcoplasmic reticulum (SR) (Shareef et al., 2014). Hart failure, which is the end stage of MI, cardiac hypertrophy, and hypertension, is characterized by the down-regulation of the alpha-myosin heavy chain (α-MHC) gene and SERCA genes and reactivation of specific fetal cardiac genes including atrial natriuretic factor and brain natriuretic peptide (Duygu et al., 2013). In addition to SERCA, altered levels and activities of many other $Ca^{2+}$ cycling proteins are related to the heart failure. Among these $Ca^{2+}$ cycling proteins responsible for the release and reuptake of intracellular $Ca^{2+}$ that drives muscle contraction and relaxation are the ryanodine receptor 2, cardiac (RyR2), $Ca^{2+}$ release channel macromolecular complexes, and phospholamban (Marks, 2013).
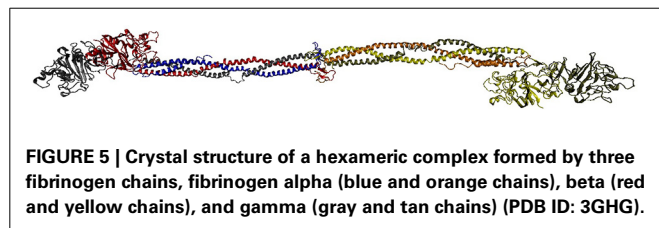
Altered expression of the angiotensin-converting enzyme 2 (ACE2) is associated with cardiac and vascular diseases, such as failing human hearts and atherosclerotic vessels. ACE2 degrades angiotensin II, which is the main effector of the classic renin-angiotensin system (RAS) that plays an important role in the pathophysiology of CVD (Burrell et al., 2013). Elevated levels of plasma plasminogen activator inhibitor-1 (PAI-1) have been shown to precede MI in patients (Juhan-Vague et al., 1996; Ploplis, 2011). In addition for being aberrantly expressed in various cancers, levels of the Kruppel-like factor 5 (KLF5) are altered in all vascular smooth muscle cell (VSMC)-related diseases, such as atherosclerosis, restenosis after angioplasty, cardiac hypertrophy, and hypertension (Dong and Chen, 2009).

Alteration in the composition of ECM proteins, such as claudin-5, occludin, zona occludens proteins (ZO-1 and ZO-2),

increased levels of some ECM proteins (such as osteopontin, chondroitin sulfate proteoglycan neurocan, and fibrinogen), and decreased levels of ECM receptors, such as integrins and dystroglycan, are associated with the altered blood-brain barrier (BBB) found in neurologic diseases, such as stroke (Baeten and Akassoglou, 2011). Osteopontin is an acidic hydrophilic glycophosphoprotein that functions as a cell attachment protein and was predicted to be mostly disordered and shown to be a typical IDP in solution by multidimensional nuclear magnetic resonance spectroscopy, synchrotron radiation circular dichroism spectroscopy, and small-angle X-ray scattering (Platzer et al., 2011). Many of the listed above proteins are predicted to possess long IDPRs. Massive parts of neurocan (residues 350–1050), fibrinogen alpha (residues 100–650), beta (residues 1–250), and gamma chains (flexible residues with disorder scores close to 0.6 are found in regions 65–177 and 318–349), occludin (residues 300–522), ZO-1 (residues 100–400 and 600–1600), ZO-2 (residues 100–500 and 900–1190), and dystroglycan (residues 300–500 and 800–895) are predicted to be highly disordered. In agreement with these predictions, no crystal structure is available for the monomeric forms of any fibrinogen chain, whereas crystal structure was solved for a hexameric complex formed by all three fibrinogen chains (PDB ID: 3GHG, **Figure 5**). Importantly, many residues were missing in this crystal structure (e.g., residues 1–26 and 201–562 in fibrinogen alpha chain, residues 1–57 and 458–461 in fibrinogen beta chain, and residues 1–13 and 395–411 in fibrinogen gamma chain) indicating that the corresponding regions remain high flexibility even in their bound forms.

### Altered expression of IDPs in diabetes

Type 2 diabetes, a heterogeneous disorder with hyperglycaemia caused by impaired insulin secretion and decreased insulin sensitivity is characterized by aberrant expression of several important proteins in β-cells and pancreatic islets (Ostenson and Efendic, 2007). Among the islet proteins, whose levels are altered in type 2 diabetes, are GLUT2, glucokinase, phosphofructokinase, glucose-6-phosphoisomerase, m-glycerophosphate dehydrogenase, pyruvate kinase, KCNJ11/Kir.2, SUR1, PDX-1, Foxo-1, and IRS-2 (Ostenson and Efendic, 2007). Markedly decreased expression of a number of SNARE complex proteins [such as syntaxin-1A, SNAP-25, VAMP-2, nSec1 (MUNC18), synaptophysin, and synaptotagmin V] was also found in pancreatic islets of type 2 diabetic patients (Ostenson et al., 2006). It has been established that the non-bound SNAP-25 is mostly disordered in solution (Fasshauer et al., 1997). In agreement with these data, recent comprehensive bioinformatics analysis revealed that human SNARE proteins possess

**FIGURE 5 | Crystal structure of a hexameric complex formed by three fibrinogen chains, fibrinogen alpha (blue and orange chains), beta (red and yellow chains), and gamma (gray and tan chains) (PDB ID: 3GHG).**

substantial amount of predicted disorder (Pietrosemoli et al., 2013).

One of the hallmarks of the type 2 diabetes is diabetic dyslipidemia associated with the increased production of very low density lipoprotein (VLDL) (Verges, 2010). This abnormality is caused by the aberrant levels of hepatic apolipoprotein B (ApoB), a protein playing crucial role in the VLDL production by the liver (Verges, 2010). ApoB is a very large (4536 residues), mostly ordered protein that is predicted to have several relatively long IDPRs (e.g., residues 299–340 and 506–546). Skeletal muscle of type 2 diabetic patients shows reduced expression of mitochondrial proteins peroxisome proliferator-activated receptor gamma coactivator 1-α (PGC-1α), peroxisome proliferator-activated receptor gamma coactivator 1-β (PGC-1β), and the mitochondrial fusion protein mitofusin 2 (Mfn2) (Zorzano et al., 2010). Both PGC-1α and PGC-1β are predicted to be mostly disordered, and Mfn2 is expected to be mostly ordered, this protein is predicted to have multiple short IDPRs. Abnormal expression of insulin-like growth factor (IGF) binding proteins (IGFBPs) was detected in diabetes and could be used as a sensitive marker of insulin resistance defined as decreased sensitivity and/or responsiveness to metabolic actions of insulin promoting glucose disposal (Ruan and Lai, 2010). Several different IGFBPs are found in the serum, other biological fluids, and tissue extracts. The members of this superfamily of homologous proteins form an important link between the insulin and IGF systems (Ruan and Lai, 2010). Many members of this IGFBP superfamily are predicted to be either mostly disordered or at least possess long IDPRs.

One of the serious complications of diabetes is diabetic nephropathy (Molitch et al., 2003). In fact, 20–30% of patients with type 1 and or type diabetes mellitus possess nephropathy, which develops in 40–50% of patients with a 20-year history of type 1 diabetes mellitus (Thrailkill et al., 2009). One of the hallmarks of diabetic nephropathy is altered expression or activation of matrix metalloproteinases (MMPs), such as collagenases, gelatinases, stromelysins, matrilysins, and membrane-type MMPs, which have the capacity to breakdown all components of the ECM (Thrailkill et al., 2009). Another protein associated with diabetic nephropathy is angiotensin I-converting enzyme (ACE), elevated levels of which are found in this pathology (Lely et al., 2007). Although all MMPs and ACE are enzymes and are expected to be mostly ordered, they do possess regions of predicted disorder.

### NON-GENETIC FACTORS PROMOTING PATHOGENICITY OF IDPs
#### Abnormal posttranslational modifications
**Abnormal PTMs and cancer.** Functions of many IDPs and IDPRs are controlled, modulated, and regulated by various PTMs. Therefore, aberrant PTMs are commonly associated with several human diseases. In fact, all major PTMs, such as glycosylation, phosphorylation, acetylation, ubiquitination, methylation, and palmitylation, have been observed to be altered in cancer, affecting key cellular pathways including signal transduction, cell membrane receptor function, and protein-protein interactions (Markiv et al., 2012). For example, abnormal glycosylation of some glycoproteins due to deregulated glycosyltransferases and

glycosidases is known to be a common phenomenon of many malignancies, including CRC, where elevated levels of the cell-surface α2,6-linked sialic acids have been linked to metastatic spread and therapeutic resistance of this cancer (Park and Lee, 2013). The widespread and diverse PTMs of histones, important nuclear IDPs (Peng et al., 2012) that are crucial for regulated gene expression and for a variety of epigenetic mechanisms, are under very tight and complex spatial and temporal control (Campbell and Turner, 2013). This spatial and temporal regulation of histone modifications is distorted in malignancies on both a genome-wide and discrete gene loci levels (Campbell and Turner, 2013). For example, excessive aberrant acetylation and methylation of specific histone residues have been found in CRC (Gargalionis et al., 2012). Also, alterations of different PTMs at lysine residues (such as acetylation, methylation, ubiquitination, and sumoylation) of proteins involved in DNA repair is often associated with genomic instability, which is the major cause of different diseases, especially cancer (Chatterjee et al., 2012). It has been even proposed that the analysis of histone modifications in circulating nucleosomes can be used for the diagnosis and estimation of prognosis in the CRC patients (Gezer and Holdenrieder, 2014).

Aberrant glycosylation is a well-established event in oncogenesis and cancer progression (Tuccillo et al., 2014), and many biomarkers used for diagnosis, prognosis, and prediction of many cancers are N-linked glycosylated proteins (Drake et al., 2006; Tian and Zhang, 2013). For example, malignant transformation of cells is associated with aberrant glycosylation of mucins, abnormal branching of N-glycans, and increased presence of sialic acid on proteins and glycolipids (Hauselmann and Borsig, 2014).

In addition to translational regulation discussed above, cancer-related activities of Sox2 protein are controlled by various PTMs (Liu et al., 2013). In fact, many sites of Sox2 can be modified through phosphorylation, acetylation, ubiquitination, methylation, and SUMOylation, and based on which sites are modified, Sox2 displays different normal and pathological activities (Liu et al., 2013). Ubiquitination-deubiquitination of a master transcriptional repressor that acts as a tumor suppressor or oncogene in diverse types of cancers, the RE-1 silencing transcription factor (REST) or neuron-restrictive silencer factor (NRSF), controls levels of this protein in cancer, and abnormal upregulation of REST has been found in medulloblastoma, neuroblastoma, and glioblastoma (Huang and Bao, 2012). As many other transcription factors (Liu et al., 2006; Guo et al., 2012), REST/NRSF is predicted to be highly disordered.

It is important to remember that alterations in PTMs of many disease-related proteins are typically produced by alterations of modifying enzymes. For example, aberrant phosphorylation, acetylation, methylation, sumoylation, and ubiquitination of the AR found in prostate cancer is caused by alterations of enzymes that modify the AR (Gioeli and Paschal, 2012). Also, histone acetyltransferases (HATs) and histone deacetylases (HDACs) are two classes of enzymes regulating histone acetylation whose altered activity has been identified in several cancers (Di Gennaro et al., 2004).

**Aberrant PTMs and neurodegenerative diseases.** In Huntington's disease (HD), a genetic neurodegenerative disorder caused by

CAG expansions in the gene encoding Huntingtin protein (Htt), alterations of several histone PTMs are found, including phosphorylation, acetylation, methylation, ubiquitination, and polyamination (Moumne et al., 2013). Various PTMs of Htt itself, such as phosphorylation, sumoylation, ubiquitination, acetylation, proteolytic cleavage, and palmitylation, are also significantly altered in HD, resulting in changes in clinical phenotypes (Ehrnhoefer et al., 2011). In AD, which is a neurodegenerative disorder characterized by the progressive cognitive decline and by accumulation of insoluble aggregates of two proteins in the brain, amyloid-β (Aβ) and the microtubule-associated protein tau, Aβ levels and tau aggregation are impacted by altered sumoylation (Lee et al., 2013). Aberrant phosphorylation of the microtubule-associated protein tau is known to be associated with AD pathology and pathogenesis of other tauopathies (Hernandez and Avila, 2007). In fact, in AD, tau is abnormally hyperphosphorylated to a stoichiometry of at least three-fold greater than normal tau. This hyperphosphorylation is believed to be a major driving force for pathological tau aggregation, leading to the formation of a histopathological hallmark of the disease, paired helical filaments assembled from neurofibrillary tangles. Abnormal hyperphosphorylation and concomitant aggregation of tau is also a characteristic feature of several other tauopathies (Wang et al., 2013). Abnormal hyperphosphorylation of a RNA/DNA binding protein TDP-43 (TAR DNA binding protein 43) was shown to accumulate in the cytoplasm of neuronal cells of patients affected by fronto temporal lobar degenerations (Buratti and Baralle, 2008). TDP-43, which is known to control both normal and pathological RNA splicing events, is predicted to have a long disordered C-tail (residues 250–414). In Huntington disease, mutations in huntingtin (which normally undergoes different PTMs, such as phosphorylation, SUMOylation, ubiquitination, acetylation, proteolytic cleavage, and palmitoylation) significantly alters PTMs leading to changes in the clinical phenotype (Ehrnhoefer et al., 2011).

Prion diseases are characterized by the conversion of a soluble form of the glycoprotein prion protein ($PrP^C$) to the abnormal (infectious) scrapie form ($PrP^{Sc}$) (Gains and LeBlanc, 2007). NMR analysis revealed that human $PrP^C$ is a hybrid protein that contains a C-terminal globular domain extending from residues 125–228, and an N-terminal disordered tail (Zahn et al., 2000). Recent study revealed that the transmissible spongiform encephalopathy (TSE) strain characteristics are affected by the glycosylation status of PrP (Cancellotti et al., 2013). This conclusion is based on the comparative study of the infectious properties of the isolates of three TSE strains passaged through transgenic mice with PrP devoid of glycans at the first, second or both N-glycosylation sites. This analysis revealed that the strain-specific characteristics of the TSE strain changed when $PrP^{Sc}$ was devoid of one or both glycans suggesting that the PTMs of PrP may play a role in altering the infectious properties of a TSE strain (Cancellotti et al., 2013).

***Altered PTMs in cardiovascular diseases.*** Altered PTMs of the voltage-gated Na channel isoform 1.5 (NaV1.5) contribute to the aberrant functions of this important protein associated with acquired cardiac disorders, including arrhythmias and heart failure (Herren et al., 2013). NaV1.5 is the pore forming α-subunit of the voltage-gated cardiac Na channel, which is responsible for the initiation and propagation of cardiac action potentials. Abnormal PTMs affect localization and gating potential of this protein (Herren et al., 2013). Abnormal phosphorylation of myosin light chain and troponins I and T, and, potentially, altered oxidation and glycation of sarcomeric proteins represent very important mechanisms underlying the myofilament dysfunction in DCM (LeWinter, 2005). Earlier, troponin I was shown to possess functionally important disordered regions (Hoffman et al., 2006). Furthermore, subsequent comprehensive bioinformatics analysis revealed that different troponin I isoforms are characterized by different abundance and distribution of intrinsic disorder (Hoffman and Sykes, 2008). This isoform-specific disorder variability was suggested to have potential mechanistic significance being responsible for modulation of the extent to which conformational fluctuations in tropomyosin are communicated to the troponin complex (Hoffman and Sykes, 2008). Troponin T is predicted to be mostly disordered.

Since constitutive nitric oxide (NO) synthases (NOS) are among the crucial factors responsible for the maintenance of myocardial $Ca^{2+}$ homeostasis, myocardial relaxation and distensibility, and protection from arrhythmia and abnormal stress stimuli, the aberrant regulation and dysfunction of these important proteins in hypertension, hemodynamic overload, and atrial fibrillation are known to lead to the production of superoxide instead of NO (Carnicer et al., 2013; Shibata et al., 2013). Deregulation of NOS1 and NOS3 phosphorylation and glutathionylation are potent contributors to vascular disease, myocardial ischemia, reperfusion injury, MI, cardiac hypertrophy, and failure (Carnicer et al., 2013). NOS1 and NOS3 are predicted to have long IDPRs.

***Abnormal PTMs in diabetes.*** Maintenance of metabolic homeostasis is controlled by some nuclear receptors, e.g., by the subfamily of the peroxisome proliferator-activated receptor (PPARs), which consists of three members (PPAR-α, PPAR-γ, and PPAR-δ) encoded by different genes and is involved in a wide range of physiological processes affecting lipid homeostasis, inflammatory responses, adipogenesis, insulin sensitivity, reproduction, wound healing, and carcinogenesis (Escher and Wahli, 2000; Lazennec et al., 2000; Mandard et al., 2004) PPARs has similar structural topology, possessing an N-terminal ligand-independent transactivation domain (AF-1), a two zinc-finger DNA-binding domain, a hinge domain, and a C-terminal ligand-binding domain containing a ligand responsive activation domain (AF-2). Structural analysis of PPAR-γ revealed that this protein has several long regions of missing electron density (PDB ID: 3DZY), such as residues 1–211, 238–298, and 443–477 (Chandra et al., 2008; Fuxreiter et al., 2011). The N-tails of PPAR-α (residues 1–96) and PPAR-δ (residues 1–70) are predicted to be disordered. Curiously, obesity (which one of the major underlying cause of the metabolic syndrome) was shown to promote phosphorylation of PPAR-γ at Ser273 (pSer273), and this PTM is correlated with the dysregulation of a subset of PPAR-γ target genes (such as the insulin-sensitizing adipokines such as adiponectin and adipsin), many of which are dysregulated in obesity (Choi et al., 2010).

Glycation of fibrinogen is a very common PTM detected in diabetes (Hammer et al., 1989; Henschen-Edman, 2001). This is an important finding since even modest degrees of fibrinogen modification can alter the rate of assembly of fibrin monomers into a fibrin clot and the fiber structure and packing, thereby contributing to the increased atherothrombotic risk associated with hyperhomocysteinemia and diabetes (Hoffman, 2008). The peculiarities of disorder distribution in various chains of fibrinogen were already discussed (see section "Alterations of IDP expression in cardiovascular diseases").

Although the modification of nuclear and cytoplasmic proteins at their serine and threonine residues by the post-translational attachment of O-linked N-acetylglucosamine (O-GlcNAc) is crucial for the regulation of various cellular processes, prolonged increases in O-GlcNAcylation and sustained increases in O-GlcNAc levels have been implicated in glucose toxicity and insulin resistance (McLarty et al., 2013). It has been also emphasized that there is a complex interplay between phosphorylation and O-GlcNAcylation, as a result of which proteins can be both O-GlcNAcylated and phosphorylated, since O-GlcNAc and phosphate moieties compete for the same protein residues, serine, and threonine (McLarty et al., 2013). Obviously, this contributes a new level of complexity for protein regulation by PTMs under conditions of globally elevated GlcNAcylation, since increased GlcNAcylation was shown to affect phosphate stoichiometry at most of the sites involved in active phosphorylation-dephosphorylation cycling (Wang et al., 2008). Among 711 phosphopeptides for which phosphorylation dynamics was analyzed, GlcNAcylation resulted in lower phosphorylation at 280 sites and caused increased phosphorylation at 148 sites (Wang et al., 2008). Three hundred and eighty one proteins affected by such dual PTMization possessed a wide spectrum of biological functions, such as chaperones, cytoskeleton regulatory proteins, metabolic enzymes, kinases, transcription factors, and RNA processing proteins (Wang et al., 2008). From the list of functions it is clear that many of these proteins are either IDPs or hybrid proteins possessing long functional IDPRs. In fact, protein involved in chaperon functions, RNA binding and processing, and regulation of cytoskeleton and transcription are all expected to be disordered based on the associative study correlating Swiss-Prot functional keywords and protein intrinsic disorder (Vucetic et al., 2007; Xie et al., 2007a,b), and on existing experimental data for some members of these functional classes.

### Abnormal proteolytic degradation

It is known that proteolytic digestion is orders of magnitude faster in unstructured as compared to structured protein regions (Polverino de Laureto et al., 1995; Fontana et al., 1997, 1986, 2004; Iakoucheva et al., 2001; de Laureto et al., 2006). Therefore, it is extremely important for the protein cleavage process that the sites of cleavage would be located in regions that lack structure or possess high structural flexibility.

Among the multiple modifications affecting aggregation propensity of the intrinsically disordered soluble protein tau and thereby contributing to the neurofibrillary pathology in AD and other tauopathies is abnormal truncation (Kovacech and Novak, 2010). In fact, the progression of AD is at least partially linked

to the existence of several site-specific tau cleavages generating a multitude of various truncated forms. It has been shown that tau truncation alone is sufficient to induce the complete cascade of neurofibrillary pathology suggesting that proteolytical abnormalities in the stressed neurons and production of aberrant tau cleavage products are important events in the AD pathogenesis (Kovacech and Novak, 2010).

High levels of the truncated form Gli3$^R$ of a glioma-associated oncogene 3 (Gli3) are related to various human limb malformations (Ruiz i Altaba, 1999; Hui and Angers, 2011). Glioma-associated oncogene family members 1, 2, and 3 (Gli1, Gli2, and Gli3) are specific transcription factors for controlling the signal transduction of the Hedgehog (Hh) pathway, one of the key regulatory networks involved in animal development (Jiang and Hui, 2008). A peculiar feature of the Gli family members is that the regulated proteolytic processing can convert some Gli proteins from a full-length transcriptional activator form into a truncated repressor form (Hui and Angers, 2011). The appearance of the Gl3$^R$ form represents a result of limited proteolysis of Gli3 by the proteasome (Hui and Angers, 2011). Being transcriptional activators possessing multiple zinc finger motifs, all three members of the Gli family are predicted to be extensively disordered.

Proteolytic processing of cell-surface proteoglycans (PGs) known as ectodomain shedding of syndecans is related to the facilitation of cancer and promotion of the cancer cell motility and invasion thereby increasing aggressiveness of various tumors (Theocharis et al., 2010). Among the PGs with aberrant proteolytic degradation patterns in cancers are versican, aggrecan, brevican, decorin, perlecan, glypicans, and syndecans (Theocharis et al., 2010). The majority of these proteins are characterized by moderate to high levels of predicted disorder. In the crystal structure of human glypican-1 (PDB ID: 4AD7), residues 24–28, 350–361, 406–412, and 476–529 are located within the regions with missing electron density. Similarly, abnormal proteolytic degradation of collagens [which are known to be enriched in intrinsic disorder (Peysselon et al., 2011)] contributes to the cancer pathogenesis via multiple routs via affecting the efficiency of cancer cell invasion, activation of integrins, cancer cell proliferation, angiogenesis (Egeblad et al., 2010).

### Defective trafficking of IDPs and hybrid proteins

It has been pointed out that defects in trafficking of some proteins might represent an important mechanism in many diseases (Delisle et al., 2004). Many of these pathogenic proteins with defective trafficking are IDPs or hybrid proteins with long IDPRs. An illustrative example is given by the low-density lipoprotein receptor (LDLR), mutations in which are associated with familial hypercholesterolemia (Delisle et al., 2004). It was pointed out that ≈50% of LDLR mutations act to disrupt receptor protein trafficking from the ER to the Golgi complex (Hobbs et al., 1990; Goldstein and Brown, 2001). Structural and bioinformatics analysis revealed that LDLR possesses long IDPRs. For example, residues 4–275, 457–464, and 714–788 (PDB ID: 3M0C) comprise regions with missing electron density, which typically correspond to IDPRs (Le Gall et al., 2007).

In cystis fibrosis, multiple mutations in cystic fibrosis transmembrane conductance regulator (CFTR) cause defective protein

trafficking, where the mutated CFTR is retained in the ER (Cheng et al., 1990). The most well-studied mutation affecting normal trafficking of CFTR is the deletion of the codon for a phenylalanine at position 508 (ΔF508) of CFTR protein and is found in 70% of affected patients (Cheng et al., 1990; Delisle et al., 2004). Region 708–831 of CFTR, which is the functional regulatory (R) domain phosphorylation of which domain initiates Cl⁻ channel activity of CFTR, was shown to be predominantly disordered in solution (Ostedgaard et al., 2000).

Mutations in connexin 32 (Cx32), which is an integral gap junction membrane protein forming channels for the transmission of electrical signals and diffusion of small ions and molecules between coupled cells, are known to be associated with the Charcot-Marie-Tooth disease (CMTX) (Bergoffen et al., 1993; Delisle et al., 2004). Some CMTX mutations found in this complex polygenic neuropathic disorder, that constitutes the most common form of inheritable disease in the peripheral nervous system, result in the Cx32 protein that is synthesized but not properly transported to the plasma membrane (Delisle et al., 2004). C-terminal tail of Cx32 protein (residues 221–283) is predicted to be extensively disordered.

In the norm, the vasopressin-regulated aquaporin-2 channel (AQP2) is found in the kidney at distinct sites along nephrons and collecting ducts (Nielsen et al., 2002). Several mutations in AQP2 are associated with impaired transport from the ER and development of the non-X-linked nephrogenic diabetes insipidus (NDI), a hereditary malady characterized by the inability to concentrate the urine resulting in excessive urine production and thirst (Tamarappoo and Verkman, 1998) Being a transmembrane protein, AQP2 is expected to be mostly structured. However, computational analysis revealed that it has two long predicted IDPRs, residues 147–164 and 238–271.

## CONCLUDING REMARKS

Intrinsic disorder is very common in proteins, plays various roles in numerous protein functions and is tightly controlled in the healthy cells. Among the major functions of IDPs are recognition as well as regulation and control of various signaling events. Normally, function and abundance of IDPs are tightly controlled. The breaking these proteins out of control brings havoc and initiate various pathological events. In fact, IDPs and hybrid proteins containing ordered domains and functional IDPRs are commonly involved in the pathogenesis of human diseases. The crucial roles of IDPs and hybrid proteins in promoting and supporting the disease states are obvious from several illustrative examples of well-characterized disease-related IDPs, such as p53, α-synuclein, PTEN, tau protein, prion protein, etc., as well as from the results of comprehensive bioinformatics studies. The high degree of involvement of IDPs in the pathogenesis of various diseases is determined by the unique structural and functional properties of these proteins. IDPs and IDPRs frequently serve as major cellular regulators, recognizers and signal transducers. Their normal functionality is tightly controlled and modulated via a wide spectrum of PTMs and AS. Many IDPs/IDPRs can fold (completely or partially) upon interaction with corresponding binding partners and possess multiple binding specificities, enabling them to participate in one-to-many and many-to-one interactions.

Distortion of any of the mechanisms controlling IDP/IDPR functionality can be detrimental. Some disease-related proteins have an intrinsic propensity to form pathologic conformations, whereas other proteins require some external factors, such as impaired interactions with chaperones, intracellular or extracellular matrices, other proteins, small molecules, and additional endogenous factors, to gain conformational alterations leading to increased propensities for misfolding and dysfunction. Often, protein pathogenicity originates from altered splicing, chromosomal translocations, abnormal expression levels, or impaired trafficking. Formation of pathologic conformations can also be triggered by aberrant PTMs and increased degradation propensities. In other words, any cellular event that affects the functionality, foldability, abundance, or cellular distribution of these important players may cause pathological transformations.

## REFERENCES

Adamia, S., Haibe-Kains, B., Pilarski, P. M., Bar-Natan, M., Pevzner, S., Avet-Loiseau, H., et al. (2014). A Genome-wide aberrant RNA splicing in patients with acute myeloid leukemia identifies novel potential disease markers and therapeutic targets. *Clin. Cancer Res.* 20, 1135–1145. doi: 10.1158/1078-0432.CCR-13-0956

Adler, A. S., McCleland, M. L., Yee, S., Yaylaoglu, M., Hussain, S., Cosino, E., et al. (2014). An integrative analysis of colon cancer identifies an essential function for PRPF6 in tumor growth. *Genes Dev.* 28, 1068–1084. doi: 10.1101/gad.237206.113

Ahn, J., Byeon, I. J., Byeon, C. H., and Gronenborn, A. M. (2009). Insight into the structural basis of pro- and antiapoptotic p53 modulation by ASPP proteins. *J. Biol. Chem.* 284, 13812–13822. doi: 10.1074/jbc.M808821200

Alderson, T. R., and Markley, J. L. (2013). Biophysical characterization of alpha-synuclein and its controversial structure. *Intrinsically Disord. Proteins* 1, 18–39. doi: 10.4161/idp.26255

Anderson, C. W., and Appella, E. (2004). "Signaling to the p53 tumor suppressor through pathways activated by genotoxic and nongenotoxic stress," in *Handbook of Cell Signaling*, eds R. A. Bradshaw and E. A. Dennis (New York, NY: Academic Press), 237–247.

Andreadis, A. (2012). Tau splicing and the intricacies of dementia. *J. Cell. Physiol.* 227, 1220–1225. doi: 10.1002/jcp.22842

Asher, G., Tsvetkov, P., Kahana, C., and Shaul, Y. (2005). A mechanism of ubiquitin-independent proteasomal degradation of the tumor suppressors p53 and p73. *Genes Dev.* 19, 316–321. doi: 10.1101/gad.319905

Babu, M. M., van der Lee, R., de Groot, N. S., and Gsponer, J. (2011). Intrinsically disordered proteins: regulation and disease. *Curr. Opin. Struct. Biol.* 21, 432–440. doi: 10.1016/j.sbi.2011.03.011

Baens, M., Peeters, P., Guo, C. Y., Aerssens, J., and Marynen, P. (1996). Genomic organization of TEL: the human ETS-variant gene 6. *Genome Res.* 6, 404–413. doi: 10.1101/gr.6.5.404

Baeten, K. M., and Akassoglou, K. (2011). Extracellular matrix and matrix receptors in blood-brain barrier formation and stroke. *Dev. Neurobiol.* 71, 1018–1039. doi: 10.1002/dneu.20954

Baker, J. M., Hudson, R. P., Kanelis, V., Choy, W. Y., Thibodeau, P. H., Thomas, P. J., et al. (2007). CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices. *Nat. Struct. Mol. Biol.* 14, 738–745. doi: 10.1038/nsmb1278

Ballerini, P., Struski, S., Cresson, C., Prade, N., Toujani, S., Deswarte, C., et al. (2012). RET fusion genes are associated with chronic myelomonocytic leukemia and enhance monocytic differentiation. *Leukemia* 26, 2384–2389. doi: 10.1038/leu.2012.109

Bellotti, V., Mangione, P., and Stoppini, M. (1999). Biological activity and pathological implications of misfolded proteins. *Cell. Mol. Life Sci.* 55, 977–991. doi: 10.1007/s000180050348

Bergoffen, J., Scherer, S. S., Wang, S., Scott, M. O., Bone, L. J., Paul, D. L., et al. (1993). Connexin mutations in X-linked Charcot-Marie-Tooth disease. *Science* 262, 2039–2042. doi: 10.1126/science.8266101

Beyer, K. (2006). Alpha-synuclein structure, posttranslational modification and alternative splicing as aggregation enhancers. *Acta Neuropathol.* 112, 237–251. doi: 10.1007/s00401-006-0104-6

Beyer, K., and Ariza, A. (2013). Alpha-Synuclein posttranslational modification and alternative splicing as a trigger for neurodegeneration. *Mol. Neurobiol.* 47, 509–524. doi: 10.1007/s12035-012-8330-5

Biamonti, G., Catillo, M., Pignataro, D., Montecucco, A., and Ghigna, C. (2014). The alternative splicing side of cancer. *Semin. Cell Dev. Biol.* doi: 10.1016/j.semcdb.2014.03.016. [Epub ahead of print].

Bischof, D., Pulford, K., Mason, D. Y., and Morris, S. W. (1997). Role of the nucleophosmin (NPM) portion of the non-Hodgkin's lymphoma-associated NPM-anaplastic lymphoma kinase fusion protein in oncogenesis. *Mol. Cell. Biol.* 17, 2312–2325.

Blumenthal, G. M., and Dennis, P. A. (2008). PTEN hamartoma tumor syndromes. *Eur. J. Hum. Genet.* 16, 1289–1300. doi: 10.1038/ejhg.2008.162

Bonomi, S., Gallo, S., Catillo, M., Pignataro, D., Biamonti, G., and Ghigna, C. (2013). Oncogenic alternative splicing switches: role in cancer progression and prospects for therapy. *Int. J. Cell Biol.* 2013, 962038. doi: 10.1155/2013/962038

Bossi, D., Carlomagno, F., Pallavicini, I., Pruneri, G., Trubia, M., Raviele, P. R., et al. (2014). Functional characterization of a novel FGFR1OP-RET rearrangement in hematopoietic malignancies. *Mol. Oncol.* 8, 221–231. doi: 10.1016/j.molonc.2013.11.004

Bourdon, J. C. (2007). p53 Family isoforms. *Curr. Pharm. Biotechnol.* 8, 332–336. doi: 10.2174/138920107783018444

Bourdon, J. C., Fernandes, K., Murray-Zmijewski, F., Liu, G., Diot, A., Xirodimas, D. P., et al. (2005). p53 isoforms can regulate p53 transcriptional activity. *Genes Dev.* 19, 2122–2137. doi: 10.1101/gad.1339905

Breydo, L., Wu, J. W., and Uversky, V. N. (2012). Alpha-synuclein misfolding and Parkinson's disease. *Biochim. Biophys. Acta* 1822, 261–285. doi: 10.1016/j.bbadis.2011.10.002

Buljan, M., Chalancon, G., Dunker, A. K., Bateman, A., Balaji, S., Fuxreiter, M., et al. (2013). Alternative splicing of intrinsically disordered regions and rewiring of protein interactions. *Curr. Opin. Struct. Biol.* 23, 443–450. doi: 10.1016/j.sbi.2013.03.006

Buljan, M., Chalancon, G., Eustermann, S., Wagner, G. P., Fuxreiter, M., Bateman, A., et al. (2012). Tissue-specific splicing of disordered segments that embed binding motifs rewires protein interaction networks. *Mol. Cell* 46, 871–883. doi: 10.1016/j.molcel.2012.05.039

Buratti, E., and Baralle, F. E. (2008). Multiple roles of TDP-43 in gene expression, splicing regulation, and human disease. *Front. Biosci.* 13:867–878. doi: 10.2741/2727

Burrell, L. M., Harrap, S. B., Velkoska, E., and Patel, S. K. (2013). The ACE2 gene: its potential as a functional candidate for cardiovascular disease. *Clin. Sci. (Lond.)* 124, 65–76. doi: 10.1042/CS20120269

Bustos, D. M., and Iglesias, A. A. (2006). Intrinsic disorder is a key characteristic in partners that bind 14-3-3 proteins. *Proteins* 63, 35–42. doi: 10.1002/prot.20888

Campbell, M. J., and Turner, B. M. (2013). Altered histone modifications in cancer. *Adv. Exp. Med. Biol.* 754, 81–107. doi: 10.1007/978-1-4419-9967-2_4

Cancellotti, E., Mahal, S. P., Somerville, R., Diack, A., Brown, D., Piccardo, P., et al. (2013). Post-translational changes to PrP alter transmissible spongiform encephalopathy strain properties. *EMBO J.* 32, 756–769. doi: 10.1038/emboj.2013.6

Cardenas, A. M., Ardiles, A. O., Barraza, N., Baez-Matus, X., and Caviedes, P. (2012). Role of tau protein in neuronal damage in Alzheimer's disease and Down syndrome. *Arch. Med. Res.* 43, 645–654. doi: 10.1016/j.arcmed.2012.10.012

Carlino, D., De Vanna, M., and Tongiorgi, E. (2013). Is altered BDNF biosynthesis a general feature in patients with cognitive dysfunctions? *Neuroscientist* 19, 345–353. doi: 10.1177/1073858412469444

Carnicer, R., Crabtree, M. J., Sivakumaran, V., Casadei, B., and Kass, D. A. (2013). Nitric oxide synthases in heart failure. *Antioxid. Redox Signal.* 18, 1078–1099. doi: 10.1089/ars.2012.4824

Carracedo, A., Alimonti, A., and Pandolfi, P. P. (2011). PTEN level in tumor suppression: how much is too little? *Cancer Res.* 71, 629–633. doi: 10.1158/0008-5472.CAN-10-2488

Casu, F., Duggan, B. M., and Hennig, M. (2013). The arginine-rich RNA-binding motif of HIV-1 Rev is intrinsically disordered and folds upon RRE binding. *Biophys. J.* 105, 1004–1017. doi: 10.1016/j.bpj.2013.07.022

Chakrabortee, S., Tripathi, R., Watson, M., Schierle, G. S., Kurniawan, D. P., Kaminski, C. F., et al. (2012). Intrinsically disordered proteins as molecular shields. *Mol. Biosyst.* 8, 210–219. doi: 10.1039/c1mb05263b

Chandra, V., Huang, P., Hamuro, Y., Raghuram, S., Wang, Y., Burris, T. P., et al. (2008). Structure of the intact PPAR-gamma-RXR- nuclear receptor complex on DNA. *Nature* 456, 350–356. doi: 10.1038/nature07413

Chatterjee, S., Senapati, P., and Kundu, T. K. (2012). Post-translational modifications of lysine in DNA-damage repair. *Essays Biochem.* 52, 93–111. doi: 10.1042/bse0520093

Chen, G. L., and Miller, G. M. (2013). Extensive alternative splicing of the repressor element silencing transcription factor linked to cancer. *PLoS ONE* 8:e62217. doi: 10.1371/journal.pone.0062217

Chen, H., Xu, C., Jin, Q., and Liu, Z. (2014). S100 protein family in human cancer. *Am. J. Cancer Res.* 4, 89–115.

Chen, H. M., Lin, Y. H., Cheng, Y. M., Wing, L. Y., and Tsai, S. J. (2013). Overexpression of integrin-beta1 in leiomyoma promotes cell spreading and proliferation. *J. Clin. Endocrinol. Metab.* 98, E837–E846. doi: 10.1210/jc.2012-3647

Chen, J., and Weiss, W. A. (2014). Alternative splicing in cancer: implications for biology and therapy. *Oncogene.* doi: 10.1038/onc.2013.570. [Epub ahead of print].

Chen, R., Tan, Y., Wang, M., Wang, F., Yao, Z., Dong, L., et al. (2011). Development of glycoprotein capture-based label-free method for the high-throughput screening of differential glycoproteins in hepatocellular carcinoma. *Mol. Cell. Proteomics* 10, M110.006445. doi: 10.1074/mcp.M110.006445

Cheng, S. H., Gregory, R. J., Marshall, J., Paul, S., Souza, D. W., White, G. A., et al. (1990). Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. *Cell* 63, 827–834. doi: 10.1016/0092-8674(90)90148-8

Cheng, Y., LeGall, T., Oldfield, C. J., Dunker, A. K., and Uversky, V. N. (2006). Abundance of intrinsic disorder in protein associated with cardiovascular disease. *Biochemistry* 45, 10448–10460. doi: 10.1021/bi060981d

Chipuk, J. E., Fisher, J. C., Dillon, C. P., Kriwacki, R. W., Kuwana, T., and Green, D. R. (2008). Mechanism of apoptosis induction by inhibition of the anti-apoptotic BCL-2 proteins. *Proc. Natl. Acad. Sci. U.S.A.* 105, 20327–20332. doi: 10.1073/pnas.0808036105

Choi, J. H., Banks, A. S., Estall, J. L., Kajimura, S., Bostrom, P., Laznik, D., et al. (2010). Anti-diabetic drugs inhibit obesity-linked phosphorylation of PPARgamma by Cdk5. *Nature* 466, 451–456. doi: 10.1038/nature09291

Coelho Ribeiro Mde, L., Espinosa, J., Islam, S., Martinez, O., Thanki, J. J., Mazariegos, S., et al. (2013). Malleable ribonucleoprotein machine: protein intrinsic disorder in the Saccharomyces cerevisiae spliceosome. *PeerJ* 1, e2. doi: 10.7717/peerj.2

Colak, R., Kim, T., Michaut, M., Sun, M., Irimia, M., Bellay, J., et al. (2013). Distinct types of disorder in the human proteome: functional implications for alternative splicing. *PLoS Comput. Biol.* 9:e1003030. doi: 10.1371/journal.pcbi.1003030

Collins, M. O., Yu, L., Campuzano, I., Grant, S. G., and Choudhary, J. S. (2008). Phosphoproteomic analysis of the mouse brain cytosol reveals a predominance of protein phosphorylation in regions of intrinsic sequence disorder. *Mol. Cell. Proteomics* 7, 1331–1348. doi: 10.1074/mcp.M700564-MCP200

Cozzetto, D., and Jones, D. T. (2013). The contribution of intrinsic disorder prediction to the elucidation of protein function. *Curr. Opin. Struct. Biol.* 23, 467–472. doi: 10.1016/j.sbi.2013.02.001

Cryns, K., Sivakumaran, T. A., Van den Ouweland, J. M., Pennings, R. J., Cremers, C. W., Flothmann, K., et al. (2003). Mutational spectrum of the WFS1 gene in Wolfram syndrome, nonsyndromic hearing impairment, diabetes mellitus, and psychiatric disease. *Hum. Mutat.* 22, 275–287. doi: 10.1002/humu.10258

Cuanalo-Contreras, K., Mukherjee, A., and Soto, C. (2013). Role of protein misfolding and proteostasis deficiency in protein misfolding diseases and aging. *Int. J. Cell Biol.* 2013:638083. doi: 10.1155/2013/638083

David, C. J., and Manley, J. L. (2010). Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes Dev.* 24, 2343–2364. doi: 10.1101/gad.1973010

De Braekeleer, E., Douet-Guilbert, N., Morel, F., Le Bris, M. J., Basinko, A., and De Braekeleer, M. (2012). ETV6 fusion genes in hematological malignancies: a review. *Leukemia Res.* 36, 945–961. doi: 10.1016/j.leukres.2012.04.010

De Jonge, N., Garcia-Pino, A., Buts, L., Haesaerts, S., Charlier, D., Zangger, K., et al. (2009). Rejuvenation of CcdB-poisoned gyrase by an intrinsically disordered protein domain. *Mol. Cell* 35, 154–163. doi: 10.1016/j.molcel.2009.05.025

de la Grange, P., Gratadou, L., Delord, M., Dutertre, M., and Auboeuf, D. (2010). Splicing factor and exon profiling across human tissues. *Nucleic Acids Res.* 38, 2825–2838. doi: 10.1093/nar/gkq008

de Laureto, P. P., Tosatto, L., Frare, E., Marin, O., Uversky, V. N., and Fontana, A. (2006). Conformational properties of the SDS-bound state of alpha-synuclein probed by limited proteolysis: unexpected rigidity of the acidic C-terminal tail. *Biochemistry* 45, 11523–11531. doi: 10.1021/bi052614s

Delisle, B. P., Anson, B. D., Rajamani, S., and January, C. T. (2004). Biology of cardiac arrhythmias: ion channel protein trafficking. *Circ. Res.* 94, 1418–1428. doi: 10.1161/01.RES.0000128561.28701.ea

Demarest, S. J., Martinez-Yamout, M., Chung, J., Chen, H., Xu, W., Dyson, H. J., et al. (2002). Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* 415, 549–553. doi: 10.1038/415549a

Didenko, T., Duarte, A. M., Karagoz, G. E., and Rudiger, S. G. (2012). Hsp90 structure and function studied by NMR spectroscopy. *Biochim. Biophys. Acta* 1823, 636–647. doi: 10.1016/j.bbamcr.2011.11.009

Di Domenico, T., Walsh, I., Martin, A. J., and Tosatto, S. C. (2012). MobiDB: a comprehensive database of intrinsic protein disorder annotations. *Bioinformatics* 28, 2080–2081. doi: 10.1093/bioinformatics/bts327

Di Domenico, T., Walsh, I., and Tosatto, S. C. (2013). Analysis and consensus of currently available intrinsic protein disorder annotation sources in the MobiDB database. *BMC Bioinformatics* 14(Suppl. 7):S3. doi: 10.1186/1471-2105-14-S7-S3

Di Gennaro, E., Bruzzese, F., Caraglia, M., Abruzzese, A., and Budillon, A. (2004). Acetylation of proteins as novel target for antitumor therapy: review article. *Amino Acids* 26, 435–441. doi: 10.1007/s00726-004-0087-3

Dobson, C. M. (1999). Protein misfolding, evolution and disease. *Trends Biochem. Sci.* 24, 329–332. doi: 10.1016/S0968-0004(99)01445-0

Dong, J. T., and Chen, C. (2009). Essential role of KLF5 transcription factor in cell proliferation and differentiation and its implications for human diseases. *Cell. Mol. Life Sci.* 66, 2691–2706. doi: 10.1007/s00018-009-0045-z

Dosztanyi, Z., Chen, J., Dunker, A. K., Simon, I., and Tompa, P. (2006). Disorder and sequence repeats in hub proteins and their implications for network evolution. *J. Proteome Res.* 5, 2985–2995. doi: 10.1021/pr060171o

Drake, R. R., Schwegler, E. E., Malik, G., Diaz, J., Block, T., Mehta, A., et al. (2006). Lectin capture strategies combined with mass spectrometry for the discovery of serum glycoprotein biomarkers. *Mol. Cell. Proteomics* 5, 1957–1967. doi: 10.1074/mcp.M600176-MCP200

Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradovic, Z. (2002). Intrinsic disorder and protein function. *Biochemistry* 41, 6573–6582. doi: 10.1021/bi012159+

Dunker, A. K., Cortese, M. S., Romero, P., Iakoucheva, L. M., and Uversky, V. N. (2005). Flexible nets: the roles of intrinsic disorder in protein interaction networks. *FEBS J.* 272, 5129–5148. doi: 10.1111/j.1742-4658.2005.04948.x

Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., et al. (2001). Intrinsically disordered protein. *J. Mol. Graph. Model.* 19, 26–59. doi: 10.1016/S1093-3263(00)00138-8

Dunker, A. K., Obradovic, Z., Romero, P., Garner, E. C., and Brown, C. J. (2000). Intrinsic protein disorder in complete genomes. *Genome Inform. Ser. Workshop Genome Inform.* 11, 161–171.

Duygu, B., Poels, E. M., and da Costa Martins, P. A. (2013). Genetics and epigenetics of arrhythmia and heart failure. *Front. Genet.* 4:219. doi: 10.3389/fgene.2013.00219

Dyson, H. J. (2011). Expanding the proteome: disordered and alternatively folded proteins. *Q. Rev. Biophys.* 44, 467–518. doi: 10.1017/S0033583511000060

Dyson, H. J., and Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* 6, 197–208. doi: 10.1038/nrm1589

Ebert, M. O., Bae, S. H., Dyson, H. J., and Wright, P. E. (2008). NMR relaxation study of the complex formed between CBP and the activation domain of the nuclear hormone receptor coactivator ACTR. *Biochemistry* 47, 1299–1308. doi: 10.1021/bi701767j

Egeblad, M., Rasch, M. G., and Weaver, V. M. (2010). Dynamic interplay between the collagen scaffold and tumor evolution. *Curr. Opin. Cell Biol.* 22, 697–706. doi: 10.1016/j.ceb.2010.08.015

Ehrnhoefer, D. E., Sutton, L., and Hayden, M. R. (2011). Small changes, big impact: posttranslational modifications and function of huntingtin in Huntington disease. *Neuroscientist* 17, 475–492. doi: 10.1177/1073858410390378

Ekman, D., Light, S., Bjorklund, A. K., and Elofsson, A. (2006). What properties characterize the hub proteins of the protein-protein interaction network of Saccharomyces cerevisiae? *Genome Biol.* 7, R45. doi: 10.1186/gb-2006-7-6-r45

Elia, A. E., Petrucci, S., Fasano, A., Guidi, M., Valbonesi, S., Bernardini, L., et al. (2013). Alpha-synuclein gene duplication: marked intrafamilial variability in two novel pedigrees. *Mov. Disord.* 28, 813–817. doi: 10.1002/mds.25518

Eriksen, J. L., Przedborski, S., and Petrucelli, L. (2005). Gene dosage and pathogenesis of Parkinson's disease. *Trends Mol. Med.* 11, 91–96. doi: 10.1016/j.molmed.2005.01.001

Escher, P., and Wahli, W. (2000). Peroxisome proliferator-activated receptors: insight into multiple cellular functions. *Mutat. Res.* 448, 121–138. doi: 10.1016/S0027-5107(99)00231-6

Fasshauer, D., Bruns, D., Shen, B., Jahn, R., and Brunger, A. T. (1997). A structural change occurs upon binding of syntaxin to SNAP-25. *J. Biol. Chem.* 272, 4582–4590. doi: 10.1074/jbc.272.7.4582

Felkin, L. E., Narita, T., Germack, R., Shintani, Y., Takahashi, K., Sarathchandra, P., et al. (2011). Calcineurin splicing variant calcineurin Abeta1 improves cardiac function after myocardial infarction without inducing hypertrophy. *Circulation* 123, 2838–2847. doi: 10.1161/CIRCULATIONAHA.110.012211

Feng, H. Z., and Jin, J. P. (2010). Coexistence of cardiac troponin T variants reduces heart efficiency. *Am. J. Physiol. Heart Circ. Physiol.* 299, H97–H105. doi: 10.1152/ajpheart.01105.2009

Ferreon, A. C., Ferreon, J. C., Wright, P. E., and Deniz, A. A. (2013). Modulation of allostery by protein intrinsic disorder. *Nature* 498, 390–394. doi: 10.1038/nature12294

Ferreon, J. C., Lee, C. W., Arai, M., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2009a). Cooperative regulation of p53 by modulation of ternary complex formation with CBP/p300 and HDM2. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6591–6596. doi: 10.1073/pnas.0811023106

Ferreon, J. C., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2009b). Structural basis for subversion of cellular control mechanisms by the adenoviral E1A oncoprotein. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13260–13265. doi: 10.1073/pnas.0906770106

Fischer, E. (1894). Einfluss der configuration auf die wirkung der enzyme. *Ber. Dt. Chem. Ges.* 27, 2985–2993. doi: 10.1002/cber.18940270364

Follis, A. V., Chipuk, J. E., Fisher, J. C., Yun, M. K., Grace, C. R., Nourse, A., et al. (2013). PUMA binding induces partial unfolding within BCL-xL to disrupt p53 binding and promote apoptosis. *Nat. Chem. Biol.* 9, 163–168. doi: 10.1038/nchembio.1166

Follis, A. V., Galea, C. A., and Kriwacki, R. W. (2012). Intrinsic protein flexibility in regulation of cell proliferation: advantages for signaling and opportunities for novel therapeutics. *Adv. Exp. Med. Biol.* 725, 27–49. doi: 10.1007/978-1-4614-0659-4_3

Follis, A. V., Llambi, F., Ou, L., Baran, K., Green, D. R., and Kriwacki, R. W. (2014). The DNA-binding domain mediates both nuclear and cytosolic functions of p53. *Nat. Struct. Mol. Biol.* 21, 535–543. doi: 10.1038/nsmb.2829

Fontana, A., de Laureto, P. P., Spolaore, B., Frare, E., Picotti, P., and Zambonin, M. (2004). Probing protein structure by limited proteolysis. *Acta Biochim. Pol.* 51, 299–321.

Fontana, A., Fassina, G., Vita, C., Dalzoppo, D., Zamai, M., and Zambonin, M. (1986). Correlation between sites of limited proteolysis and segmental mobility in thermolysin. *Biochemistry* 25, 1847–1851. doi: 10.1021/bi00356a001

Fontana, A., Polverino de Laureto, P., De Filippis, V., Scaramella, E., and Zambonin, M. (1997). Probing the partly folded states of proteins by limited proteolysis. *Fold. Des.* 2, R17–R26. doi: 10.1016/S1359-0278(97)00010-2

Frye, J. J., Brown, N. G., Petzold, G., Watson, E. R., Grace, C. R., Nourse, A., et al. (2013). Electron microscopy structure of human APC/C(CDH1)-EMI1 reveals multimodal mechanism of E3 ligase shutdown. *Nat. Struct. Mol. Biol.* 20, 827–835. doi: 10.1038/nsmb.2593

Fu, R. H., Liu, S. P., Huang, S. J., Chen, H. J., Chen, P. R., Lin, Y. H., et al. (2013). Aberrant alternative splicing events in Parkinson's disease. *Cell Transplant.* 22, 653–661. doi: 10.3727/096368912X655154

Fukuchi, S., Amemiya, T., Sakamoto, S., Nobe, Y., Hosoda, K., Kado, Y., et al. (2014). IDEAL in 2014 illustrates interaction networks composed of intrinsically disordered proteins and their binding partners. *Nucleic Acids Res.* 42, D320–D325. doi: 10.1093/nar/gkt1010

Fukuchi, S., Sakamoto, S., Nobe, Y., Murakami, S. D., Amemiya, T., Hosoda, K., et al. (2012). IDEAL: intrinsically disordered proteins with extensive annotations and literature. *Nucleic Acids Res.* 40, D507–D511. doi: 10.1093/nar/gkr884

Fuxreiter, M., Simon, I., and Bondos, S. (2011). Dynamic protein-DNA recognition: beyond what can be seen. *Trends Biochem. Sci.* 36, 415–423. doi: 10.1016/j.tibs.2011.04.006

Fuxreiter, M., Tompa, P., Simon, I., Uversky, V. N., Hansen, J. C., and Asturias, F. J. (2008). Malleable machines take shape in eukaryotic transcriptional regulation. *Nat. Chem. Biol.* 4, 728–737. doi: 10.1038/nchembio.127

Gains, M. J., and LeBlanc, A. C. (2007). Canadian association of neurosciences review: prion protein and prion diseases: the good and the bad. *Can. J. Neurol. Sci.* 34, 126–145.

Galea, C. A., Wang, Y., Sivakolundu, S. G., and Kriwacki, R. W. (2008a). Regulation of cell division by intrinsically unstructured proteins: intrinsic flexibility, modularity, and signaling conduits. *Biochemistry* 47, 7598–7609. doi: 10.1021/bi8006803

Galea, C. A., Nourse, A., Wang, Y., Sivakolundu, S. G., Heller, W. T., and Kriwacki, R. W. (2008b). Role of intrinsic flexibility in signal transduction mediated by the cell cycle regulator, p27 Kip1. *J. Mol. Biol.* 376, 827–838. doi: 10.1016/j.jmb.2007.12.016

Gargalionis, A. N., Piperi, C., Adamopoulos, C., and Papavassiliou, A. G. (2012). Histone modifications as a pathogenic mechanism of colorectal tumorigenesis. *Int. J. Biochem. Cell Biol.* 44, 1276–1289. doi: 10.1016/j.biocel.2012.05.002

Gasperini, R. J., Klaver, D. W., Hou, X., Aguilar, M. I., and Small, D. H. (2012). Mechanisms of transthyretin aggregation and toxicity. *Subcell. Biochem.* 65, 211–224. doi: 10.1007/978-94-007-5416-4_9

Gezer, U., and Holdenrieder, S. (2014). Post-translational histone modifications in circulating nucleosomes as new biomarkers in colorectal cancer. *In Vivo* 28, 287–292.

Gioeli, D., and Paschal, B. M. (2012). Post-translational modification of the androgen receptor. *Mol. Cell. Endocrinol.* 352, 70–78. doi: 10.1016/j.mce.2011.07.004

Goh, K. I., Cusick, M. E., Valle, D., Childs, B., Vidal, M., and Barabasi, A. L. (2007). The human disease network. *Proc. Natl. Acad. Sci. U.S.A.* 104, 8685–8690. doi: 10.1073/pnas.0701361104

Goldstein, J. L., and Brown, M. S. (2001). Molecular medicine. The cholesterol quartet. *Science* 292, 1310–1312. doi: 10.1126/science.1061815

Gorbatenko, A., Olesen, C. W., Boedtkjer, E., and Pedersen, S. F. (2014). Regulation and roles of bicarbonate transporters in cancer. *Front. Physiol.* 5:130. doi: 10.3389/fphys.2014.00130

Greenfield, S. A., Zimmermann, M., and Bond, C. E. (2008). Non-hydrolytic functions of acetylcholinesterase. The significance of C-terminal peptides. *FEBS J.* 275, 604–611. doi: 10.1111/j.1742-4658.2007.06235.x

Grimmler, M., Wang, Y., Mund, T., Cilensek, Z., Keidel, E. M., Waddell, M. B., et al. (2007). Cdk-inhibitory activity and stability of p27Kip1 are directly regulated by oncogenic tyrosine kinases. *Cell* 128, 269–280. doi: 10.1016/j.cell.2006.11.047

Gsponer, J., Futschik, M. E., Teichmann, S. A., and Babu, M. M. (2008). Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. *Science* 322, 1365–1368. doi: 10.1126/science.1163581

Guo, X., Bulyk, M. L., and Hartemink, A. J. (2012). Intrinsic disorder within and flanking the DNA-binding domains of human transcription factors. *Pac. Symp. Biocomput.* 104–115.

Halliday, A. C., and Greenfield, S. A. (2012). From protein to peptides: a spectrum of non-hydrolytic functions of acetylcholinesterase. *Protein Pept. Lett.* 19, 165–172. doi: 10.2174/092986612799080149

Hammer, M. R., John, P. N., Flynn, M. D., Bellingham, A. J., and Leslie, R. D. (1989). Glycated fibrinogen: a new index of short-term diabetic control. *Ann. Clin. Biochem.* 26(pt 1), 58–62. doi: 10.1177/000456328902600108

Hauselmann, I., and Borsig, L. (2014). Altered tumor-cell glycosylation promotes metastasis. *Front. Oncol.* 4:28. doi: 10.3389/fonc.2014.00028

Haynes, C., Oldfield, C. J., Ji, F., Klitgord, N., Cusick, M. E., Radivojac, P., et al. (2006). Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS Comput. Biol.* 2:e100. doi: 10.1371/journal.pcbi.0020100

Hegyi, H., Buday, L., and Tompa, P. (2009). Intrinsic structural disorder confers cellular viability on oncogenic fusion proteins. *PLoS Comput. Biol.* 5:e1000552. doi: 10.1371/journal.pcbi.1000552

Hegyi, H., and Tompa, P. (2012). Increased structural disorder of proteins encoded on human sex chromosomes. *Mol. Biosyst.* 8, 229–236. doi: 10.1039/c1mb05285c

Henschen-Edman, A. H. (2001). Fibrinogen non-inherited heterogeneity and its relationship to function in health and disease. *Ann. N.Y Acad. Sci.* 936, 580–593. doi: 10.1111/j.1749-6632.2001.tb03546.x

Hernandez, F., and Avila, J. (2007). Tauopathies. *Cell. Mol. Life Sci.* 64, 2219–2233. doi: 10.1007/s00018-007-7220-x

Hernandez, L., Pinyol, M., Hernandez, S., Bea, S., Pulford, K., Rosenwald, A., et al. (1999). TRK-fused gene (TFG) is a new partner of ALK in anaplastic large cell lymphoma producing two structurally different TFG-ALK translocations. *Blood* 94, 3265–3268.

Herren, A. W., Bers, D. M., and Grandi, E. (2013). Post-translational modifications of the cardiac Na channel: contribution of CaMKII-dependent phosphorylation to acquired arrhythmias. *Am. J. Physiol. Heart Circ. Physiol.* 305, H431–H445. doi: 10.1152/ajpheart.00306.2013

Hipp, M. S., Park, S. H., and Hartl, F. U. (2014). Proteostasis impairment in protein-misfolding and -aggregation diseases. *Trends Cell Biol.* doi: 10.1016/j.tcb.2014.05.003

Hobbs, H. H., Russell, D. W., Brown, M. S., and Goldstein, J. L. (1990). The LDL receptor locus in familial hypercholesterolemia: mutational analysis of a membrane protein. *Annu. Rev. Genet.* 24, 133–170. doi: 10.1146/annurev.ge.24.120190.001025

Hoffman, M. (2008). Alterations of fibrinogen structure in human disease. *Cardiovasc. Hematol. Agents Med. Chem.* 6, 206–211. doi: 10.2174/187152508784871981

Hoffman, R. M., Blumenschein, T. M., and Sykes, B. D. (2006). An interplay between protein disorder and structure confers the $Ca^{2+}$ regulation of striated muscle. *J. Mol. Biol.* 361, 625–633. doi: 10.1016/j.jmb.2006.06.031

Hoffman, R. M., and Sykes, B. D. (2008). Isoform-specific variation in the intrinsic disorder of troponin I. *Proteins* 73, 338–350. doi: 10.1002/prot.22063

Hollander, M. C., Blumenthal, G. M., and Dennis, P. A. (2011). PTEN loss in the continuum of common cancers, rare syndromes and mouse models. *Nat. Rev. Cancer* 11, 289–301. doi: 10.1038/nrc3037

Hopkins, B. D., Fine, B., Steinbach, N., Dendy, M., Rapp, Z., Shaw, J., et al. (2013). A secreted PTEN Phosphatase That Enters Cells to alter signaling and survival. *Science* 341, 399–402. doi: 10.1126/science.1234907

Hu, Y., Dobi, A., Sreenath, T., Cook, C., Tadase, A. Y., Ravindranath, L., et al. (2008). Delineation of TMPRSS2-ERG splice variants in prostate cancer. *Clin. Cancer Res.* 14, 4719–4725. doi: 10.1158/1078-0432.CCR-08-0531

Huang, Z., and Bao, S. (2012). Ubiquitination and deubiquitination of REST and its roles in cancers. *FEBS Lett.* 586, 1602–1605. doi: 10.1016/j.febslet.2012.04.052

Hui, C. C., and Angers, S. (2011). Gli proteins in development and disease. *Annu. Rev. Cell Dev. Biol.* 27, 513–537. doi: 10.1146/annurev-cellbio-092910-154048

Iakoucheva, L. M., Brown, C. J., Lawson, J. D., Obradovic, Z., and Dunker, A. K. (2002). Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* 323, 573–584. doi: 10.1016/S0022-2836(02)00969-5

Iakoucheva, L. M., Kimzey, A. L., Masselon, C. D., Bruce, J. E., Garner, E. C., Brown, C. J., et al. (2001). Identification of intrinsic order and disorder in the DNA repair protein XPA. *Protein Sci.* 10, 560–571. doi: 10.1110/ps.29401

Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z., et al. (2004). The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* 32, 1037–1049. doi: 10.1093/nar/gkh253

Imai, Y., Philippe, N., Sesti, G., Accili, D., and Taylor, S. I. (1997). Expression of variant forms of insulin receptor substrate-1 identified in patients with noninsulin-dependent diabetes mellitus. *J. Clin. Endocrinol. Metab.* 82, 4201–4207.

Ishihara, T., Fukuda, I., Morita, A., Takinami, Y., Okamoto, H., Nishimura, S., et al. (2011). Development of quantitative plasma N-glycoproteomics using label-free 2-D LC-MALDI MS and its applicability for biomarker discovery in hepatocellular carcinoma. *J. Proteomics* 74, 2159–2168. doi: 10.1016/j.jprot.2011.06.010

Iwai, A., Masliah, E., Yoshimoto, M., Ge, N., Flanagan, L., de Silva, H. A., et al. (1995). The precursor protein of non-A beta component of Alzheimer's disease amyloid is a presynaptic protein of the central nervous system. *Neuron* 14, 467–475. doi: 10.1016/0896-6273(95)90302-X

Jiang, J., and Hui, C. C. (2008). Hedgehog signaling in development and cancer. *Dev. Cell* 15, 801–812. doi: 10.1016/j.devcel.2008.11.010

Joerger, A. C., and Fersht, A. R. (2007). Structural biology of the tumor suppressor p53 and cancer-associated mutants. *Adv. Cancer Res.* 97, 1–23. doi: 10.1016/S0065-230X(06)97001-8

Joerger, A. C., and Fersht, A. R. (2008). Structural biology of the tumor suppressor p53. *Annu. Rev. Biochem.* 77, 557–582. doi: 10.1146/annurev.biochem.77.060806.091238

Juhan-Vague, I., Pyke, S. D., Alessi, M. C., Jespersen, J., Haverkate, F., and Thompson, S. G. (1996). Fibrinolytic factors and the risk of myocardial infarction or sudden death in patients with angina pectoris. ECAT study group. european concerted action on thrombosis and disabilities. *Circulation* 94, 2057–2063. doi: 10.1161/01.CIR.94.9.2057

Kahali, B., and Ghosh, T. C. (2013). Disorderness in *Escherichia coli* proteome: perception of folding fidelity and protein-protein interactions. *J. Biomol. Struct. Dyn.* 31, 472–476. doi: 10.1080/07391102.2012.706071

Kalnina, Z., Zayakin, P., Silina, K., and Line, A. (2005). Alterations of pre-mRNA splicing in cancer. *Genes Chromosomes Cancer* 42, 342–357. doi: 10.1002/gcc.20156

Karagoz, G. E., Duarte, A. M., Akoury, E., Ippel, H., Biernat, J., Moran Luengo, T., et al. (2014). Hsp90-Tau complex reveals molecular basis for specificity in chaperone action. *Cell* 156, 963–974. doi: 10.1016/j.cell.2014.01.037

Kelly, J. W. (1998). The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr. Opin. Struct. Biol.* 8, 101–106. doi: 10.1016/S0959-440X(98)80016-X

Kirilyuk, A., Shimoji, M., Catania, J., Sahu, G., Pattabiraman, N., Giordano, A., et al. (2012). An intrinsically disordered region of the acetyltransferase p300 with similarity to prion-like domains plays a role in aggregation. *PLoS ONE* 7:e48243. doi: 10.1371/journal.pone.0048243

Korneta, I., and Bujnicki, J. M. (2012). Intrinsic disorder in the human spliceosomal proteome. *PLoS Comput. Biol.* 8:e1002641. doi: 10.1371/journal.pcbi.1002641

Kostic, M., Matt, T., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2006). Solution structure of the Hdm2 C2H2C4 RING, a domain critical for ubiquitination of p53. *J. Mol. Biol.* 363, 433–450. doi: 10.1016/j.jmb.2006.08.027

Kovacech, B., and Novak, M. (2010). Tau truncation is a productive posttranslational modification of neurofibrillary degeneration in Alzheimer's disease. *Curr. Alzheimer Res.* 7, 708–716. doi: 10.2174/156720510793611556

Krasowski, M. D., Reschly, E. J., and Ekins, S. (2008). Intrinsic disorder in nuclear hormone receptors. *J. Proteome Res.* 7, 4359–4372. doi: 10.1021/pr8003024

Kurotani, A., Tokmakov, A. A., Kuroda, Y., Fukami, Y., Shinozaki, K., and Sakurai, T. (2014). Correlations between predicted protein disorder and post-translational modifications in plants. *Bioinformatics* 30, 1095–1103. doi: 10.1093/bioinformatics/btt762

Kvansakul, M., and Hinds, M. G. (2014). The structural biology of BH3-Only proteins. *Methods Enzymol.* 544, 49–74. doi: 10.1016/B978-0-12-417158-9.00003-0

Lacronique, V., Boureux, A., Valle, V. D., Poirel, H., Quang, C. T., Mauchauffe, M., et al. (1997). A TEL-JAK2 fusion protein with constitutive kinase activity in human leukemia. *Science* 278, 1309–1312. doi: 10.1126/science.278.5341.1309

Ladd, A. N. (2013). CUG-BP, Elav-like family (CELF)-mediated alternative splicing regulation in the brain during health and disease. *Mol. Cell. Neurosci.* 56, 456–464. doi: 10.1016/j.mcn.2012.12.003

Ladomery, M. (2013). Aberrant alternative splicing is another hallmark of cancer. *Int. J. Cell Biol.* 2013, 463786. doi: 10.1155/2013/463786

Lamant, L., Dastugue, N., Pulford, K., Delsol, G., and Mariame, B. (1999). A new fusion gene TPM3-ALK in anaplastic large cell lymphoma created by a (1;2)(q25;p23) translocation. *Blood* 93, 3088–3095.

Lane, D. P. (1992). p53, guardian of the genome. *Nature* 358, 15–16. doi: 10.1038/358015a0

Lara-Pezzi, E., Dopazo, A., and Manzanares, M. (2012). Understanding cardiovascular disease: a journey through the genome (and what we found there). *Dis. Model Mech.* 5, 434–443. doi: 10.1242/dmm.009787

Lara-Pezzi, E., Gomez-Salinero, J., Gatto, A., and Garcia-Pavia, P. (2013). The alternative heart: impact of alternative splicing in heart disease. *J. Cardiovasc. Transl. Res.* 6, 945–955. doi: 10.1007/s12265-013-9482-z

Lazennec, G., Canaple, L., Saugy, D., and Wahli, W. (2000). Activation of peroxisome proliferator-activated receptors (PPARs) by their ligands and protein kinase A activators. *Mol. Endocrinol.* 14, 1962–1975. doi: 10.1210/mend.14.12.0575

Lee, C. W., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2010). Structure of the p53 transactivation domain in complex with the nuclear receptor coactivator binding domain of CREB binding protein. *Biochemistry* 49, 9964–9971. doi: 10.1021/bi1012996

Lee, H. K., Choi, Y. S., Park, Y. A., and Jeong, S. (2006). Modulation of oncogenic transcription and alternative splicing by beta-catenin and an RNA aptamer in colon cancer cells. *Cancer Res.* 66, 10560–10566. doi: 10.1158/0008-5472.CAN-06-2526

Lee, L., Sakurai, M., Matsuzaki, S., Arancio, O., and Fraser, P. (2013). SUMO and Alzheimer's disease. *Neuromolecular Med.* 15, 720–736. doi: 10.1007/s12017-013-8257-7

Le Gall, T., Romero, P. R., Cortese, M. S., Uversky, V. N., and Dunker, A. K. (2007). Intrinsic disorder in the Protein Data Bank. *J. Biomol. Struct. Dyn.* 24, 325–342. doi: 10.1080/07391102.2007.10507123

Lely, A. T., Luik, P. T., and Navis, G. (2007). Angiotensin I-converting enzyme: a pathogenetic role in diabetic renal damage? *Curr. Diabetes Rev.* 3, 41–52. doi: 10.2174/157339907779802049

Le Quesne, J. P., Spriggs, K. A., Bushell, M., and Willis, A. E. (2010). Dysregulation of protein synthesis and disease. *J. Pathol.* 220, 140–151. doi: 10.1002/path.2627

LeWinter, M. M. (2005). Functional consequences of sarcomeric protein abnormalities in failing myocardium. *Heart Fail. Rev.* 10, 249–257. doi: 10.1007/s10741-005-5254-4

Li, X., Gilkes, D., Li, B., Cheng, Q., Pernazza, D., Lawrence, H., et al. (2012). Abnormal MDMX degradation in tumor cells due to ARF deficiency. *Oncogene* 31, 3721–3732. doi: 10.1038/onc.2011.534

Liu, J., Perumal, N. B., Oldfield, C. J., Su, E. W., Uversky, V. N., and Dunker, A. K. (2006). Intrinsic disorder in transcription factors. *Biochemistry* 45, 6873–6888. doi: 10.1021/bi0602718

Liu, K., Lin, B., Zhao, M., Yang, X., Chen, M., Gao, A., et al. (2013). The multiple roles for Sox2 in stem cell maintenance and tumorigenesis. *Cell. Signal.* 25, 1264–1271. doi: 10.1016/j.cellsig.2013.02.013

Lobanov, M. Y., Shoemaker, B. A., Garbuzynskiy, S. O., Fong, J. H., Panchenko, A. R., and Galzitskaya, O. V. (2010). ComSin: database of protein structures in bound (complex) and unbound (single) states in relation to their intrinsic disorder. *Nucleic Acids Res.* 38, D283–D287. doi: 10.1093/nar/gkp963

Lokody, I. (2014). Alternative splicing: aberrant splicing promotes colon tumour growth. *Nat. Rev. Cancer* 14, 382–383. doi: 10.1038/nrc3753

Lorson, C. L., Rindt, H., and Shababi, M. (2010). Spinal muscular atrophy: mechanisms and therapeutic strategies. *Hum. Mol. Genet.* 19, R111–R118. doi: 10.1093/hmg/ddq147

Ma, K., Forbes, J. G., Gutierrez-Cruz, G., and Wang, K. (2006). Titin as a giant scaffold for integrating stress and Src homology domain 3-mediated signaling pathways: the clustering of novel overlap ligand motifs in the elastic PEVK segment. *J. Biol. Chem.* 281, 27539–27556. doi: 10.1074/jbc.M604525200

Makarenko, I., Opitz, C. A., Leake, M. C., Neagoe, C., Kulke, M., Gwathmey, J. K., et al. (2004). Passive stiffness changes caused by upregulation of compliant titin isoforms in human dilated cardiomyopathy hearts. *Circ. Res.* 95, 708–716. doi: 10.1161/01.RES.0000143901.37063.2f

Malaney, P., Pathak, R. R., Xue, B., Uversky, V. N., and Dave, V. (2013a). Intrinsic disorder in PTEN and its interactome confers structural plasticity and functional versatility. *Sci. Rep.* 3, 2035. doi: 10.1038/srep02035

Malaney, P., Uversky, V. N., and Dave, V. (2013b). The PTEN Long N-tail is intrinsically disordered: increased viability for PTEN therapy. *Mol. BioSyst.* 9, 2877–2888. doi: 10.1039/c3mb70267g

Mandard, S., Muller, M., and Kersten, S. (2004). Peroxisome proliferator-activated receptor alpha target genes. *Cell. Mol. Life Sci.* 61, 393–416. doi: 10.1007/s00018-003-3216-3

Marino-Enriquez, A., and Dal Cin, P. (2013). ALK as a paradigm of oncogenic promiscuity: different mechanisms of activation and different fusion partners drive tumors of different lineages. *Cancer Genet.* 206, 357–373. doi: 10.1016/j.cancergen.2013.07.001

Mark, W. Y., Liao, J. C., Lu, Y., Ayed, A., Laister, R., Szymczyna, B., et al. (2005). Characterization of segments from the central region of BRCA1: an intrinsically disordered scaffold for multiple protein-protein and protein-DNA interactions? *J. Mol. Biol.* 345, 275–287. doi: 10.1016/j.jmb.2004.10.045

Markiv, A., Rambaruth, N. D., and Dwek, M. V. (2012). Beyond the genome and proteome: targeting protein modifications in cancer. *Curr. Opin. Pharmacol.* 12, 408–413. doi: 10.1016/j.coph.2012.04.003

Marks, A. R. (2013). Calcium cycling proteins and heart failure: mechanisms and therapeutics. *J. Clin. Invest.* 123, 46–52. doi: 10.1172/JCI62834

Martinez-Yamout, M. A., Venkitakrishnan, R. P., Preece, N. E., Kroon, G., Wright, P. E., and Dyson, H. J. (2006). Localization of sites of interaction between p23 and Hsp90 in solution. *J. Biol. Chem.* 281, 14457–14464. doi: 10.1074/jbc.M601759200

McBurney, M. W., Clark-Knowles, K. V., Caron, A. Z., and Gray, D. A. (2013). SIRT1 is a highly networked protein that mediates the adaptation to chronic physiological stress. *Genes Cancer* 4, 125–134. doi: 10.1177/1947601912474893

McEwan, I. J. (2012). Intrinsic disorder in the androgen receptor: identification, characterisation and drugability. *Mol. BioSyst.* 8, 82–90. doi: 10.1039/c1mb05249g

McLarty, J. L., Marsh, S. A., and Chatham, J. C. (2013). Post-translational protein modification by O-linked N-acetyl-glucosamine: its role in mediating the adverse effects of diabetes on the heart. *Life Sci.* 92, 621–627. doi: 10.1016/j.lfs.2012.08.006

Mei, Y., Su, M., Soni, G., Salem, S., Colbert, C. L., and Sinha, S. C. (2014). Intrinsically disordered regions in autophagy proteins. *Proteins* 82, 565–578. doi: 10.1002/prot.24424

Midic, U., and Obradovic, Z. (2012). Intrinsic disorder in putative protein sequences. *Proteome Sci.* 10(Suppl. 1):S19. doi: 10.1186/1477-5956-10-S1-S19

Midic, U., Oldfield, C. J., Dunker, A. K., Obradovic, Z., and Uversky, V. N. (2009a). Protein disorder in the human diseasome: unfoldomics of human genetic diseases. *BMC Genomics* 10(Suppl. 1):S12. doi: 10.1186/1471-2164-10-S1-S12

Midic, U., Oldfield, C. J., Dunker, A. K., Obradovic, Z., and Uversky, V. N. (2009b). Unfoldomics of human genetic diseases: illustrative examples of ordered and intrinsically disordered members of the human diseasome. *Protein Pept. Lett.* 16, 1533–1547. doi: 10.2174/092986609789839377

Mills, J. D., and Janitz, M. (2012). Alternative splicing of mRNA in the molecular pathology of neurodegenerative diseases. *Neurobiol. Aging* 33, 1012.e11–1012.e24. doi: 10.1016/j.neurobiolaging.2011.10.030

Minde, D. P., Radli, M., Forneris, F., Maurice, M. M., and Rudiger, S. G. (2013). Large extent of disorder in Adenomatous Polyposis Coli offers a strategy to guard Wnt signalling against point mutations. *PLoS ONE* 8:e77257. doi: 10.1371/journal.pone.0077257

Mitrea, D. M., Grace, C. R., Buljan, M., Yun, M. K., Pytel, N. J., Satumba, J., et al. (2014). Structural polymorphism in the N-terminal oligomerization domain of NPM1. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4466–4471. doi: 10.1073/pnas.1321007111

Mitrea, D. M., and Kriwacki, R. W. (2013). Regulated unfolding of proteins in signaling. *FEBS Lett.* 587, 1081–1088. doi: 10.1016/j.febslet.2013.02.024

Mitrea, D. M., Yoon, M. K., Ou, L., and Kriwacki, R. W. (2012). Disorder-function relationships for the cell cycle regulatory proteins p21 and p27. *Biol. Chem.* 393, 259–274. doi: 10.1515/hsz-2011-0254

Mohan, A., Sullivan, W. J. Jr., Radivojac, P., Dunker, A. K., and Uversky, V. N. (2008). Intrinsic disorder in pathogenic and non-pathogenic microbes: discovering and analyzing the unfoldomes of early-branching eukaryotes. *Mol. BioSyst.* 4, 328–340. doi: 10.1039/b719168e

Moldoveanu, T., Grace, C. R., Llambi, F., Nourse, A., Fitzgerald, P., Gehring, K., et al. (2013). BID-induced structural changes in BAK promote apoptosis. *Nat. Struct. Mol. Biol.* 20, 589–597. doi: 10.1038/nsmb.2563

Molitch, M. E., DeFronzo, R. A., Franz, M. J., Keane, W. F., Mogensen, C. E., and Parving, H. H. (2003). Diabetic nephropathy. *Diabetes Care* 26(Suppl. 1), S94–S98. doi: 10.2337/diacare.26.2007.S94

Moreau, K. L., and King, J. A. (2012). Protein misfolding and aggregation in cataract disease and prospects for prevention. *Trends Mol. Med.* 18, 273–282. doi: 10.1016/j.molmed.2012.03.005

Moumne, L., Betuing, S., and Caboche, J. (2013). Multiple aspects of gene dysregulation in huntington's disease. *Front. Neurol.* 4:127. doi: 10.3389/fneur.2013.00127

Muggia, F., Safra, T., and Dubeau, L. (2011). BRCA genes: lessons learned from experimental and clinical cancer. *Ann. Oncol.* 22(Suppl. 1), i7–i10. doi: 10.1093/annonc/mdq659

Mulligan, V. K., and Chakrabartty, A. (2013). Protein misfolding in the late-onset neurodegenerative diseases: common themes and the unique case of amyotrophic lateral sclerosis. *Proteins* 81, 1285–1303. doi: 10.1002/prot.24285

Naro, C., and Sette, C. (2013). Phosphorylation-Mediated regulation of alternative splicing in cancer. *Int. J. Cell Biol.* 2013, 151839. doi: 10.1155/2013/151839

Ng, K. P., Potikyan, G., Savene, R. O., Denny, C. T., Uversky, V. N., and Lee, K. A. (2007). Multiple aromatic side chains within a disordered structure are critical for transcription and transforming activity of EWS family oncoproteins. *Proc. Natl. Acad. Sci. U.S.A.* 104, 479–484. doi: 10.1073/pnas.0607007104

Nielsen, S., Frokiaer, J., Marples, D., Kwon, T. H., Agre, P., and Knepper, M. A. (2002). Aquaporins in the kidney: from molecules to medicine. *Physiol. Rev.* 82, 205–244. doi: 10.1152/physrev.00024.2001

Norholm, A. B., Hendus-Altenburger, R., Bjerre, G., Kjaergaard, M., Pedersen, S. F., and Kragelund, B. B. (2011). The intracellular distal tail of the Na$^+$/H$^+$ exchanger NHE1 is intrinsically disordered: implications for NHE1 trafficking. *Biochemistry* 50, 3469–3480. doi: 10.1021/bi1019989

Noutsou, M., Duarte, A. M., Anvarian, Z., Didenko, T., Minde, D. P., Kuper, I., et al. (2011). Critical scaffolding regions of the tumor suppressor Axin1 are natively unfolded. *J. Mol. Biol.* 405, 773–786. doi: 10.1016/j.jmb.2010.11.013

Oates, M. E., Romero, P., Ishida, T., Ghalwash, M., Mizianty, M. J., Xue, B., et al. (2013). D(2)P(2): database of disordered protein predictions. *Nucleic Acids Res.* 41, D508–D516. doi: 10.1093/nar/gks1226

Obradovic, Z., Peng, K., Vucetic, S., Radivojac, P., Brown, C. J., and Dunker, A. K. (2003). Predicting intrinsic disorder from amino acid sequence. *Proteins* 53(Suppl. 6), 566–572. doi: 10.1002/prot.10532

Obradovic, Z., Peng, K., Vucetic, S., Radivojac, P., and Dunker, A. K. (2005). Exploiting heterogeneous sequence properties improves prediction of protein disorder. *Proteins* 61(Suppl. 7), 176–182. doi: 10.1002/prot.20735

Okumura, N., Yoshida, H., Kitagishi, Y., Nishimura, Y., and Matsuda, S. (2011). Alternative splicings on p53, BRCA1 and PTEN genes involved in breast cancer. *Biochem. Biophys. Res. Commun.* 413, 395–399. doi: 10.1016/j.bbrc.2011.08.098

Oldfield, C. J., Cheng, Y., Cortese, M. S., Brown, C. J., Uversky, V. N., and Dunker, A. K. (2005). Comparing and combining predictors of mostly disordered proteins. *Biochemistry* 44, 1989–2000. doi: 10.1021/bi047993o

Oldfield, C. J., Meng, J., Yang, J. Y., Yang, M. Q., Uversky, V. N., and Dunker, A. K. (2008). Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. *BMC Genomics* 9(Suppl. 1):S1. doi: 10.1186/1471-2164-9-S1-S1

Omenn, G. S., Guan, Y., and Menon, R. (2014). A new class of protein cancer biomarker candidates: differentially expressed splice variants of ERBB2 (HER2/neu) and ERBB1 (EGFR) in breast cancer cell lines. *J. Proteomics* 107C, 103–112. doi: 10.1016/j.jprot.2014.04.012

Omenn, G. S., Yocum, A. K., and Menon, R. (2010). Alternative splice variants, a new class of protein cancer biomarker candidates: findings in pancreatic cancer and breast cancer with systems biology implications. *Dis. Markers* 28, 241–251. doi: 10.1155/2010/705847

Orban, T. I., and Olah, E. (2001). Expression profiles of BRCA1 splice variants in asynchronous and in G1/S synchronized tumor cell lines. *Biochem. Biophys. Res. Commun.* 280, 32–38. doi: 10.1006/bbrc.2000.4068

Ostedgaard, L. S., Baldursson, O., Vermeer, D. W., Welsh, M. J., and Robertson, A. D. (2000). A functional R domain from cystic fibrosis transmembrane conductance regulator is predominantly unstructured in solution. *Proc. Natl. Acad. Sci. U.S.A.* 97, 5657–5662. doi: 10.1073/pnas.100588797

Ostenson, C. G., and Efendic, S. (2007). Islet gene expression and function in type 2 diabetes; studies in the Goto-Kakizaki rat and humans. *Diabetes Obes. Metab.* 9(Suppl. 2), 180–186. doi: 10.1111/j.1463-1326.2007.00787.x

Ostenson, C. G., Gaisano, H., Sheu, L., Tibell, A., and Bartfai, T. (2006). Impaired gene and protein expression of exocytotic soluble N-ethylmaleimide attachment protein receptor complex proteins in pancreatic islets of type 2 diabetic patients. *Diabetes* 55, 435–440. doi: 10.2337/diabetes.55.02.06.db04-1575

Ou, L., Waddell, M. B., and Kriwacki, R. W. (2012). Mechanism of cell cycle entry mediated by the intrinsically disordered protein p27(Kip1). *ACS Chem. Biol.* 7, 678–682. doi: 10.1021/cb200487h

Pal, S., Gupta, R., and Davuluri, R. V. (2012). Alternative transcription and alternative splicing in cancer. *Pharmacol. Ther.* 136, 283–294. doi: 10.1016/j.pharmthera.2012.08.005

Pancsa, R., and Tompa, P. (2012). Structural disorder in eukaryotes. *PLoS ONE* 7:e34687. doi: 10.1371/journal.pone.0034687

Park, J. J., and Lee, M. (2013). Increasing the alpha 2, 6 Sialylation of Glycoproteins may contribute to metastatic spread and therapeutic resistance in colorectal cancer. *Gut. Liver* 7, 629–641. doi: 10.5009/gnl.2013.7.6.629

Park, S., Mathis, K. W., and Lee, I. K. (2014). The physiological roles of apolipoprotein J/clusterin in metabolic and cardiovascular diseases. *Rev. Endocr. Metab. Disord.* 15, 45–53. doi: 10.1007/s11154-013-9275-3

Patil, A., and Nakamura, H. (2006). Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. *FEBS Lett.* 580, 2041–2045. doi: 10.1016/j.febslet.2006.03.003

Pejaver, V., Hsu, W. L., Xin, F., Dunker, A. K., Uversky, V. N., and Radivojac, P. (2014). The structural and functional signatures of proteins that undergo multiple events of post-translational modification. *Protein Sci.* 23, 1077–1093. doi: 10.1002/pro.2494

Peng, Z., Mizianty, M. J., Xue, B., Kurgan, L., and Uversky, V. N. (2012). More than just tails: intrinsic disorder in histone proteins. *Mol. BioSyst.* 8, 1886–1901. doi: 10.1039/c2mb25102g

Peng, Z., Xue, B., Kurgan, L., and Uversky, V. N. (2013). Resilience of death: intrinsic disorder in proteins involved in the programmed cell death. *Cell Death Differ.* 20, 1257–1267. doi: 10.1038/cdd.2013.65

Pentony, M. M., and Jones, D. T. (2010). Modularity of intrinsic disorder in the human proteome. *Proteins* 78, 212–221. doi: 10.1002/prot.22504

Permyakov, S. E., Ismailov, R. G., Xue, B., Denesyuk, A. I., Uversky, V. N., and Permyakov, E. A. (2011). Intrinsic disorder in S100 proteins. *Mol. Biosyst.* 7, 2164–2180. doi: 10.1039/c0mb00305k

Peterson, E. A., and Petty, E. M. (2010). Conquering the complex world of human septins: implications for health and disease. *Clin. Genet.* 77, 511–524. doi: 10.1111/j.1399-0004.2010.01392.x

Peysselon, F., Xue, B., Uversky, V. N., and Ricard-Blum, S. (2011). Intrinsic disorder of the extracellular matrix. *Mol. Biosyst.* 7, 3353–3365. doi: 10.1039/c1mb05316g

Philips, A. V., and Cooper, T. A. (2000). RNA processing and human disease. *Cell. Mol. Life Sci.* 57, 235–249. doi: 10.1007/PL00000687

Pietrosemoli, N., Pancsa, R., and Tompa, P. (2013). Structural disorder provides increased adaptability for vesicle trafficking pathways. *PLoS Comput. Biol.* 9:e1003144. doi: 10.1371/journal.pcbi.1003144

Pihlstrom, L., and Toft, M. (2011). Genetic variability in SNCA and Parkinson's disease. *Neurogenetics* 12, 283–293. doi: 10.1007/s10048-011-0292-7

Platzer, G., Schedlbauer, A., Chemelli, A., Ozdowy, P., Coudevylle, N., Auer, R., et al. (2011). The metastasis-associated extracellular matrix protein osteopontin forms transient structure in ligand interaction sites. *Biochemistry* 50, 6113–6124. doi: 10.1021/bi200291e

Ploplis, V. A. (2011). Effects of altered plasminogen activator inhibitor-1 expression on cardiovascular disease. *Curr. Drug Targets* 12, 1782–1789. doi: 10.2174/138945011797635803

Polverino de Laureto, P., De Filippis, V., Di Bello, M., Zambonin, M., and Fontana, A. (1995). Probing the molten globule state of alpha-lactalbumin by limited proteolysis. *Biochemistry* 34, 12596–12604. doi: 10.1021/bi00039a015

Proctor, D. T., Coulson, E. J., and Dodd, P. R. (2011). Post-synaptic scaffolding protein interactions with glutamate receptors in synaptic dysfunction and Alzheimer's disease. *Prog. Neurobiol.* 93, 509–521. doi: 10.1016/j.pneurobio.2011.02.002

Qiu, H., Dai, H., Jain, K., Shah, R., Hong, C., Pain, J., et al. (2008). Characterization of a novel cardiac isoform of the cell cycle-related kinase that is regulated during heart failure. *J. Biol. Chem.* 283, 22157–22165. doi: 10.1074/jbc.M710459200

Radivojac, P., Vucetic, S., O'Connor, T. R., Uversky, V. N., Obradovic, Z., and Dunker, A. K. (2006). Calmodulin signaling: analysis and prediction of a disorder-dependent molecular recognition. *Proteins* 63, 398–410. doi: 10.1002/prot.20873

Rahmutulla, B., Matsushita, K., Satoh, M., Seimiya, M., Tsuchida, S., Kubo, S., et al. (2013). Alternative splicing of FBP-interacting repressor coordinates c-Myc, P27Kip1/cyclinE and Ku86/XRCC5 expression as a molecular sensor for bleomycin-induced DNA damage pathway. *Oncotarget* 5, 2404–2017.

Rajagopalan, K., Mooney, S. M., Parekh, N., Getzenberg, R. H., and Kulkarni, P. (2011). A majority of the cancer/testis antigens are intrinsically disordered proteins. *J. Cell. Biochem.* 112, 3256–3267. doi: 10.1002/jcb.23252

Roccato, E., Pagliardini, S., Cleris, L., Canevari, S., Formelli, F., Pierotti, M. A., et al. (2003). Role of TFG sequences outside the coiled-coil domain in TRK-T3 oncogenic activation. *Oncogene* 22, 807–818. doi: 10.1038/sj.onc.1206189

Rochet, J. C., and Lansbury, P. T. Jr. (2000). Amyloid fibrillogenesis: themes and variations. *Curr. Opin. Struct. Biol.* 10, 60–68. doi: 10.1016/S0959-440X(99)00049-4

Rodriguez, F., Arsene-Ploetze, F., Rist, W., Rudiger, S., Schneider-Mergener, J., Mayer, M. P., et al. (2008). Molecular basis for regulation of the heat shock transcription factor sigma32 by the DnaK and DnaJ chaperones. *Mol. Cell* 32, 347–358. doi: 10.1016/j.molcata.2008.01.004

Romero, P. R., Zaidi, S., Fang, Y. Y., Uversky, V. N., Radivojac, P., Oldfield, C. J., et al. (2006). Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms. *Proc. Natl. Acad. Sci. U.S.A.* 103, 8390–8395. doi: 10.1073/pnas.0507916103

Rotem, S., Katz, C., and Friedler, A. (2007). Insights into the structure and protein-protein interactions of the pro-apoptotic protein ASPP2. *Biochem. Soc. Trans.* 35(pt 5), 966–969. doi: 10.1042/BST0350966

Ruan, W., and Lai, M. (2010). Insulin-like growth factor binding protein: a possible marker for the metabolic syndrome? *Acta Diabetol.* 47, 5–14. doi: 10.1007/s00592-009-0142-3

Rudiger, S., Freund, S. M., Veprintsev, D. B., and Fersht, A. R. (2002). CRINEPT-TROSY NMR reveals p53 core domain bound in an unfolded form to the chaperone Hsp90. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11085–11090. doi: 10.1073/pnas.132393699

Ruiz i Altaba, A. (1999). Gli proteins encode context-dependent positive and negative functions: implications for development and disease. *Development* 126, 3205–3216.

Safar, J. G. (2012). Molecular pathogenesis of sporadic prion diseases in man. *Prion* 6, 108–115. doi: 10.4161/pri.18666

Schad, E., Tompa, P., and Hegyi, H. (2011). The relationship between proteome size, structural disorder and organism complexity. *Genome Biol.* 12, R120. doi: 10.1186/gb-2011-12-12-r120

Schiano, C., Casamassimi, A., Rienzo, M., de Nigris, F., Sommese, L., and Napoli, C. (2014). Involvement of Mediator complex in malignancy. *Biochim. Biophys. Acta* 1845, 66–83. doi: 10.1016/j.bbcan.2013.12.001

Sesti, G. (2000). Insulin receptor variant forms and type 2 diabetes mellitus. *Pharmacogenomics* 1, 49–61. doi: 10.1517/14622416.1.1.49

Sesti, G., Federici, M., Lauro, D., Sbraccia, P., and Lauro, R. (2001). Molecular mechanism of insulin resistance in type 2 diabetes mellitus: role of the insulin receptor variant forms. *Diabetes Metab. Res. Rev.* 17, 363–373. doi: 10.1002/dmrr.225

Sette, C. (2013). Alternative splicing programs in prostate cancer. *Int. J. Cell Biol.* 2013, 458727. doi: 10.1155/2013/458727

Shang, L. L., Pfahnl, A. E., Sanyal, S., Jiao, Z., Allen, J., Banach, K., et al. (2007). Human heart failure is associated with abnormal C-terminal splicing variants in the cardiac sodium channel. *Circ. Res.* 101, 1146–1154. doi: 10.1161/CIRCRESAHA.107.152918

Shareef, M. A., Anwer, L. A., and Poizat, C. (2014). Cardiac SERCA2A/B: therapeutic targets for heart failure. *Eur. J. Pharmacol.* 724, 1–8. doi: 10.1016/j.ejphar.2013.12.018

Sharma, A., Costantini, S., and Colonna, G. (2013). The protein-protein interaction network of the human Sirtuin family. *Biochim. Biophys. Acta* 1834, 1998–2009. doi: 10.1016/j.bbapap.2013.06.012

Shibata, K., Shimokawa, H., Yanagihara, N., Otsuji, Y., and Tsutsui, M. (2013). Nitric oxide synthases and heart failure - lessons from genetically manipulated mice. *J. UOEH* 35, 147–158. doi: 10.7888/juoeh.35.147

Shimizu, K., and Toh, H. (2009). Interaction between intrinsically disordered proteins frequently occurs in a human protein-protein interaction network. *J. Mol. Biol.* 392, 1253–1265. doi: 10.1016/j.jmb.2009.07.088

Shkreta, L., Bell, B., Revil, T., Venables, J. P., Prinos, P., Elela, S. A., et al. (2013). Cancer-associated perturbations in alternative Pre-messenger RNA splicing. *Cancer Treat Res.* 158, 41–94. doi: 10.1007/978-3-642-31659-3_3

Sickmeier, M., Hamilton, J. A., LeGall, T., Vacic, V., Cortese, M. S., Tantos, A., et al. (2007). DisProt: the database of disordered proteins. *Nucleic Acids Res.* 35, D786–D793. doi: 10.1093/nar/gkl893

Singh, G. P., and Dash, D. (2007). Intrinsic disorder in yeast transcriptional regulatory network. *Proteins* 68, 602–605. doi: 10.1002/prot.21497

Singh, G. P., Ganapathi, M., and Dash, D. (2007). Role of intrinsic disorder in transient interactions of hub proteins. *Proteins* 66, 761–765. doi: 10.1002/prot.21281

Sivadas, V. P., Gulati, S., Varghese, B. T., Balan, A., and Kannan, S. (2014). The early manifestation, tumor-specific occurrence and prognostic significance of TGFBR2 aberrant splicing in oral carcinoma. *Exp. Cell Res.* doi: 10.1016/j.yexcr.2014.05.004. [Epub ahead of print].

Soda, M., Choi, Y. L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., et al. (2007). Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature* 448, 561–566. doi: 10.1038/nature05945

Splawski, I., Timothy, K. W., Sharpe, L. M., Decher, N., Kumar, P., Bloise, R., et al. (2004). Ca(V)1.2 calcium channel dysfunction causes a multisystem disorder including arrhythmia and autism. *Cell* 119, 19–31. doi: 10.1016/j.cell.2004.09.011

Srebrow, A., and Kornblihtt, A. R. (2006). The connection between splicing and cancer. *J. Cell Sci.* 119(pt 13), 2635–2641. doi: 10.1242/jcs.03053

Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., et al. (2005). Function of alternative splicing. *Gene* 344, 1–20. doi: 10.1016/j.gene.2004.10.022

Suskiewicz, M. J., Sussman, J. L., Silman, I., and Shaul, Y. (2011). Context-dependent resistance to proteolysis of intrinsically disordered proteins. *Protein Sci.* 20, 1285–1297. doi: 10.1002/pro.657

Takahashi, M. (2001). The GDNF/RET signaling pathway and human diseases. *Cytokine Growth Factor Rev.* 12, 361–373. doi: 10.1016/S1359-6101(01)00012-0

Tamarappoo, B. K., and Verkman, A. S. (1998). Defective aquaporin-2 trafficking in nephrogenic diabetes insipidus and correction by chemical chaperones. *J. Clin. Invest.* 101, 2257–2267. doi: 10.1172/JCI2303

Tan, F., Thiele, C. J., and Li, Z. (2014). Collapsin response mediator proteins: potential diagnostic and prognostic biomarkers in cancers (Review). *Oncol. Lett.* 7, 1333–1340. doi: 10.3892/ol.2014.1909

Theocharis, A. D., Skandalis, S. S., Tzanakakis, G. N., and Karamanos, N. K. (2010). Proteoglycans in health and disease: novel roles for proteoglycans in malignancy and their pharmacological targeting. *FEBS J.* 277, 3904–3923. doi: 10.1111/j.1742-4658.2010.07800.x

Thrailkill, K. M., Clay Bunn, R., and Fowlkes, J. L. (2009). Matrix metalloproteinases: their potential role in the pathogenesis of diabetic nephropathy. *Endocrine* 35, 1–10. doi: 10.1007/s12020-008-9114-6

Tian, Y., Yao, Z., Roden, R. B., and Zhang, H. (2011a). Identification of glycoproteins associated with different histological subtypes of ovarian tumors using quantitative glycoproteomics. *Proteomics* 11, 4677–4687. doi: 10.1002/pmic.201000811

Tian, Y., Bova, G. S., and Zhang, H. (2011b). Quantitative glycoproteomic analysis of optimal cutting temperature-embedded frozen tissues identifying glycoproteins associated with aggressive prostate cancer. *Anal. Chem.* 83, 7013–7019. doi: 10.1021/ac200815q

Tian, Y., and Zhang, H. (2013). Characterization of disease-associated N-linked glycoproteins. *Proteomics* 13, 504–511. doi: 10.1002/pmic.201200333

Tokuriki, N., Oldfield, C. J., Uversky, V. N., Berezovsky, I. N., and Tawfik, D. S. (2009). Do viral proteins possess unique biophysical features? *Trends Biochem. Sci.* 34, 53–59. doi: 10.1016/j.tibs.2008.10.009

Tompa, P. (2002). Intrinsically unstructured proteins. *Trends Biochem. Sci.* 27, 527–533. doi: 10.1016/S0968-0004(02)02169-2

Tompa, P. (2005). The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett.* 579, 3346–3354. doi: 10.1016/j.febslet.2005.03.072

Tompa, P., Dosztanyi, Z., and Simon, I. (2006). Prevalent structural disorder in *E. coli* and *S. cerevisiae* proteomes. *J. Proteome Res.* 5, 1996–2000. doi: 10.1021/pr0600881

Tompa, P., and Kalmar, L. (2010). Power law distribution defines structural disorder as a structural element directly linked with function. *J. Mol. Biol.* 403, 346–350. doi: 10.1016/j.jmb.2010.07.044

Toth-Petroczy, A., Oldfield, C. J., Simon, I., Takagi, Y., Dunker, A. K., Uversky, V. N., et al. (2008). Malleable machines in transcription regulation: the mediator complex. *PLoS Comput. Biol.* 4:e1000243. doi: 10.1371/journal.pcbi.1000243

Trinei, M., Lanfrancone, L., Campo, E., Pulford, K., Mason, D. Y., Pelicci, P. G., et al. (2000). A new variant anaplastic lymphoma kinase (ALK)-fusion protein (ATIC-ALK) in a case of ALK-positive anaplastic large cell lymphoma. *Cancer Res.* 60, 793–798.

Tsvetkov, P., Asher, G., Paz, A., Reuven, N., Sussman, J. L., Silman, I., et al. (2008). Operational definition of intrinsically unstructured protein sequences based on susceptibility to the 20S proteasome. *Proteins* 70, 1357–1366. doi: 10.1002/prot.21614

Tsvetkov, P., Myers, N., Moscovitz, O., Sharon, M., Prilusky, J., and Shaul, Y. (2012). Thermo-resistant intrinsically disordered proteins are efficient 20S proteasome substrates. *Mol. Biosyst.* 8, 368–373. doi: 10.1039/c1mb05283g

Tsvetkov, P., Reuven, N., and Shaul, Y. (2009a). The nanny model for IDPs. *Nat. Chem. Biol.* 5, 778–781. doi: 10.1038/nchembio.233

Tsvetkov, P., Reuven, N., Prives, C., and Shaul, Y. (2009b). Susceptibility of p53 unstructured N terminus to 20 s proteasomal degradation programs the stress response. *J. Biol. Chem.* 284, 26234–26242. doi: 10.1074/jbc.M109.040493

Tuccillo, F. M., de Laurentiis, A., Palmieri, C., Fiume, G., Bonelli, P., Borrelli, A., et al. (2014). Aberrant glycosylation as biomarker for cancer: focus on CD43. *Biomed. Res. Int.* 2014, 742831. doi: 10.1155/2014/742831

Uversky, V. N. (2003). A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders. *J. Biomol. Struct. Dyn.* 21, 211–234. doi: 10.1080/07391102.2003.10506918

Uversky, V. N. (2007). Neuropathology, biochemistry, and biophysics of alpha-synuclein aggregation. *J. Neurochem.* 103, 17–37. doi: 10.1111/j.1471-4159.2007.04764.x

Uversky, V. N. (2008a). Amyloidogenesis of natively unfolded proteins. *Curr. Alzheimer Res.* 5, 260–287. doi: 10.2174/156720508784533312

Uversky, V. N. (2008b). Alpha-synuclein misfolding and neurodegenerative diseases. *Curr. Protein Pept. Sci.* 9, 507–540. doi: 10.2174/138920308785915218

Uversky, V. N. (2009). Intrinsic disorder in proteins associated with neurodegenerative diseases. *Front. Biosci. (Landmark Ed.)* 14:5188–5238. doi: 10.2741/3594

Uversky, V. N. (2010a). The mysterious unfoldome: structureless, underappreciated, yet vital part of any given proteome. *J. Biomed. Biotechnol.* 2010, 568068. doi: 10.1155/2010/568068

Uversky, V. N. (2010b). Targeting intrinsically disordered proteins in neurodegenerative and protein dysfunction diseases: another illustration of the D(2) concept. *Expert. Rev. Proteomics* 7, 543–564. doi: 10.1586/epr.10.36

Uversky, V. N. (2012). Intrinsically disordered proteins and novel strategies for drug discovery. *Expert. Opin. Drug. Discov.* 7, 475–488. doi: 10.1517/17460441.2012.686489

Uversky, V. N. (2013). Intrinsic disorder-based protein interactions and their modulators. *Curr. Pharm. Des.* 19, 4191–4213. doi: 10.2174/1381612811319230005

Uversky, V. N. (2014). The triple power of D(3): protein intrinsic disorder in degenerative diseases. *Front. Biosci. (Landmark Ed)* 19:181–258. doi: 10.2741/4204

Uversky, V. N., Dave, V., Iakoucheva, L. M., Malaney, P., Metallo, S. J., Pathak, R. R., et al. (2014). Pathological unfoldomics of uncontrolled chaos: intrinsically disordered proteins and human diseases. *Chem. Rev.* 114, 6844–6879. doi: 10.1021/cr400713r

Uversky, V. N., and Dunker, A. K. (2008). Biochemistry. Controlled chaos. *Science* 322, 1340–1341. doi: 10.1126/science.1167453

Uversky, V. N., and Dunker, A. K. (2010). Understanding protein non-folding. *Biochim. Biophys. Acta* 1804, 1231–1264. doi: 10.1016/j.bbapap.2010.01.017

Uversky, V. N., and Eliezer, D. (2009). Biophysics of Parkinson's disease: structure and aggregation of alpha-synuclein. *Curr. Protein Pept. Sci.* 10, 483–499. doi: 10.2174/138920309789351921

Uversky, V. N., and Fink, A. L. (2004). Conformational constraints for amyloid fibrillation: the importance of being unfolded. *Biochim. Biophys. Acta.* 1698, 131–153. doi: 10.1016/j.bbapap.2003.12.008

Uversky, V. N., Gillespie, J. R., and Fink, A. L. (2000). Why are natively unfolded proteins unstructured under physiologic conditions? *Proteins* 41, 415–427. doi: 10.1002/1097-0134(20001115)41:3%3C415::AID-PROT130%3E3.3.CO;2-Z

Uversky, V. N., Oldfield, C. J., and Dunker, A. K. (2005). Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J. Mol. Recognit.* 18, 343–384. doi: 10.1002/jmr.747

Uversky, V. N., Oldfield, C. J., and Dunker, A. K. (2008). Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu. Rev. Biophys.* 37, 215–246. doi: 10.1146/annurev.biophys.37.032807.125924

Uversky, V. N., Oldfield, C. J., Midic, U., Xie, H., Xue, B., Vucetic, S., et al. (2009). Unfoldomics of human diseases: linking protein intrinsic disorder with diseases. *BMC Genomics* 10(Suppl. 1):S7. doi: 10.1186/1471-2164-10-S1-S7

Uversky, V. N., Roman, A., Oldfield, C. J., and Dunker, A. K. (2006). Protein intrinsic disorder and human papillomaviruses: increased amount of disorder in E6 and E7 oncoproteins from high risk HPVs. *J. Proteome Res.* 5, 1829–1842. doi: 10.1021/pr0602388

Uversky, V. N., Talapatra, A., Gillespie, J. R., and Fink, A. L. (1999a). Protein deposits as the molecular basis of amyloidosis. I. Systemic amyloidoses. *Med. Sci. Monitor* 5, 1001–1012.

Uversky, V. N., Talapatra, A., Gillespie, J. R., and Fink, A. L. (1999b). Protein deposits as the molecular basis of amyloidosis. II. Localized amyloidosis and neurodegenerative disordres. *Med. Sci. Monitor* 5, 1238–1254.

Vacic, V., and Iakoucheva, L. M. (2012). Disease mutations in disordered regions–exception to the rule? *Mol. Biosyst.* 8, 27–32. doi: 10.1039/c1mb05251a

Vacic, V., Markwick, P. R., Oldfield, C. J., Zhao, X., Haynes, C., Uversky, V. N., et al. (2012). Disease-associated mutations disrupt functionally important regions of intrinsic protein disorder. *PLoS Comput. Biol.* 8:e1002709. doi: 10.1371/journal.pcbi.1002709

Varadi, M., Kosol, S., Lebrun, P., Valentini, E., Blackledge, M., Dunker, A. K., et al. (2014). pE-DB: a database of structural ensembles of intrinsically disordered and of unfolded proteins. *Nucleic Acids Res.* 42, D326–D335. doi: 10.1093/nar/gkt960

Vavouri, T., Semple, J. I., Garcia-Verdugo, R., and Lehner, B. (2009). Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity. *Cell* 138, 198–208. doi: 10.1016/j.cell.2009.04.029

Verges, B. (2010). Abnormal hepatic apolipoprotein B metabolism in type 2 diabetes. *Atherosclerosis* 211, 353–360. doi: 10.1016/j.atherosclerosis.2010.01.028

Vousden, K. H., and Prives, C. (2009). Blinded by the light: the growing complexity of p53. *Cell* 137, 413–431. doi: 10.1016/j.cell.2009.04.037

Vucetic, S., Obradovic, Z., Vacic, V., Radivojac, P., Peng, K., Iakoucheva, L. M., et al. (2005). DisProt: a database of protein disorder. *Bioinformatics* 21, 137–140. doi: 10.1093/bioinformatics/bth476

Vucetic, S., Xie, H., Iakoucheva, L. M., Oldfield, C. J., Dunker, A. K., Obradovic, Z., et al. (2007). 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.* 6, 1899–1916. doi: 10.1021/pr060393m

Walker, L. C., and LeVine, H. 3rd. (2012). Corruption and spread of pathogenic proteins in neurodegenerative diseases. *J. Biol. Chem.* 287, 33109–33115. doi: 10.1074/jbc.R112.399378

Wang, J. Z., Xia, Y. Y., Grundke-Iqbal, I., and Iqbal, K. (2013). Abnormal hyperphosphorylation of tau: sites, regulation, and molecular mechanism of neurofibrillary degeneration. *J. Alzheimers Dis.* 33(Suppl. 1), S123–S139. doi: 10.3233/JAD-2012-129031

Wang, Y., Fisher, J. C., Mathew, R., Ou, L., Otieno, S., Sublet, J., et al. (2011). Intrinsic disorder mediates the diverse regulatory functions of the Cdk inhibitor p21. *Nat. Chem. Biol.* 7, 214–221. doi: 10.1038/nchembio.536

Wang, Z., Gucek, M., and Hart, G. W. (2008). Cross-talk between GlcNAcylation and phosphorylation: site-specific phosphorylation dynamics in response to globally elevated O-GlcNAc. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13793–13798. doi: 10.1073/pnas.0806216105

Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., and Jones, D. T. (2004). Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* 337, 635–645. doi: 10.1016/j.jmb.2004.02.002

Weatheritt, R. J., and Gibson, T. J. (2012). Linear motifs: lost in (pre)translation. *Trends Biochem. Sci.* 37, 333–341. doi: 10.1016/j.tibs.2012.05.001

Wegiel, J., Gong, C. X., and Hwang, Y. W. (2011). The role of DYRK1A in neurodegenerative diseases. *FEBS J.* 278, 236–245. doi: 10.1111/j.1742-4658.2010.07955.x

Wells, M., Tidow, H., Rutherford, T. J., Markwick, P., Jensen, M. R., Mylonas, E., et al. (2008). Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5762–5767. doi: 10.1073/pnas.0801353105

Wiggins, C. M., Tsvetkov, P., Johnson, M., Joyce, C. L., Lamb, C. A., Bryant, N. J., et al. (2011). BIM(EL), an intrinsically disordered protein, is degraded by 20S proteasomes in the absence of poly-ubiquitylation. *J. Cell Sci.* 124(pt 6), 969–977. doi: 10.1242/jcs.058438

Willems, S., Hoefer, I., and Pasterkamp, G. (2012). The role of the Interleukin 1 receptor-like 1 (ST2) and Interleukin-33 pathway in cardiovascular disease and cardiovascular risk assessment. *Minerva Med.* 103, 513–524.

Wojciak, J. M., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2009). Structural basis for recruitment of CBP/p300 coactivators by STAT1 and STAT2 transactivation domains. *EMBO J.* 28, 948–958. doi: 10.1038/emboj.2009.30

Wright, P. E., and Dyson, H. J. (1999). Intrinsically unstructured proteins: reassessing the protein structure-function paradigm. *J. Mol. Biol.* 293, 321–331. doi: 10.1006/jmbi.1999.3110

Xie, H., Vucetic, S., Iakoucheva, L. M., Oldfield, C. J., Dunker, A. K., Obradovic, Z., et al. (2007a). Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. *J. Proteome Res.* 6, 1917–1932. doi: 10.1021/pr060394e

Xie, H., Vucetic, S., Iakoucheva, L. M., Oldfield, C. J., Dunker, A. K., Uversky, V. N., et al. (2007b). Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. *J. Proteome Res.* 6, 1882–1898. doi: 10.1021/pr060392u

Xu, Y., Gao, X. D., Lee, J. H., Huang, H., Tan, H., Ahn, J., et al. (2014). Cell type-restricted activity of hnRNPM promotes breast cancer metastasis via regulating alternative splicing. *Genes Dev.* 28, 1191–1203. doi: 10.1101/gad.241968.114

Xue, B., Dunker, A. K., and Uversky, V. N. (2012a). Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. *J. Biomol. Struct. Dyn.* 30, 137–149. doi: 10.1080/07391102.2012.675145

Xue, B., Dunker, A. K., and Uversky, V. N. (2012b). The roles of intrinsic disorder in orchestrating the Wnt-pathway. *J. Biomol. Struct. Dyn.* 29, 843–861. doi: 10.1080/073911012010525024

Xue, B., Oldfield, C. J., Van, Y. Y., Dunker, A. K., and Uversky, V. N. (2012c). Protein intrinsic disorder and induced pluripotent stem cells. *Mol. Biosyst.* 8, 134–150. doi: 10.1039/c1mb05163f

Xue, B., Romero, P. R., Noutsou, M., Maurice, M. M., Rudiger, S. G., William, A. M. Jr., et al. (2013). Stochastic machines as a colocalization mechanism for scaffold protein function. *FEBS Lett.* 587, 1587–1591. doi: 10.1016/j.febslet.2013.04.006

Xue, B., Williams, R. W., Oldfield, C. J., Dunker, A. K., and Uversky, V. N. (2010a). Archaic chaos: intrinsically disordered proteins in Archaea. *BMC Syst. Biol.* 4(Suppl. 1):S1. doi: 10.1186/1752-0509-4-S1-S1

Xue, B., Dunbrack, R. L., Williams, R. W., Dunker, A. K., and Uversky, V. N. (2010b). PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochim. Biophys. Acta* 1804, 996–1010. doi: 10.1016/j.bbapap.2010.01.011

Yates, C. M., and Sternberg, M. J. E. (2013). The effects of non-synonymous single nucleotide polymorphisms (nsSNPs) on protein protein interactions. *J. Mol. Biol.* 425, 3949–3963. doi: 10.1016/j.jmb.2013.07.012

Yoon, M. K., Mitrea, D. M., Ou, L., and Kriwacki, R. W. (2012). Cell cycle regulation by the intrinsically disordered proteins p21 and p27. *Biochem. Soc. Trans.* 40, 981–988. doi: 10.1042/BST20120092

Zahn, R., Liu, A., Luhrs, T., Riek, R., von Schroetter, C., Lopez Garcia, F., et al. (2000). NMR solution structure of the human prion protein. *Proc. Natl. Acad. Sci. U.S.A.* 97, 145–150. doi: 10.1073/pnas.97.1.145

Zeng, X., Hood, B. L., Sun, M., Conrads, T. P., Day, R. S., Weissfeld, J. L., et al. (2010). Lung cancer serum biomarker discovery using glycoprotein capture and liquid chromatography mass spectrometry. *J. Proteome Res.* 9, 6440–6449. doi: 10.1021/pr100696n

Zhan, J., Easton, J. B., Huang, S., Mishra, A., Xiao, L., Lacy, E. R., et al. (2007). Negative regulation of ASK1 by p21Cip1 involves a small domain that includes Serine 98 that is phosphorylated by ASK1 *in vivo*. *Mol. Cell. Biol.* 27, 3530–3541. doi: 10.1128/MCB.00086-06

Zhang, L., Smit-McBride, Z., Pan, X., Rheinhardt, J., and Hershey, J. W. (2008). An oncogenic role for the phosphorylated h-subunit of human translation initiation factor eIF3. *J. Biol. Chem.* 283, 24047–24060. doi: 10.1074/jbc.M800956200

Zhang, S. S., and Shaw, R. M. (2013). Multilayered regulation of cardiac ion channels. *Biochim. Biophys. Acta* 1833, 876–885. doi: 10.1016/j.bbamcr.2012.10.020

Zimmermann, M. (2013). Neuronal AChE splice variants and their non-hydrolytic functions: redefining a target of AChE inhibitors? *Br. J. Pharmacol.* 170, 953–967. doi: 10.1111/bph.12359

Zorzano, A., Hernandez-Alvarez, M. I., Palacin, M., and Mingrone, G. (2010). Alterations in the mitochondrial regulatory pathways constituted by the nuclear co-factors PGC-1alpha or PGC-1beta and mitofusin 2 in skeletal muscle in type 2 diabetes. *Biochim. Biophys. Acta* 1797, 1028–1033. doi: 10.1016/j.bbabio.2010.02.017

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.