CHEMICAL REVIEWS

Review

pubs.acs.org/CR

Pathological Unfoldomics of Uncontrolled Chaos: Intrinsically Disordered Proteins and Human Diseases

Vladimir N. Uversky,*[,†,‡,§] Vrushank Davé,[∥,⊥] Lilia M. Iakoucheva,[#] Prerna Malaney,[∥] Steven J. Metallo,[▽] Ravi Ramesh Pathak,[∥] and Andreas C. Joerger*[,▲]

[†]Department of Molecular Medicine and USF Health Byrd Alzheimer's Research Institute and [∥]Department of Pathology and Cell Biology, Morsani College of Medicine, University of South Florida, Tampa, Florida 33612, United States

[‡]Institute for Biological Instrumentation, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia

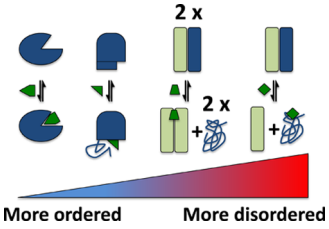
[§]Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah 22254, Saudi Arabia

[⊥]Department of Molecular Oncology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida 33612, United States

[#]Department of Psychiatry, University of California San Diego, La Jolla, California 92093, United States

[▽]Department of Chemistry, Georgetown University, Washington, District of Columbia 20057, United States

[▲]Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, United Kingdom



CONTENTS

1. INTRODUCTION

Many biologically important proteins lack stable tertiary and/or secondary structure under physiological conditions in vitro as a whole or in part.[1−5] These intrinsically disordered proteins

Special Issue: 2014 Intrinsically Disordered Proteins (IDPs)

Received: December 15, 2013

 © XXXX American Chemical Society dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX−XXX

(IDPs), or intrinsically disordered protein regions (IDPRs) of hybrid proteins possessing both structured and disordered domains, do not have unique well-defined 3D structures, existing instead as collapsed or extended dynamically mobile conformational ensembles. Therefore, natural proteins can be found in one of three major protein forms: functional and folded, nonfunctional and misfolded, or functional and intrinsically disordered. Although IDPs and IDPRs are highly dynamic, their structures can be described reasonably well by a rather limited number of lower-energy conformations.[6,7] The structural plasticity and conformational adaptability of IDPs/IDPRs and their intrinsic lack of rigid structure leads to a number of exceptional functional advantages, providing them with unique capabilities to act in functional modes not achievable by ordered proteins.[5] As a result, intrinsic disorder is a common feature of proteins involved in signaling, regulation, and recognition, and IDPs/IDPRs play diverse roles in modulation and control of their binding partners' functions and in promoting the assembly of supramolecular complexes. The biological actions of IDPs/IDPRs, which frequently serve as major regulators of their binding partners, are controlled by extensive posttranslational modifications (PTMs), such as phosphorylation, acetylation, ubiquitination, and sumoylation,[5] and by alternative splicing.[8] In fact, many IDPs/IDPRs are known to contain multiple functional elements that contribute to their ability to be involved in interaction with, regulation of, and control by multiple structurally unrelated partners.[9] Given the existence of multiple functions in a single disordered protein, and given that each functional element is typically relatively short, alternative splicing could readily generate sets of protein isoforms with highly diverse regulatory elements.[8] The complexity of the disorder-based interactomes is further increased by the capacity of a single IDPR to bind to multiple partners, gaining very different structures in the bound state.[10]

IDPs 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 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 that allow them to exert different functions in different cellular contexts according to a specific conformational state.[5]

Although the field of protein disorder has started from careful analysis of a very limited number of biologically active proteins without unique structures (which, for a long time, were taken as rare exceptions from the general "one sequence−one unique structure−one unique function" paradigm),[1−4] applications of various disorder predictors to different proteomes revealed that IDPs are highly abundant in nature,[11−16] and the overall amount of disorder in proteins increases from bacteria to archaea to eukaryota, with over half of all eukaryotic proteins predicted to contain extended IDPRs.[11,12,15−17] One explanation for this trend is a change in the cellular requirements for certain protein functions, particularly cellular signaling. In support of this hypothesis, an analysis of a eukaryotic signal protein database indicated that the majority of known signal transduction proteins were predicted to contain significant regions of disorder.[18]

A detailed study focused on the intricate mechanisms of IDP regulation inside the cell was recently conducted by Gsponer et al.[19] These authors grouped all the *Saccharomyces cerevisiae* proteins into three classes according to their predicted disorder propensities and evaluated the correlations between intrinsic disorder and the various regulation steps of protein synthesis and degradation.[19] Although the transcriptional rates of mRNAs encoding IDPs and ordered proteins were comparable, IDP-encoding transcripts were generally less abundant than transcripts encoding ordered proteins because of increased decay rates of IDP mRNAs.[19] Also, IDPs were found to be less abundant than ordered proteins because of lower rates of protein synthesis and shorter protein half-lives.[19] Curiously, IDPs were shown to be substrates of twice as many kinases as ordered proteins. Furthermore, the vast majority of kinases whose substrates were IDPs were either regulated in a cell-cycle-dependent manner or activated upon exposure to specific stimuli or stress.[19] Similar regulation trends were also found in proteomes of *Schizosaccharomyces pombe* and *Homo sapiens*,[19] suggesting that both unicellular and multicellular organisms use evolutionarily conserved mechanisms to regulate the availability of their IDPs. This tight regulation is directly related to the major roles of IDPs/IDPRs in signaling, where it is crucial for a given protein to be available in appropriate amounts and not to be present longer than needed.[19] It was also pointed out[5] that although the abundance of many IDPs may be closely regulated, some disordered proteins could be present in cells in large amounts or/and for long periods of time, either due to specific PTMs or via interactions with other factors. These events could promote changes in cellular localization of IDPs or protect them from degradation.[3,20−23] Taken together, these data highlight that the chaos seemingly associated with highly flexible and promiscuous IDPs/IDPRs is under tight control.[24]

2. ENRICHMENT OF INTRINSICALLY DISORDERED PROTEINS IN HUMAN DISEASES: UNCONTROLLED CHAOS OR DISORDER IN DISORDERS CONCEPT

Although IDPs and IDPRs are normally under very tight control, rigorous investigation of IDP functions and dysfunctions conducted over the past decade has led to the recognition that they are prevalent among disease-related proteins, and numerous cases are known in which the malfunction of a protein is associated with the development of particular pathological conditions. In fact, a broad range of human diseases is linked to the failure of a specific peptide or protein to adopt its functional conformational state. This leads to protein misfolding, loss of normal function, gain of toxic function, and/or protein aggregation.[25,26] Each of these diseases originates from the dysfunction of a particular protein. 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 the propensity to misfold. Often, misfolding and dysfunction originate from point mutation(s) or result from protein exposure to internal or external toxins. Furthermore, they 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.[27] The common involvement of IDPs/IDPRs in the pathogenesis of numerous human maladies (disorders) gave rise to the "disorder in disorders" or D^2 concept,[28] according to

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX−XXX

which these proteins are abundantly involved in the development of various diseases because of their unique structural and functional properties. Such diseases, therefore, may originate from the misidentification, misregulation, misfolding, and missignaling of causative IDPs/IDPRs.[25−31]

We were among the first to point to the involvement of IDPs in human diseases.[32] Applying the predictor of protein disordered regions, PONDR VL-XT, to a data set of cancer-associated proteins, we observed a significant enrichment of proteins with IDPRs among these proteins compared to other eukaryotic proteins. Examples of cancer proteins with experimentally confirmed IDPRs include p53,[33] BRCA1,[34] EWS,[35] HPV proteins,[36] and PTEN[37] among others. Recently, a comprehensive computational analysis revealed that a majority of cancer/testis antigens (CTAs), members of an interesting group of heterogeneous proteins that are typically expressed in the testis but aberrantly expressed in several types of cancer, are IDPs.[38] Some of these CTAs can bind DNA and affect cell growth in a dosage-dependent manner, whereas other CTAs serve as hubs in protein regulatory networks.[38]

Following these initial observations, the involvement of IDPs in other human diseases has been intensively investigated. The most notable diseases that involve IDPs are human neurodegenerative diseases. For example, Parkinson's disease, dementia with Lewy bodies, Alzheimer's disease (AD), and Down's syndrome are all characterized by the accumulation of aggregates of the α-synuclein protein that serves as a classical example of an IDP. The disorder of α-synuclein has been experimentally validated by a variety of biochemical and biophysical methods confirming that α-synuclein 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.[39] Other IDPs implicated in neurodegenerative diseases include amyloid β and τ proteins (AD), prions (Creutzfeldt−Jakob disease, scrapie, bovine spongiform encephalopathy), and ataxin (spinocerebellar ataxia).[31]

Besides cancer and neurodegenerative diseases, IDPs have also been implicated in cardiovascular diseases (hirudin and thrombin);[40] type II diabetes (amylin);[28] acquired immunodeficiency syndrome, AIDS [hyman immunodeficiency virus (HIV) Rev protein];[41] and cystic fibrosis (cystic fibrosis transmembrane conductance regulator, CFTR).[42] However, when the entire network of human diseases[43] was analyzed in terms of its disorder content, it was observed that there is a wide variability of predicted disorder among different diseases,[29] possibly due to variability of the disease candidate proteins selected for analysis. Whole-exome and whole-genome sequencing is beginning to provide candidate genes for many complex human diseases. Refinement of the candidate gene lists in the future may open new opportunities for followup analysis of the human unfoldome.

2.1. Computational Approaches for Estimating IDP Abundance in Different Diseases

The intensive involvement of IDPs in pathogenesis of human diseases has been investigated in computational/bioinformatics studies specifically designed to estimate the abundance of IDPs in various pathological conditions. The first computational approach is based on assembly of specific data sets of proteins associated with a given disease and computational analysis of these data sets by use of a number of disorder predictors.[31,32,36,40,44,45] 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.[32] Similar analyses revealed that the percentage of proteins with 30 or more consecutive disordered residues was 61% for proteins associated with cardiovascular disease (CVD).[40] Many CVD-related proteins were predicted to be entirely disordered, with 101 proteins from the CVD data set predicted to have a total of almost 200 specific disorder-based binding motifs (thus about 2 binding sites per protein).[40] Finally, data set analysis revealed that in addition to being abundant in cancer- and CVD-related proteins, intrinsic disorder is commonly found in maladies such as neurodegenerative diseases and diabetes.[25,28]

In a second approach, the abundance of intrinsic disorder was analyzed in the human diseasome,[29] which is a complex network that systematically links the human disease phenome with the human disease genome.[43] 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.[29]

Finally, a third approach is based on evaluation of the association between a particular protein function (including disease-specific functional keywords) and the level of intrinsic disorder in a set of proteins known to carry out this function.[46,47] This analysis revealed that many diseases are strongly correlated with proteins predicted to be disordered.[22,46,47] Contrary to this, no disease-associated proteins were found to be strongly correlated with absence of disorder.[22]

3. REGULATION OF INTRINSICALLY DISORDERED PROTEINS AND DISEASE

Physiological protein function and the ability to be converted from a normal protein to a pathological form depend on multiple factors that can be grouped into two major classes, genetic and nongenetic. Genetic factors include pathological mutations (see section 4), aberrant splicing, chromosomal translocation, alternative transcription, and altered alternative splicing. Nongenetic 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.

3.1. Genetic Factors: Chromosomal Translocation

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. Several forms of cancer, such as acute myelogenous leukemia (AML), acute lymphoblastic leukemia (ALL), chronic myelogenous leukemia (CML), and Ewing's sarcoma (EWS), 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.[48] 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 represent an important factor that helps fusion proteins to escape detection

dx.doi.org/10.1021/cr400713r | Chem. Rev. XXXX, XXX, XXX−XXX

by cellular surveillance mechanisms that eliminate misfolded proteins and to live long enough to manifest their altered function(s).[48] The authors found that these translocation-generated fusions enable 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.[48] Another mechanism is related to fusion of a dimerization/oligomerization domain with the kinase domain, generating an oligomeric hybrid protein. Subunits within such an oligomer are engaged in multiple mutual intermolecular phosphorylation reactions that promote autoactivation and generate novel binding sites for signaling proteins. Examples of this mechanism include TFG−ALK (TRK-fused gene−anaplastic lymphoma kinase fusion) related to anaplastic large-cell lymphomas,[49] constitutively activated TEL−Jak2 fusion (ETS translocation variant 6−Janus tyrosine kinase 2 fusion) with kinase activity in human leukemia,[50] and 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.[51] Finally, chromosomal translocation can affect transcription factors, as illustrated by the EWS-ATF or EWS-Fli1 hybrids, where the DNA-binding element of transcription factors ATF1 or Fli1 is fused to the disordered transactivation domain of the EWS oncogene to generate an aberrant transcription factor related to Ewing sarcoma.[35]

3.2. Genetic Factors: Aberrant RNA Splicing

3.2.1. Intrinsic Disorder and Alternative Splicing. Alternative splicing 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 alternatively spliced mRNA.[52] Recently, it has been established that regions of alternative splicing are enriched in intrinsic disorder.[8] The finding that alternatively spliced regions of mRNA encode IDPRs with greater frequency than structured regions suggests a link between alternative splicing 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.[8] Furthermore, associating alternative splicing with protein disorder enables time- and tissue-specific modulations of protein functions. Since disorder is frequently utilized in protein binding regions, having alternative splicing of pre-mRNA coupled to regions of protein disorder can lead to tissue-specific signaling and regulatory diversity.[8,53] In agreement with this hypothesis, recent bioinformatics analysis clearly showed that tissue-specific splicing of disordered segments with embedded binding motifs is responsible for rewiring of protein interaction networks and signaling pathways.[54,55]

3.2.2. Altered Alternative Splicing and Diseases. Although the flexibility of alternative splicing constitutes an evolutionary advantage for higher eukaryotes, it also represents a risk. In fact, strong evidence indicates that defective regulation of alternative splicing correlates with onset and progression of human cancers.[56−59] Alterations in alternative splicing might generate multiple mRNA variants from a single oncogene, thus producing protein isoforms with different or even opposing functions and thereby contributing to the heterogeneity of various cancers, such as prostate cancer[60] or AML.[61] The phenomenon of cancer-associated (or cancer-promoting) aberrant splicing is widespread. For example, ∼29% of genome-wide expressed genes were shown to be differentially and recurrently spliced in AML patients compared to healthy individuals.[61] Among these differentially spliced genes were genes encoding several oncogenes, tumor suppressor proteins, splicing factors, heterogeneous nuclear ribonucleoproteins, and proteins involved in apoptosis, cell proliferation, and spliceosome assembly.[61] Among these targets of aberrant splicing in AML, there are many proteins with known enrichment in intrinsic disorder, such as proteins related to apoptosis[62] as well as proteins involved in spliceosome assembly.[63,64] Some of the crucial proteins affected by alternative splicing in prostate cancer (e.g., androgen receptor, zinc finger transcription factor Krüppel-like factor 6, Bcl-x, and cyclin D1)[60] are known to contain IDPRs.[62,65]

Finally, a case of extensive alternative splicing of the TMPRSS2−ERG gene fusion represents an important illustration of the combined effects of chromosomal translocations and alternative splicing.[60,66] 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 gene 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 alternative splicing 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.[66]

Numerous studies confirmed the existence of specific differences in alternative splicing profiles between normal and cancer tissues.[67] Cancers are not the only set of diseases affected/promoted/caused by altered alternative splicing. For example, in an autosomal dominant neurodegenerative disease called frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17), patients possess a 2-fold increase in the 4R:3R ratio of τ isoforms [i.e., isoforms containing four (4R) or three (3R) microtubule binding domains, respectively] leading to enhanced aggregation causing the disease.[68] In spinal muscular atrophy (SMA), constitutive alternative splicing of the survival of motor neuron gene (*SMN*) generates the SMNΔ7 isoform lacking the region encoded by exon 7. SMNΔ7 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.[68]

3.3. Nongenetic Factors Generating Protein Pathogenicity

3.3.1. Altered Expression of IDPs and Disease. As was already pointed out, cells have evolved multiple complex mechanisms during transcription and translation to regulate the availability of IDPs.[19] Since IDPs are important players in

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX−XXX

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, nonfunctional interactions and inappropriate sequestering of proteins into nonphysiological protein complexes.[19] In agreement with this hypothesis, a careful analysis of dosage-sensitive genes (i.e., genes that are harmful when overexpressed) 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.[69,70]

3.3.2. Abnormal Posttranslational Modifications. *3.3.2.1. 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.[71] For example, abnormal glycosylation of some glycoproteins due to deregulated glycosyltransferases and glycosidases is known to be a common phenomenon of many malignancies, including colorectal cancer (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.[72] The widespread and diverse PTMs of histones, important nuclear IDPs[73] that are crucial for regulated gene expression and for a variety of epigenetic mechanisms, are under very tight and complex spatial and temporal control.[74] This spatial and temporal regulation of histone modifications is distorted in malignancies on both genome-wide and discrete gene loci levels.[74] For example, excessive aberrant acetylation and methylation of specific histone residues have been found in CRC.[75] Also, alterations of different PTMs at lysine residues (such as acetylation, methylation, ubiquitination, and sumoylation) of proteins involved in DNA repair are often associated with genomic instability, which is the major cause of different diseases, especially cancer.[76]

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 androgen receptor (AR) found in prostate cancer is caused by alterations of enzymes that modify the AR.[77] 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.[78]

3.3.2.2. Aberrant PTMs and Neurodegenerative Diseases. In Huntington's disease, 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.[79] Various PTMs of Htt itself, such as phosphorylation, sumoylation, ubiquitination, acetylation, proteolytic cleavage, and palmitylation, are also significantly altered in Huntington's disease, resulting in changes in clinical phenotypes.[80] In Alzheimer's disease (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 τ, Aβ levels and τ aggregation are impacted by altered sumoylation.[81] Aberrant phosphorylation of the microtubule-associated protein τ is known to be associated with AD pathology and pathogenesis of other neurodegenerative disorders called tauopathies.[82] In fact, in AD, τ is abnormally hyperphosphorylated to a stoichiometry of at least 3-fold greater than normal τ. This hyperphosphorylation is believed to be a major driving force for pathological τ 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 τ is also a characteristic feature of several other tauopathies.[83]

4. PATHOLOGICAL MUTATIONS OF INTRINSICALLY DISORDERED PROTEINS

4.1. Disease Mutations in Ordered Regions

How does a protein become a "disease protein"? In the majority of cases, a disease protein is annotated as such because a genetic mutation(s) identified in a patient(s) with a particular condition alters the encoded protein by either replacing its amino acid residue (i.e., missense mutation) or producing a truncated protein (i.e., nonsense mutation) or an unnaturally extended protein (i.e., frame-shift insertion/deletion mutation). Disease mutations can either be inherited or arise de novo. In order for a rare mutation to be considered as causative, it has to be observed in several patients but not in healthy control individuals. However, some mutations are not fully penetrant and could be observed in both the patients and the controls. In addition, some disease mutations are so rare that identifying them in several individuals requires sequencing of large cohorts. Determining the causality of the mutation is a challenging problem, especially for complex diseases that are often caused by rare and not fully penetrant mutations.

Historically, the functional impact of disease-associated mutations was analyzed from a structural perspective. Over the years, researchers have tried to address two important questions: (1) how a disease mutation influences protein structure and function and (2) how to distinguish a disease-causing mutation from a benign mutation or a neutral polymorphism. Next-generation sequencing technologies are producing an ever-increasing number of new mutations. As a result, the number of mutations implicated in diseases far exceeds the amount of available resources to experimentally test their functional impact, and reliance on computational methodologies is therefore unavoidable.

Various experimental and computational studies have repeatedly demonstrated that disease mutations can influence protein stability, activity, oligomerization, folding, cellular localization, and other structure-based properties. An excellent review describing numerous examples of the impact of mutations on the above properties has recently been published.[84] Significant progress has also been achieved in predicting the structural and functional impact of mutations. Many algorithms that are typically based on the information from solved or modeled protein structures combined with data on evolutionary conservation have been developed to predict the functional effect of mutations and to distinguish between

dx.doi.org/10.1021/cr400713r | Chem. Rev. XXXX, XXX, XXX−XXX

damaging and benign mutations.[85−89] These methods vary in accuracy and, regrettably, the results of their predictions with regard to pathogenicity correlate poorly with each other.[90] This raises the question of whether additional protein properties could be used for training the predictors to increase their accuracy. Only recently, new approaches were developed that, in addition to structural and evolutionary data, incorporate functional protein features such as prediction of PTMs and catalytic residues among others.[91]

4.2. Disease Mutations in Disordered Regions

The existence of IDPs and their prevalence in eukaryotic organisms, initially suggested more than a decade ago,[2−4,18] is now widely recognized. However, the general question of how disease mutations occurring in the IDPRs impact protein function remains largely unexplored. In fact, up until recently it was not known whether disease mutations could even occur in IDPRs.[92] We have investigated the prevalence of mutations in IDPRs by mapping all disease mutations and polymorphisms from the UniProt database to predicted ordered (OR) and disordered protein regions.[93] In agreement with previous observations that disease mutations affect protein structure, we observed significant enrichment of disease mutations in ORs, which was not due to overall lower disorder content of the proteins containing these mutations (Figure 1). Despite the enrichment of disease mutations in ORs, about 20% of them (corresponding to over 3300 mutations in total) mapped to the IDPRs. Our analysis of the global effect of IDPR disease mutations on PTM sites showed that some mutations may cause loss or gain of PTMs.[94,95] We also observed that disease-associated mutations lead to alterations in PTM sites more frequently than polymorphisms.[94] For example, loss of an ubiquitination site due to mutation of a lysine residue could lead to stabilization of the mutant protein and to its abnormal cellular accumulation. If such a mutation occurs in the oncoprotein, this may lead to cancer, as has been previously observed.[96] On the other hand, loss or gain of phosphorylation sites as a result of mutations could lead to hypo- or hyperphosphorylation, which may again lead to various diseases. Dysregulation of phosphorylation has been previously implicated in many human diseases.[97,98]

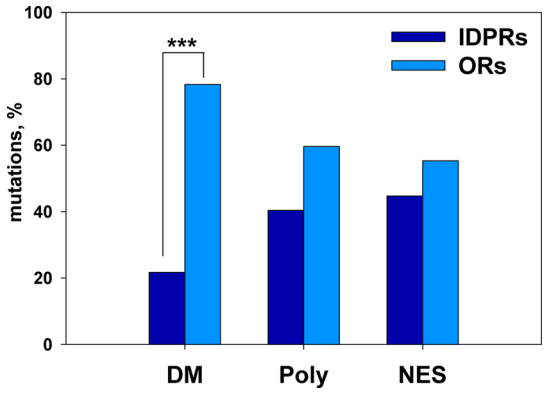


Figure 1. Distribution of disease mutations (DM) between predicted ordered regions (OR) and intrinsically disordered protein regions (IDPR) of proteins. Analysis of annotated DMs from the UniProt database shows that they are enriched in ordered regions of proteins compared to polymorphisms (Poly) or neutral evolutionary substitutions (NES). Data from ref 93 were used to create this figure.

Another important function of IDPRs is interaction with proteins, nucleic acids, and other ligands. Analysis of protein−protein interaction (PPI) networks demonstrated that IDPs frequently serve as network hubs (i.e., proteins that interact with many partners). Disorder may provide the flexibility and malleability needed to conform to differently shaped interfaces of a large number of binding partners. When disordered regions bind to multiple partners, they often undergo disorder-to-order transition via so-called molecular recognition features (MoRFs),[99] and this phenomenon was observed for both homodimeric and heterodimeric PPIs.[100] It was also proposed that disorder may increase interaction surface areas, thereby facilitating low-affinity/high-specificity binding.[101] The role of disorder in promiscuous interactions within PPI networks has recently been discussed.[102] As such, disruption of disorder by disease mutations may impair interactions with corresponding partners. PPI network disruptions should be a frequent cause of human diseases. Examples of network disruption by mutations associated with some human Mendelian disorders have recently been demonstrated.[103,104]

4.3. Disorder-to-Order and Order-to-Disorder Transition Mutations

As mentioned earlier, about 20% of annotated human disease mutations map to the IDPRs.[93,105] However, the molecular mechanisms by which these mutations impact IDPR functions remain unexplored. Since IDPRs often undergo disorder-to-order (D → O) transitions when interacting with their partners, we have hypothesized that disease mutations may also manifest their functional impact through D → O transitions. This hypothesis was tested by in silico mutating proteins carrying disease mutations and then comparing the predicted disorder scores of wild-type and mutant proteins.[93] Surprisingly, it has been observed that disease mutations lead to predicted D → O transitions more frequently than polymorphisms not associated with diseases or neutral evolutionary substitutions (Figure 2). This suggests that transitions of disordered regions into folded states may play important roles in the diseases. Further investigations of potential functional consequences of D → O mutations demonstrated that MoRFs are frequently disrupted.

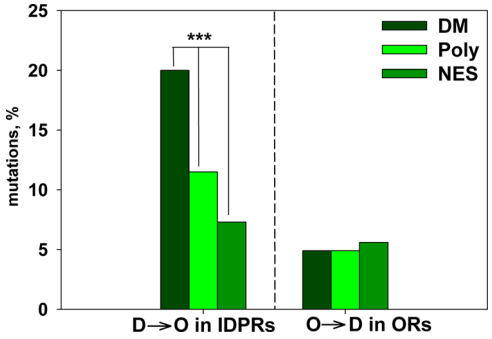


Figure 2. Conformational effects of disease mutations. Shown are the percentages of predicted order transitions for different classes of mutation in intrinsically disordered regions (IDPRs, left) and ordered regions (ORs, right) of proteins. A significantly greater percentage of disease mutations (DM) leads to a disorder-to-order (D → O) transition compared to polymorphisms (Poly) or neutral evolutionary substitutions (NES). For mutations in ordered protein regions, there is no significant difference in the percentage of order-to-disorder (O → D) transitions for the three mutant classes. Data from ref 93 were used to create this figure.

dx.doi.org/10.1021/cr400713r | Chem. Rev. XXXX, XXX, XXX−XXX

Because MoRFs mediate protein−protein interactions, it follows that PPIs may also be disrupted. Indeed, some examples from the published literature confirm both the existence of D → O mutations and their impact on protein−protein interactions. We will discuss such examples below.

Mutations of the epidermal growth factor receptor (EGFR) are associated with many types of lung cancer.[106] A recent study has demonstrated that the dimerization interface of wild-type EGFR is intrinsically disordered, as indicated by long-time-scale molecular dynamics simulations and also by H/D exchange measurements.[107] Upon receptor dimerization, the interface undergoes a D → O transition, which leads to receptor activation. Interestingly, some oncogenic mutations reduce local disorder of the EGFR interface and facilitate EGFR dimerization and its abnormal activation. For example, the L834R mutation lowers the threshold of EGFR activation, making the mutant protein more active than the wild-type variant. Furthermore, the authors demonstrated that the higher activity of the L834R cancer mutant may arise primarily from higher dimerization affinity rather than from the modest increase in its intrinsic catalytic potency. Other oncogenic mutations in EGFR also reduce its disorder, either by rigidifying the disordered β3−αC loop or shortening it (deletion mutation Del722−726) or, alternatively, by reducing the flexibility of the disordered P loop, which is dynamically coupled with the β3−αC loop (G695S mutation). In addition, the simulations demonstrate that Tyr845 phosphorylation at the activation loop suppresses intrinsic disorder and secures the αC-in conformation, suggesting a molecular mechanism for autonomous EGFR signaling.[107] These examples are in agreement with our hypothesis about the potential role of D → O mutations in human diseases.

Additional examples of D → O mutations implicated in disease are mutations in the tumor suppressor protein APC.[108] Mutations in the adenomatous polyposis coli (APC) tumor suppressor strongly predispose to development of gastrointestinal tumors. Remarkably, the large C-terminal region of APC, which spans over 2000 amino acids and includes critical regions in downregulating β-catenin, is predicted to be natively unfolded. Recently, a significant number of germline and somatic missense mutations in the central region of APC were linked to tumorigenesis in the colon as well as extraintestinal tissues. The molecular basis by which these mutations interfere with APC function remains unresolved. The authors mapped all known mutations to the APC structure and its IDPRs and proposed several mechanisms by which cancer-related missense mutations in the large disordered domain of APC may interfere with tumor suppressor activity. Among these mechanisms are alterations of protein interaction surfaces, changes in secondary structure, disruption of PTMs, and shifts of dynamics in conformational equilibria.[108] Our predictions of the impact of APC mutations suggest that at least two may act via transition mechanisms: R1348W causes a D → O transition, whereas F1197S causes an order-to-disorder (O → D) transition. Furthermore, the latter is also predicted to introduce a potential serine phosphorylation site, as indicated by our previously developed phosphorylation site predictor DisPhos (http://www.dabi.temple.edu/disphos/).[20] Although these mechanisms would have to be experimentally validated, the predictions serve as the basis for generating testable hypotheses regarding the impact of disease mutations in various proteins.

The two examples above demonstrate how cancer-associated mutations may affect IDPRs. Mutations involved in other diseases could have a similar structural impact. One example is a mutation implicated in Rett (RTT) syndrome (MIM: 312750). RTT is a neurodevelopmental disorder that occurs almost exclusively in females. It is characterized by arrested development between 6 and 18 months of age, regression of acquired skills, loss of speech, stereotypical movements (classically of the hands), microcephaly, seizures, and mental retardation. Mutations in the methyl-CpG-binding protein 2 (MECP2) are implicated in the majority of RTT cases.[109] Three D → O MECP2 mutations have been discussed previously.[92] An additional RTT mutation, R106W, although formally not crossing the disorder-to-order transition threshold at the mutation site, is of interest because it significantly decreases the disorder score of the 30-residue-long MECP2 region (Figure 3). A recent H/D exchange study on full-length MECP2 indicated that essentially the entire MECP2 polypeptide chain underwent H/D exchange at rates faster than could be measured.[110] Even its methyl DNA binding domain (MBD) exchanged rapidly, suggesting high conformational flexibility.

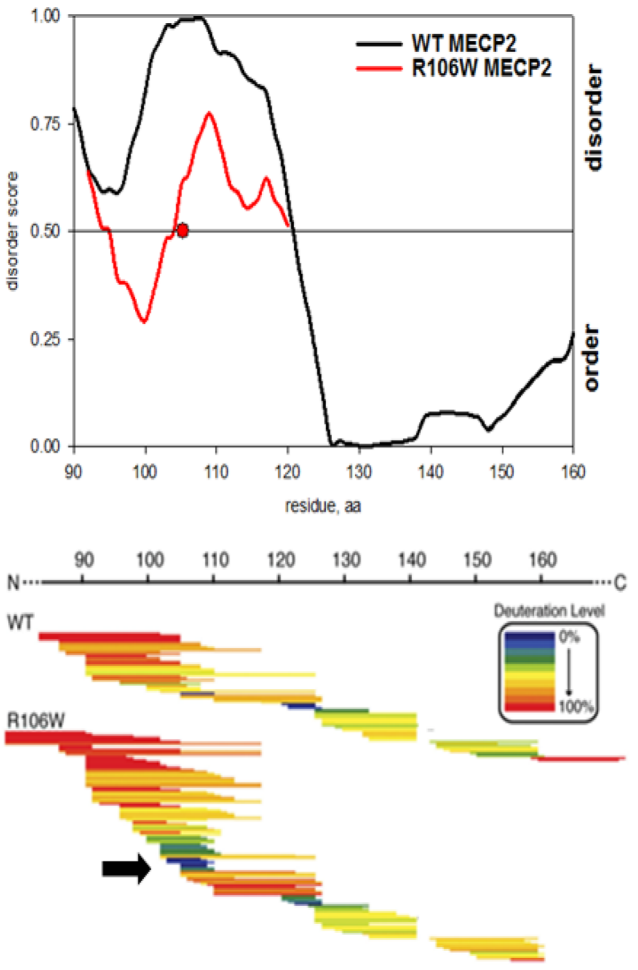


Figure 3. Effect of R106W mutation on the conformational state of MECP2. (Top) PONDR VL-XT disorder predictions for the residue 90−160 region of the wild type (WT) and R106W mutant of MECP2. The mutation (red circle) decreases the disorder score of the 95−125 MECP2 region. (Bottom) H/D exchange profiles of the WT and R106W mutant of MECP2. The black arrow points to a more slowly exchanging region at the mutation site. Adapted with permission from ref 110. Copyright 2011 American Society for Biochemistry and Molecular Biology.

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX−XXX

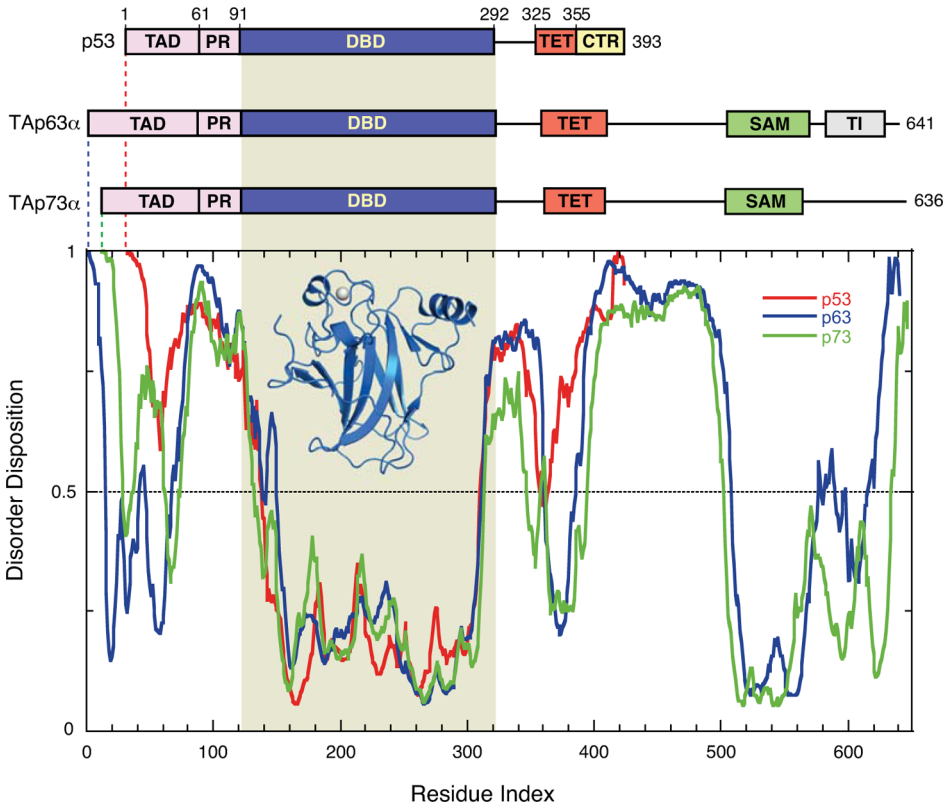


Figure 4. Schematic domain organization of p53 family members and disorder prediction. p53 consists of a natively unfolded N-terminal transactivation domain (TAD), a proline-rich region (PR), a DNA-binding domain (DBD), a flexible linker region, a tetramerization domain (TET), and the intrinsically disordered C-terminal regulatory domain (CTR). The paralogues p63 and p73 have a similar domain organization but feature an extended C-terminal region, including a structured sterile α motif (SAM) domain and in the case of p63 a transactivation inhibitory region (TI). The plot shows the disorder prediction for the three proteins by PONDR-FIT.[374] The disorder prediction curves for the three proteins are aligned on the basis of the position of DBDs, which show the highest sequence conservation. A value above 0.5 indicates structural disorder. The structure of the human p53 DBD (PDB entry 2XWR)[375] is shown as a blue ribbon diagram, and the bound zinc ion is highlighted as a gray sphere.

Binding of unmethylated DNA slowed down MBD H/D exchange by several orders of magnitude. Interestingly, R106W also led to a localized decrease of H/D exchange (Figure 3), suggesting a potential D → O transition. There are other examples of experimentally confirmed D → O and O → D transition mutations, with some of them being involved in disease.[111,112]

4.4. Whole-Exome Sequencing and Mutation Mapping to Structured and Disordered Regions

Extended studies on various human diseases using whole-exome and whole-genome sequencing technologies are currently ongoing. Millions of mutations are being identified in patients and healthy control individuals. It would not be feasible to test the functional impact of all discovered mutations experimentally. Thus, computational tools are needed to make accurate predictions and distinguish between disease-causing and benign mutations. Many such tools have already been developed, but the availability of larger mutational data sets opens the doors to further improve their accuracy. For example, as was recently shown, two of the tools, SIFT and PolyPhen, have lower accuracy when predicting the effect of mutations in disordered regions.[93] A particularly important question in this field is not only how mutations influence the protein itself but rather how they influence its interaction network. Networks influenced by mutations include physical interactions between two proteins, interactions of mutant proteins with DNA or RNA, regulatory interactions, and interactions between different splice variants. Developing predictive methods to address these questions is essential for a better understanding of human diseases.

5. ILLUSTRATIVE EXAMPLES OF PATHOGENIC INTRINISCALLY DISORDERED PROTEINS

The p53 and PTEN proteins are excellent paradigms for understanding the important role of intrinsic disorder in signaling pathways. They illustrate many key features of how intrinsic disorder and conformational flexibility facilitate binding promiscuity of proteins at the center of complex interaction networks, which enables the cell to process multiple signals from different pathways and quickly respond to genotoxic stress. In the following, we will discuss in detail the various functional and structural roles of IDPRs in these proteins and their evolution and regulation, as well as disease-related perturbations of their interaction networks and possible therapeutic strategies.

5.1. p53 Family of Transcription Factors and the Functional Role of Intrinsic Disorder in Cancer Pathways

The p53 protein is a tetrameric transcription factor that plays a key role in cell cycle control. Despite more than 70 000 publications on p53 (PubMed as of March 2014), many aspects of its structure and function remain poorly understood. p53 is best known for its role as a tumor suppressor and guardian of the genome,[113] but besides induction of cell-cycle arrest, apoptosis, or DNA repair, it is also involved in many other

dx.doi.org/10.1021/cr400713r | Chem. Rev. XXXX, XXX, XXX–XXX

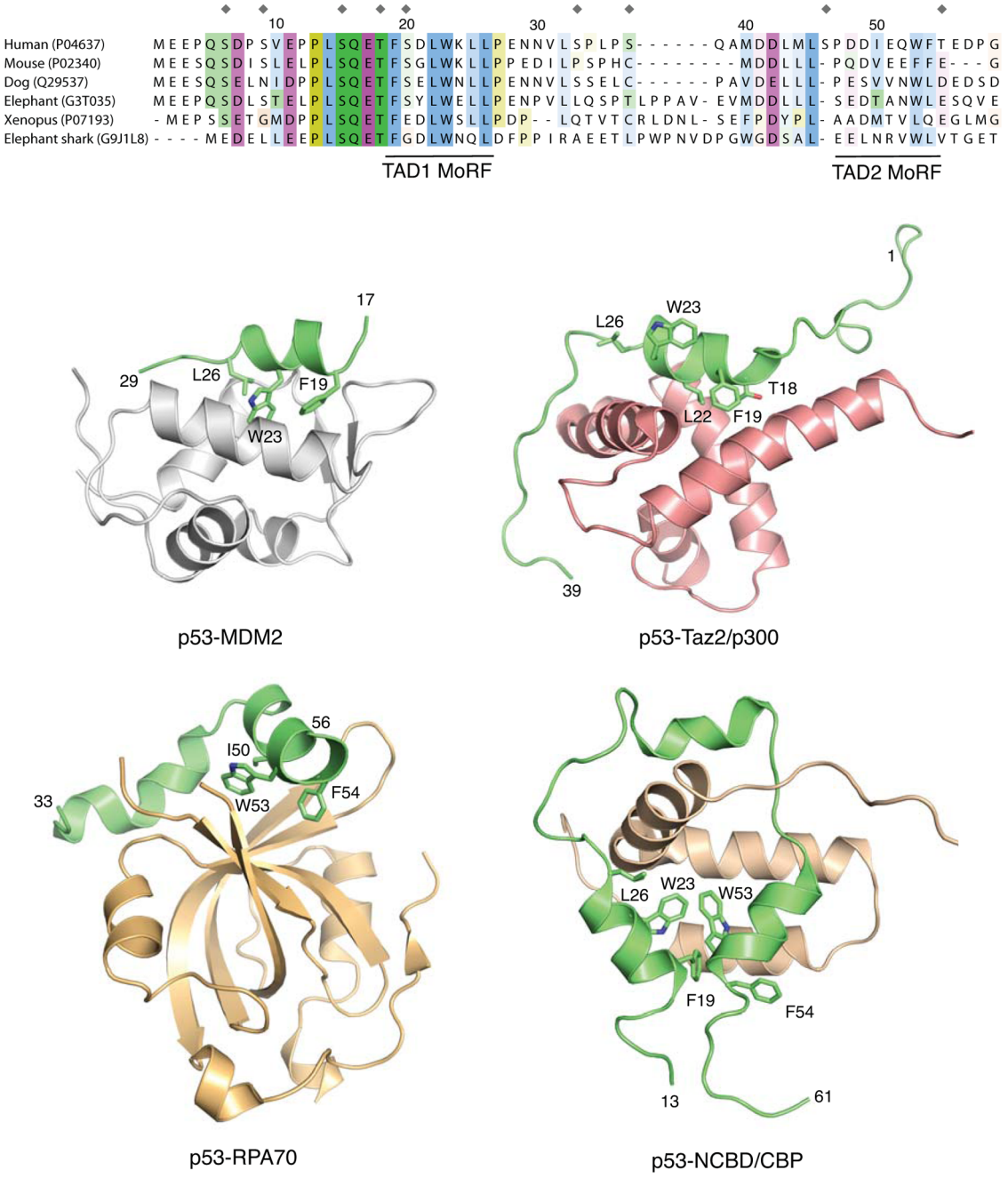


Figure 5. Molecular recognition features in p53 transactivation domain (TAD). Sequence alignment of p53 TAD from different vertebrate species shows the location of MoRFs in TAD1 and TAD2. UniProt accession numbers are given in parentheses. Gray diamond shapes indicate phosphorylation sites in the human protein.[140] Key interacting hydrophobic residues in TAD1 are highly conserved. The sequence alignment was generated by use of MUSCLE[376] and JALVIEW.[377] The two MoRFs undergo disorder-to-order transition and form an amphipathic α-helix upon binding to various target proteins in the cell cycle. Shown are complexes with MDM2 (PDB entry 1YCR),[145] Taz2 domain of p300 (PDB entry 2K8F),[147] RPA70 subunit of replication protein A (PDB entry 2B3G),[149] and nuclear coactivator binding domain (NCBD) of CBP (PDB entry 2L14).[154] Structures of the complexes are shown as cartoon representations. The p53 segment in each structure is highlighted in green, and key contact residues are shown as stick models. Structural representations in this and following figures were generated by use of PyMOL (www.pymol.org).

cellular processes, including senescence and differentiation.[114] Recent studies have also shown that p53 regulates metabolic pathways, enabling cells to survive metabolic stress.[115,116] p53 has two structurally similar homologues in vertebrates, p63 and p73, which have overlapping and distinct functions in cell cycle regulation. p63 and p73 play important roles in the control of normal development and in maintaining the fidelity of the female germline but also participate in tumorigenesis.[117] To complicate matters further, p53 family genes are expressed as multiple isoforms (with sometimes antagonistic functions) due to alternative splicing, alternative promoter usage, and alternative initiation of translation.[118]

5.1.1. Modular Domain Organization of p53 Family Proteins. Given their multifaceted functions, it is not surprising that p53 family proteins have a complex domain organization.[117,119] Human p53 protein is a homotetramer of 4

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX–XXX

× 393 residues with independently folded domains that are linked and flanked by extended disordered regions (Figure 4). About 50% of the protein is unfolded under native conditions.[120] It has an intrinsically disordered N-terminal transactivation domain (TAD) that can be further subdivided into two subdomains, TAD1 (residues 1−40) and TAD2 (residues 41−61), followed by a proline-rich region (PR). The central DNA-binding domain (DBD; residues 91−292) is structured and adopts an immunoglobulin-like β-sandwich fold. Two large loops (stabilized via coordination of a zinc ion) and a loop−sheet−helix motif form the binding surface for sequence-specific interaction with p53 target DNA sequences.[121] Four DBDs cooperatively bind to the p53 response elements, which consist of two palindromic half-sites, with significant DBD−DBD contacts both between and within half-sites.[122,123] A flexible linker region connects the DBD and the tetramerization domain. Like the N-terminal region, the C-terminal region of p53 is intrinsically disordered and features numerous PTM sites.[120,124,125]

p63 and p73 proteins have similar domain organizations (Figure 4), but their C-terminal regions are more than 200 residues longer and display additional structural features.[126] These regions of p63 and p73 contain a structured sterile α motif (SAM) domain, which functions as a putative protein interaction module.[127,128] In addition, p63 also contains a unique motif at its C-terminus that executes autoinhibitory effects on the transcriptional activity of the TAp63α isoform.[126,129]

While the individual domains are well-characterized structurally, relatively little information is available about the structures of the full-length proteins and their complexes. This is mainly due to intrinsic flexibility, which impairs crystallization. A combination of NMR to map domain−domain interactions, small-angle X-ray scattering, and electron microscopy revealed that the full-length p53 tetramer adopts an open, cross-shaped conformation with loosely coupled DBD dimers in its unbound state but forms a much more compact quaternary structure upon binding to target DNA.[130−132] This highlights the important role of the flexible linker region between the folded domains in facilitating domain−domain interactions and domain rearrangements upon binding to different interaction partners.

The intrinsically disordered PR plays an important structural role. Analysis of residual dipolar coupling (RDC) data showed that it formed polyproline II-type helical structures.[33] In addition to serving as a potential site for protein−protein interactions, it also plays a structural role by providing a relatively stiff linker region that projects the TAD away from the central DBD−DNA complex so that it can interact more efficiently with transcriptional coactivators upon binding to the promoter regions of p53 target genes.[33] Interestingly, the PR contains the most common p53 polymorphism, codon 72 Arg/Pro, which has been associated with different cancer risks.[133] Mechanistically, the role of this polymorphism in cancer predisposition is poorly understood, but there are indications that it might affect p53's protein interaction network. iASSP, an evolutionarily conserved p53 inhibitor, binds to the PR of p53-Arg72 more efficiently than to that of p53-Pro72, thus potentially modulating the apoptotic function of the two polymorphic variants.[134]

5.1.1.1. p53 Transactivation Domain: Coupling of Phosphorylation, Binding, and Folding. In unstressed cells, p53 levels are kept constitutively low through a negative feedback loop with the E3 ubiquitin ligase MDM2, a transcriptional target of p53. MDM2 promotes ubiquitination of p53, in concert with its structural homologue MDMX (also known as MDM4), leading to p53 degradation by the proteasome.[135] Unlike MDM2, MDMX has no E3 ligase activity[136] but forms heterodimers with MDM2 (via its RING domain), which have increased p53 ubiqitination activity compared to MDM2 homodimers.[137,138]

Upon DNA damage, p53 is stabilized and activated through a phosphorylation and acetylation cascade, resulting in DNA repair or apoptosis.[114,115,124,139,140] This process involves temporary suspension of the MDM2 feedback loop and recruitment of transcriptional coactivators, such as CBP and its close homologue p300. The intrinsically disordered TAD plays a key role in this activation process. It contains two MoRFs with nascent helical structure that adopt stable secondary structures upon binding to regulatory proteins.[33] This structural and conformational plasticity allows promiscuous binding to a myriad of signaling and accessory proteins that regulate p53 function in the cell cycle. Folding is context-dependent and driven by hydrophobic interactions with target proteins. p53 TAD binds to its negative regulators MDM2 and MDMX, for example,[141] but it also interacts with several domains of the transcriptional coactivators CBP/p300, thus connecting p53 to the basal transcriptional machinery.[142,143]

Both MoRFs contain conserved hydrophobic residues that are flanked by charged residues (Figure 5). Mutation of hydrophobic residues within these two sequence motifs, L22Q/W23S and W53Q/F54S, are associated with transactivation-deficient phenotypes.[144] Residues within TAD1 form an amphipathic helix upon binding to the N-terminal domain of MDM2[145] and MDMX[146] or the Taz2 domain of p300.[147] The TAD1 helix binds to a deep hydrophobic cleft on the N-terminal domain of MDM2 and MDMX via a triad of highly conserved hydrophobic residues (Phe19, Trp23, and Leu26). The TAD1 binding site on Taz2 is relatively shallow in comparison, and only two of the conserved hydrophobic residues of TAD1 (Phe19 and Leu22) contact Taz2, whereas Trp23 and Leu26 are solvent-exposed.[147]

Several structures of TAD2 complexes have been solved. A nine-residue segment within TAD2 (residues 47−55) forms an amphipathic α-helix upon binding to the pleckstrin homology domain of the Tfb1 subunit of transcription factor II human (TFIIH).[148] Again, key contacts are made by a set of hydrophobic residues (Ile50, Trp53, and Phe54). A peptide comprising p53 residues 37−57 forms two amphipathic helices in complex with RPA70, with the second helix mimicking binding of single-stranded DNA.[149,150] TAD2 also mediates interactions with transcription factor PC4, mitochondrial single-stranded DNA-binding protein, and the oligonucleotide/oligosaccharide-binding domains of BRCA2 by acting as a DNA mimetic.[151−153]

While MDM2 binds strongly to TAD1, other p53 interactors bind synergistically to both TAD subdomains to achieve tight binding.[142] Studies on the KIX domain of CBP, for example, have shown that isolated p53 TAD subdomains bind only weakly, whereas peptides containing both subdomains bind KIX tightly by simultaneously interacting with two binding surfaces.[143] Both TAD1 and TAD2 subdomains interact with the nuclear coactivator binding domain (NCBD) of CBP.[154] This example is of particular interest because both interaction partners are intrinsically disordered and fold synergistically upon binding, resulting in the formation of an intermolecular

dx.doi.org/10.1021/cr400713r | Chem. Rev. XXXX, XXX, XXX−XXX

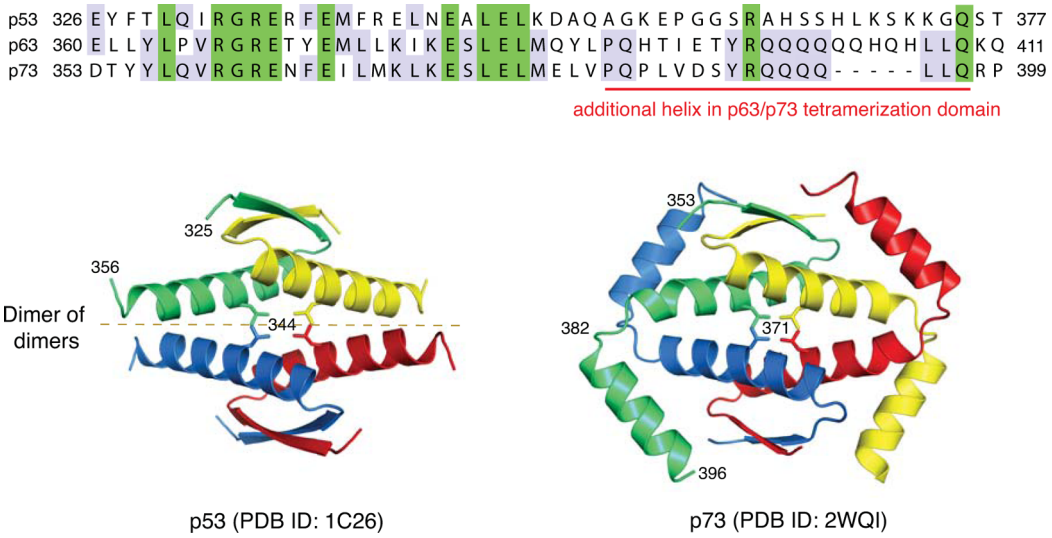


Figure 6. Structures of p53 family tetramerization domains. The sequence alignment shows the oligomerization domain regions of human p53, p63, and p73. Identical residues in all three sequences are highlighted in green, while residues that are identical in only two members are highlighted in light blue. p63 and p73 form an extended oligomerization domain and require a second helix for formation of stable tetramers. Crystal structures of p53 (PDB entry 1C26) and p73 (PDB entry 2WQI) tetramers show the assembly of tetramers as dimers of dimers with highly intertwined monomers (shown in different colors).[164,168]

hydrophobic core.[154] In this complex, p53 TAD wraps around NCBD, with residues 19−25 and 47−53 forming an amphipathic α-helix as observed in the complexes of isolated TAD1 and TAD2 subdomains with MDM2 and replication protein A, respectively, while the linker region retains a certain degree of conformational freedom (Figure 5). In all the structural examples of TAD binding mentioned above, hydrophobic interactions are the driving force of coupled folding and binding reactions, with polar interactions providing additional specificity.

While intrinsic disorder provides conformational flexibility and adaptability to bind to diverse binding partners with high specificity, an additional layer of regulation is necessary to achieve selectivity in complex interaction networks where there are many competing binding partners. Serine and to a lesser extent threonine phosphorylation sites tend to be associated with intrinsically disordered regions.[155] They are often directly involved in protein−protein interactions where they can provide a fast and reversible way of switching selectivity or favoring a particular binding partner amidst a pool of competing interactors. p53 TAD contains nine phosporylation sites (see sequence alignment in Figure 5), and their modifications play pivotal roles in the regulation of p53 activity by dynamically modulating the affinity of its protein−protein interactions in response to different environmental cues.[140] Phosphorylation of Thr18 in response to DNA damage, for example, significantly reduces its affinity for MDM2 because of electrostatic repulsions between the phosphate group and an anionic patch of the p53 binding site.[156−159] Binding affinity is further reduced upon TAD hyperphosphorylation.[158] While phosphorylation prevents MDM2 binding, thus stabilizing the p53 protein, binding to p300/CBP is significantly enhanced through phosphorylation cascades.[154,158,160] Phosphorylation of Thr18 increases p53 TAD affinity for the Taz1 domain of p300 7-fold.[158] A similar phosphorylation effect was also observed for p73.[161]

The effects of individual phosphorylation sites in p53 TAD are additive, and heptaphosphorylation increases its affinity for Taz1 by almost 2 orders of magnitude and for the CH3 domain of p300 about 40-fold.[158] Similarly, successive phosphorylation events in p53 TAD increase affinity for different domains of CBP in an additive manner.[160] Moreover, phosphorylation of p53 at Ser46 and Thr55 significantly enhances binding to Tfb1 and p62 subunits of TFIIH, again in an additive manner, most likely through the formation of phosphate-mediated salt bridges.[148] Modulation of binding affinities by dynamic phosphorylation events thus enables p53 to rapidly respond to genotoxic stress and recruit p300/CBP and TFIIH for transcriptional activation. Increased recruitment of p300/CBP and other coactivators through successive phosphorylation events may further enable fine-tuning of p53 response pathways, depending on the extent and severity of genotoxic stress.[160]

5.1.1.2. p53 Family Tetramerization Domains: Masked Molecular Recognition Features. Tetramerization is crucial for the function of p53 family members. The p53 tetramerization domain (residues 325−356 in human) forms dimers of dimers with D_2 symmetry.[162−164] The individual subunits consist of a short β-strand followed by an α-helix. A conserved glycine (Gly334) between the β-strand and the α-helix facilitates formation of a sharp turn that is stabilized through intersubunit contacts. Two monomers form a dimer via an intermolecular β-sheet and antiparallel packing of the α-helices. Two such dimers then interact in a roughly orthogonal fashion via a hydrophobic interface, forming a central four-helix bundle.[162−164] Leu344 is a key residue of this hydrophobic interface, and Leu344 side chains from all four subunits contact each other (Figure 6). Leu344 is part of a leucine-rich nuclear export signal that is occluded upon tetramerization, thus linking oligomerization and subcellular localization.[165] p53 biogenesis is thought to involve cotranslational formation of dimers on polysomes, with tetramers being formed posttranslationally.[166] These data are also consistent with p53 oligomerization kinetics and equilibria in vitro.[167]

Despite high sequence conservation of p53 family tetramerization domains, p63 and p73 require a second helix to form

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX−XXX

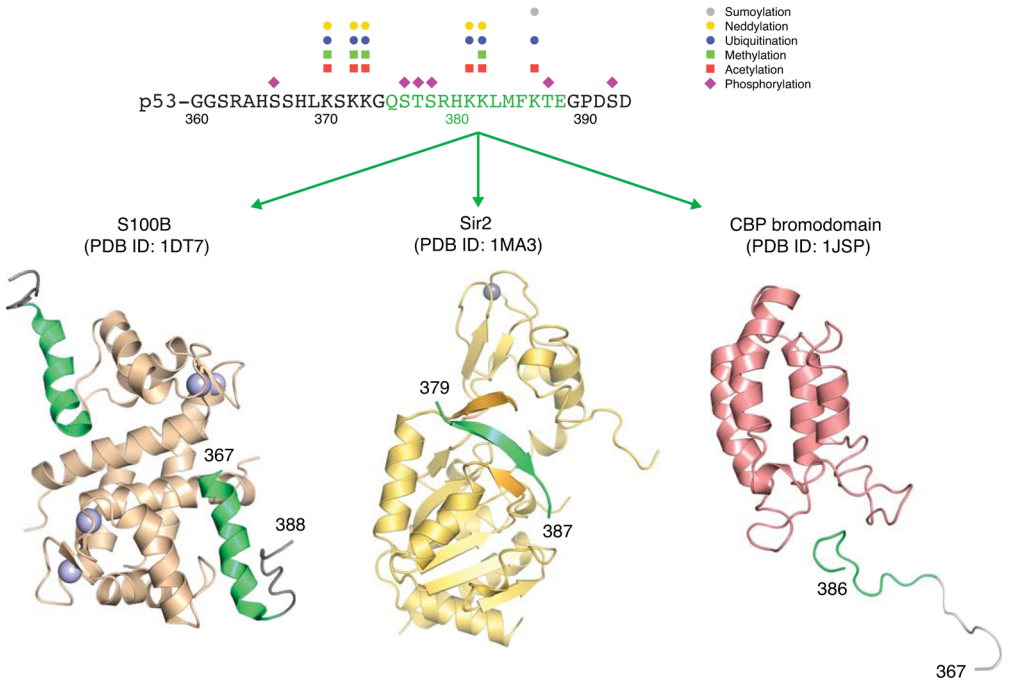


Figure 7. Chameleon sequence in the C-terminal regulatory domain of p53. A short segment in the C-terminal domain of p53 (green sequence motif) adopts different secondary structures, depending on its interaction partner, as shown for the structures in complex with calcium-loaded S110B (PDB entry 1DT7),[185] deacetylase Sir2 (PDB entry 1MA3),[187] and bromodomain of transcriptional coactivator CBP (PDB entry 1JSP).[188] Data from refs 124 and 125 were used to show sites of PTMs.

stable, transcriptionally active tetramers.[168−170] The overall building principle of the tetramers is the same as for p53, but the second helix reaches across to a neighboring dimer within the tetramer and creates additional hydrophobic and polar subunit contacts (Figure 6). The resulting tetramers are composed of highly intertwined monomers with a Z-shaped conformation that, like their p53 counterparts, form almost no intramolecular contacts and most likely exist in a conformational equilibrium in solution. The p53 family tetramerization domains can therefore be viewed as a special case of MoRFs that become frozen in a defined conformation through self-assembly. This view is also supported by studies of p53 tetramerization domain folding, which showed that the initial assembly is between highly unstructured monomers that form a transient, highly structured dimeric intermediate.[171]

Recent measurements of the dynamics of p53 oligomerization in stressed and unstressed cells show that p53 oligomerization is a highly regulated process that is not solely determined by protein concentration but also modulated by additional factors.[172] Such factors include interacting proteins and PTMs. S100 proteins, for example, regulate the oligomerization states of all three p53 family members by differentially binding to their monomeric and tetrameric forms.[173,174] Phosphorylation of the C-terminal p53 region also increases its binding affinity to 14-3-3 proteins, which shifts the oligomerization equilibrium toward active tetrameric forms.[175,176] Stability of the p53 tetramer is also modulated by additional domain−domain interaction in a phosphorylation-dependent manner. Phophorylation of Ser392 at the p53 C-terminus, for example, stabilizes the tetramer and enhances transcriptional activity.[177] This region has recently been shown to interact with the DBD of a neighboring tetramer subunit.[178]

p63 plays an important role in protecting the female germline in mice, which involves a unique regulation of its oligomerization state via concerted action of N- and C-terminal regions.[126] While p53 protein levels are kept constitutively low in unstressed cells, TAp63α is expressed at high levels in unstressed mouse oocytes.[179] It is, however, present as an inactive dimer. In this latent form, a short helical segment within the natively unfolded TAD and the C-terminal inhibitory region interact with the oligomerization domain, resulting in a closed dimeric conformation that prevents assembly of dimers into tetramers.[126] γ-Radiation-induced phosphorylation of TAp63α triggers a switch from inactive dimers to active open tetramers, resulting in apoptosis of premature oocytes.[126,179] Although the regions involved in the autoinhibition of TAp63α are largely conserved in TAp73α, the latter exhibits no such inhibitory mechanism and forms open active tetramers.[180] It has therefore been suggested that the regulation of p73 transcriptional activity might be more closely related to p53 than to its structural homologue p63.[180]

5.1.1.3. p53 C-Terminal Regulatory Domain: Binding Diversity through Chameleon Sequences and One-Dimensional Sliding on DNA. The intrinsically disordered C-terminal region of p53 is subject to extensive PTMs in both stressed and unstressed cells, ranging from phosphorylation to ubiquitination, acetylation, methylation, neddylation, and sumoylation (Figure 7).[124,125] It displays unique binding promiscuity mediated through a chameleon sequence, a special case of a MoRF, in conjunction with alternative modification patterns.[10,181] Chameleon sequences do not have pronounced preferences for either helical or β-strand conformations and can therefore adopt different secondary structures, depending on their structural context.[182−184] A comparison of different p53 structures shows that a short sequence motif within its C-terminal region can adopt α-helical, β-strand, and coiled conformations upon binding to different regulatory proteins (Figure 7). Residues 375−388 of human p53 undergo disorder-

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX−XXX

to-order transitions and adopt an α-helical structure upon binding to S110B.[185] In contrast, a conformation without regular secondary structure is observed in complex with phospho-CDK2/cyclin A.[186] Lys382-acetylated forms of the same sequence motif establish an intermolecular β-sheet structure in complex with a sirtuin deacetylase[187] and a β-turn upon binding to the bromodomain of the transcriptional coactivator CBP.[188] These examples illustrate the role of chameleon sequences in providing conformational adaptability and increased binding promiscuity and how protein interaction networks can be modulated by PTMs.

Besides serving as a protein interaction site, the C-terminal regulatory domain of p53 plays a crucial role in linear diffusion along DNA, facilitating rapid scanning for p53 targets sites. The DBDs form a sequence-specific complex with p53 response elements, whereas the C-terminal domain binds DNA non-specifically via its six lysine residues.[189,190] A model based on single-molecule and electron microscopy studies combined with in vivo data on p53 mutants suggests that fast sliding along DNA is mediated by the C-terminal domain (modulated by acetylation) and that scanning is accomplished by DBDs hopping on and off DNA until response elements are reached.[131,191,192] Such a two-state model of switching between different conformational states relies on a loose structure of the p53 protein with enough flexibility for rapid and coordinated domain movements.

5.1.2. Evolution of the p53 Family of Proteins and the p53-MDM2/MDMX Axis. The evolutionary history of the p53 family of proteins can be traced back to the beginning of multicellular life and may predate that of many other metazoan transcription factors. Two p53/p73-like proteins are predicted from the genome of the choanoflagellate *Monosiga brevicollis*, the closest living unicellular relative of metazoans.[193] p53/p73-like proteins are also found in the placozoan *Trichoplax adhaerens*,[194] one of the simplest multicellular organisms, and the starlet sea anemone *Nematostella vectensis*.[195] Upon exposure to low-level UV radiation, the p53 homologue nvp63 induces apoptosis in early gametes but not somatic cells of adult sea anemones, suggesting that the primordial function of p53 family proteins was to protect the germline from DNA damage.[195] The vertebrate p53 family with its three members p53, p63 and p73 evolved from a common p63/p73-like ancestral gene through two gene duplications and subsequent diversification.[196] All three paralogues were recently identified in the genome of the elephant shark (*Callorhinchus milii*), a member of the cartilaginous fishes, suggesting that the two gene duplications occurred at the beginning of vertebrate evolution.[197] Multiple p53 family genes are also found in some invertebrate lineages (e.g., sea anemones, flatworms, and mosquitoes) as a result of independent gene duplications.[196,198]

Most interestingly, p53 evolved at a much faster rate than p63 and p73,[197] and the diversity is most striking in their C-terminal regions, which diverged early in vertebrate evolution[197] and may be directly linked to the expansion of the p53 interactome, with p53 taking on novel roles in safeguarding genomic integrity of somatic cells. The structured DNA-binding domain is highly conserved, which reflects maintenance of DNA-binding specificity from metazoans to humans.[195,199] In contrast, the intrinsically disordered regions display high sequence divergence, supporting the notion that there is a general correlation between intrinsically disordered regions and evolutionary diversity.[200] In addition, the substitution patterns of residues in the intrinsically disordered regions are more diverse than those of residues in the structured DBD.[200] The rigid structural framework of the latter tolerates only a limited number of variations for any given residue, whereas there are fewer constraints for disordered regions, as long as the amino acid changes do not affect conformational sampling of functionally important regions or intermolecular contact sites. Phylogenetic distribution of predicted disorder in the p53 family of proteins also suggests a correlation of increased disorder with organismal complexity, with vertebrate proteins typically displaying a higher percentage of disordered regions than in metazoans.[200] This increase may relate to greater functional complexity of p53 family proteins in higher organisms.

Despite the high divergence of intrinsically disordered regions, the p53−MDM2 autoregulatory circuit is conserved in most animals. Genes of both MDM2 and p53 homologues have been identified in *Trichoplax*.[194] The N-terminal domain of the putative p53-family protein contains a sequence motif that resembles the MoRF of TAD1 in human p53, and molecular modeling suggests that this region interacts as an amphipathic helix with the N-terminal domain of the putative MDM2 protein, as observed in the complex of the human homologues.[194] MDM2-like genes are also found in cnidarians and molluscs but have not been detected in the model organisms *Drosophila melanogaster* and *Caenorhabditis elegans*, all of which contain p53 gene families.[201] This absence suggests that both species have lost key elements of the p53 pathway, including the MDM2 gene, and evolved alternative regulatory pathways. Their p53 proteins have significantly diverged from the common ancestral gene and share less sequence conservation with the human homologues than the p53 family proteins from placozoans and cnidarians.[194] With the emergence of vertebrates, we also see a duplication of the ancestral MDM2 gene, resulting in two sister proteins, MDM2 and MDMX, and a more complex regulatory network.[197,201] MDM2 and MDMX also contain large regions of intrinsic disorder,[202] further highlighting the pivotal role of intrinsic disorder in regulating the p53 pathway.

5.1.3. p53, Cancer, and Therapeutic Strategies. p53 is inactivated in virtually every cancer, either through direct mutation or via major perturbations in its regulatory pathways.[203] Different therapeutic strategies are required, depending on whether cancer cells express wild-type or mutant p53. Most p53 cancer mutations are located in the structured DBD, where they either result in the loss of an essential contact with DNA (contact mutants) or cause structural perturbations (structural mutants) that result in destabilization and functional inactivation.[204,205] Cancer mutations in intrinsically disordered regions of p53 are rare.[206] The p53 DBD is especially susceptible to inactivation by destabilizing mutations because it is marginally stable and unfolds at only slightly above body temperature, which seems to be a typical feature of vertebrate p53 proteins and may play a functional role.[207] Small molecules that bind to folded but not unfolded DBDs are poised to rescue these structural mutants. Y220C, for example, which accounts for approximately 75 000 new cancer cases per year, has been used to test such a strategy. This mutation creates a crevice on the protein surface that can be targeted by small-molecule stabilizers.[208] Screening of fragment libraries, in silico methods, and structure-guided design yielded a number of compounds that bind to this site, thereby increasing the thermodynamic and kinetic stability of the mutant.[209−212] Some of these compounds show promising biological activities in cancer cells

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX−XXX

harboring Y220C mutations and induce mutant-specific restoration of transcriptional functions, thus providing a proof of principle that small-molecule-induced reactivation of conformationally unstable p53 cancer mutants is a viable pharmacological strategy.[211,212] Alternatively, it may also be possible to target generic sites, as recently proposed for a transiently open binding pocket in the L1/S3 region of the DBD,[213] which could potentially lead to the stabilization of a whole set of different cancer mutants.

In many tumors with wild-type p53, MDM2 and MDMX are deregulated, keeping p53 levels low and blocking its transcriptional activity.[214] Several pharmacological strategies to reactivate p53 in these cancers therefore target MDM2 and MDMX, especially their interactions with p53 TAD1.[203] First, small-molecule antagonists of the p53−MDM2 interaction, known as nutlins, were published in 2004.[215] Nutlins are a family of *cis*-imidazoline analogues that bind to the p53 binding pocket in the N-terminal domain of MDM2 with high affinity by mimicking the interactions of the hydrophobic triad (Phe19, Trp23, and Leu26) of p53 TAD (Figures 5 and 8). Nutlin 3a induces expression of p53 target genes in cancer cells[215] and has become a valuable chemical probe for studying p53 pathways in living cells and organisms.[215−219]

Since the discovery of nutlins, numerous other small molecules that bind to the N-terminal domains of MDM and/or MDMX have been reported, including benzodiazepinediones,[220,221] molecules with spirooxindole core structure,[222] imidazoindoles,[223] isoindolinone-containing compounds,[224] and lithocholic acid,[225] an endogenous steroidal bile acid. A class of piperidinone inhibitors is particularly interesting from a structural biology point of view because they induce ordering of the MDM2 N-terminal lid region upon binding through direct hydrophobic and His96-mediated polar contacts.[226] A more recent structural study of MDM2 complexes with 6-chloroindole scaffolds further highlights the potential of incorporating this lid region in ligand design and exploiting transient protein states for the development of potent inhibitors of the MDM2−p53 interaction.[227]

Emerging strategies to block the p53−MDM2/MDMX interaction are hydrocarbon-stapled α-helical peptides.[228−230] Site-specific introduction of a hydrocarbon staple restricts conformational sampling and locks these p53-derived peptides in α-helical conformations, thereby increasing their target affinity. Stapled peptides have been used for a number of intracellular drug targets and have been shown to also improve pharmacological properties such as cellular uptake and resistance to proteolysis.[231] In the case of p53, the MDM2-interacting peptide comprises TAD residues 14−29 and displays 10% α-helical content in water, as determined by circular dichroism.[228] Replacing Ser20 and Pro27 with synthetic olefinic residues and generating a hydrocarbon staple by olefin metathesis increased α-helicity to almost 60% and improved its affinity for MDM2 about 400-fold.[228] Additional mutations of negatively charged residues were introduced to improve cell permeability, and the resulting stapled peptide SAH-p53-8 was shown to reactivate the p53 transcriptional pathway in cancer cell lines overexpressing MDM2 and MDMX.[228,232] The more recently developed stapled ATSP-7041 peptide is a highly potent dual inhibitor of MDM2 and MDMX. It binds to both targets with low nanomolar affinity and decreases proliferation of tumor cells in several MDM2/MDMX-overexpressing carcinomas.[230] A next-generation variant of ATSP-7041 with increased potency is scheduled to enter phase I clinical trials in 2014. Crystal structures of SAH-p53-8 and ATSP-7041 bound to MDM2 and MDMX, respectively, revealed that the hydrocarbon staple not only confers conformational stability but also makes direct contacts with the rim of the p53-binding pocket (Figure 8), thus further improving binding affinity.[230,233]

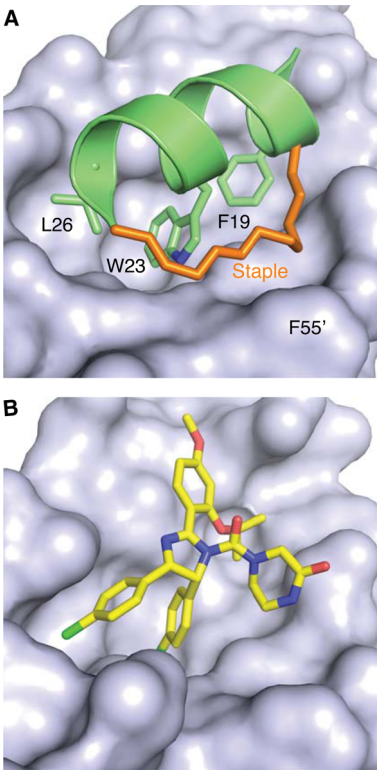


Figure 8. Targeting the p53−MDM2 interaction in cancer therapy. (A) Crystal structure of a p53-derived stapled peptide bound to the N-terminal domain of MDM2 (PDB entry 3V3B).[233] The molecular surface of MDM2 is shown in gray and the p53 helix in green, with the three key interacting side chains of the p53 helix (Phe19, Trp23, and Leu26) highlighted as stick models. The hydrocarbon staple (orange sticks) not only increases the helix propensity of the peptide but also forms specific interactions with Phe55 at the rim of the p53 binding pocket of MDM2. (B) Crystal structure of MDM2 in complex with the inhibitor nutlin 3a (PDB entry 4HG7).[378] The orientation is the same as in panel A. The rigid *cis*-imidazoline scaffold mimics key hydrophobic interactions made by the p53 helix. Two chlorophenyl groups are projected into the Trp23 and Leu26 subpockets, and a 2-propoxy group occupies the Phe19 subpocket.

Several small-molecule inhibitors of the p53−MDM2 interaction are currently in clinical trials (see Zhao et al.[234] for review). Whether any of these compounds makes it into the clinics remains to be seen. In addition to the main negative regulator MDM2, there are a number of other E3 ligases that sequester p53 for proteasomal degradation.[235] Depending on the cellular context, some of these may also be worthwhile targets. Modulating interactions of the disordered p53 C-terminus may also have therapeutic advantages, as in the case of the p53−Twist1 interaction, for example. Aberrant expression of the Twist1 transcription factor is common in sarcomas, and it was recently shown that Twist1 binds to the p53 C-terminus, preventing key PTMs and facilitating MDM2-dependent degradation of p53.[236] Disrupting the p53−Twist1 interaction might therefore result in restoration of p53 function in tumors with high Twist1 expression. Additional structural and functional insights into the p53 interactome and its

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX−XXX

deregulation in different types of cancers should provide novel avenues for targeting the intrinsically disordered regions of p53 for future cancer therapy.

5.2. PTEN, a Dual Protein/Lipid Phosphatase

PTEN (phosphatase and tensin homologue deleted on chromosome 10) is the second most frequently mutated tumor suppressor gene after p53. Somatic PTEN mutations are observed in solid tumors, while germline mutations are associated with PTEN tumor hamartoma syndromes (PTHS). Furthermore, epigenetic repression of PTEN transcription and nongenomic reduction in PTEN activity are associated with diseases affecting multiple organs.[237] PTEN, a dual protein/lipid phosphatase, regulates signaling via the pro-survival, proliferative PI3K/AKT/mTOR pathway (Figure 9),[238,239]

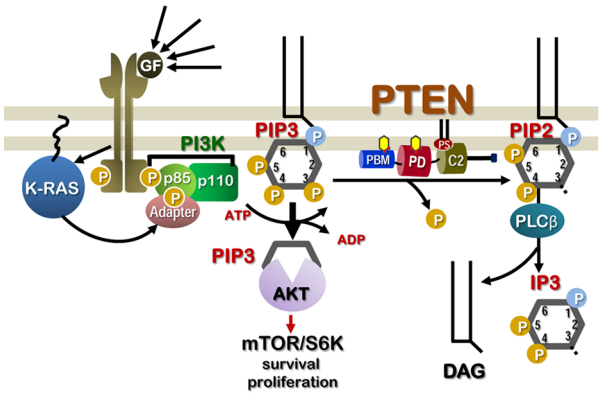


Figure 9. Role of PTEN in PI3K/AKT/mTOR signaling and cell proliferation following growth factor binding to its receptor. Growth factors such as epidermal growth factor (EGF) and insulin-like growth factor (IGF) bind to receptor tyrosine kinases (RTKs). Ligand-induced receptor dimerization and subsequent autophosphorylation of distinct tyrosine residues creates docking sites for various membrane targets via distinct adaptor molecules. Depicted here is the lipid kinase PI3K (phosphoinositide 3-kinase), consisting of regulatory subunit p85 and catalytic subunit p110, that phosphorylates PIP2 to generate the second messenger PIP3. Increase in cellular PIP3 levels results in the recruitment and activation of protein kinase AKT, thus initiating diverse cellular pathways leading to increase in proliferation, migration, gene transcription, cell cycle progression, and cell survival and changes in cell metabolism. For brevity, only the mTOR/S6K components of the pathway are shown here. Active small GTPase K-Ras also acts on PI3K and activates the mTOR pathway. Hyperactivation of the mTOR pathway causes aberrant cell proliferation and cancer. Therefore, PTEN, a lipid phosphatase, continuously removes a 3′-position phosphate group (P) in the inositol ring and converts PIP3 to PIP2. Decrease in PIP3 levels keeps the signaling via the PI3K/AKT/mTOR pathway in check. Later, PIP2 is hydrolyzed by phospholipase Cb (PLCb) to diacylglycerol (DAG) and inositol triphosphate (IP3); both molecules are involved in distinct calcium-dependent signaling events. Thus, since PTEN plays a central role in cell physiology and growth by regulating a myriad of downstream signaling events, understanding PTEN structure and its correlation to function is critical to modulating PTEN activity via targeted molecular therapies.

influences glucose metabolism, and regulates cell polarity and cellular senescence.[237] The phosphatase-independent functions of PTEN include its role in development, cell cycle regulation, maintenance of cell polarity, and genomic stability.[240] The functional versatility of PTEN is attributed to its intrinsically disordered protein regions (IDPRs), allowing it to interact with over 400 proteins in different subcellular compartments such as plasma membrane, cytoplasm, cell nucleus, and exosomes.[37] Secretory PTEN may likely interact with many unknown proteins in blood plasma and stromal compartment, altering the tumor microenvironment.[241]

5.2.1. Domain Organization and Evolution of PTEN.

Structurally, PTEN consists of (a) the N-terminal phosphatidylinositol 4,5-bisphosphate (PIP2) binding module (PBM), which allows PTEN to anchor to the membrane (its site of action) via PIP2 molecules; (b) the dual-specificity lipid and protein phosphatase domain (PD), which contains a conserved catalytic motif HCKAGKGR; (c) the C2 domain, consisting of positively charged residues that help PTEN to associate with phosphatidylserine residues in the plasma membrane; and (d) the C-terminal region (C-tail) (Figure 10B), which regulates membrane association and stability through several phosphorylation modifications. While the PD and C2 domains are structured (Figure 10A),[242] the N-terminal PBM and C-tail regions are disordered (Figure 10C).[37]

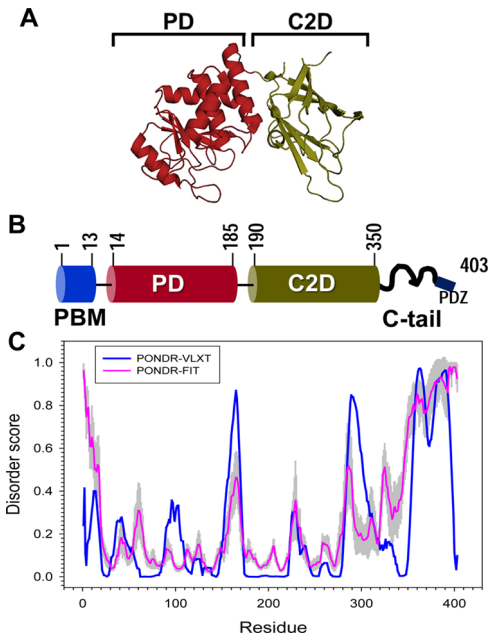


Figure 10. Domain organization of PTEN and disorder prediction. PTEN consists of four domains: PIP2 binding module (PBM), phosphatase domain (PD), C2 domain (C2D), and C-terminal tail (C-tail). Adapted from ref 37 under the Creative Commons Attribution-Non Commercial-ShareAlike 3.0 license. Copyright 2013 Nature Publishing Group. (A) PTEN crystal structure (PDB entry 1D5R)[242] consists of only the PD and C2 domains. The structures of the N-terminal PBM and C-tail are largely unknown. (B) Functional domains in the PTEN protein. The PTEN protein consists of four functional domains: PBM (residues 1−13), which helps it anchor to the PIP2 residues in the membrane (its site of enzymatic activity); catalytic domain PD (residues 14−185), which has dual lipid and protein phosphatase activity; C2 domain (residues 190−350), primarily consisting of positively charged residues to help PTEN anchor to the membrane; and C-tail region (residues 351−403), which regulates PTEN function, membrane association, and stability through a series of phosphorylation events. A PDZ binding motif is part of the C-tail and helps PTEN interact with a multitude of PDZ-domain containing proteins. (C) Disorder prediction for PTEN. PONDR-VLXT and PONDR-FIT predictors were used to determine disordered regions in the PTEN protein. A value above 0.5 indicates structural disorder. The major disordered stretches in the PTEN protein consist of the CBR3 loop in the C2 domain (residues 286−309) and the C-tail (residues 351−403).

dx.doi.org/10.1021/cr400713r | Chem. Rev. XXXX, XXX, XXX−XXX

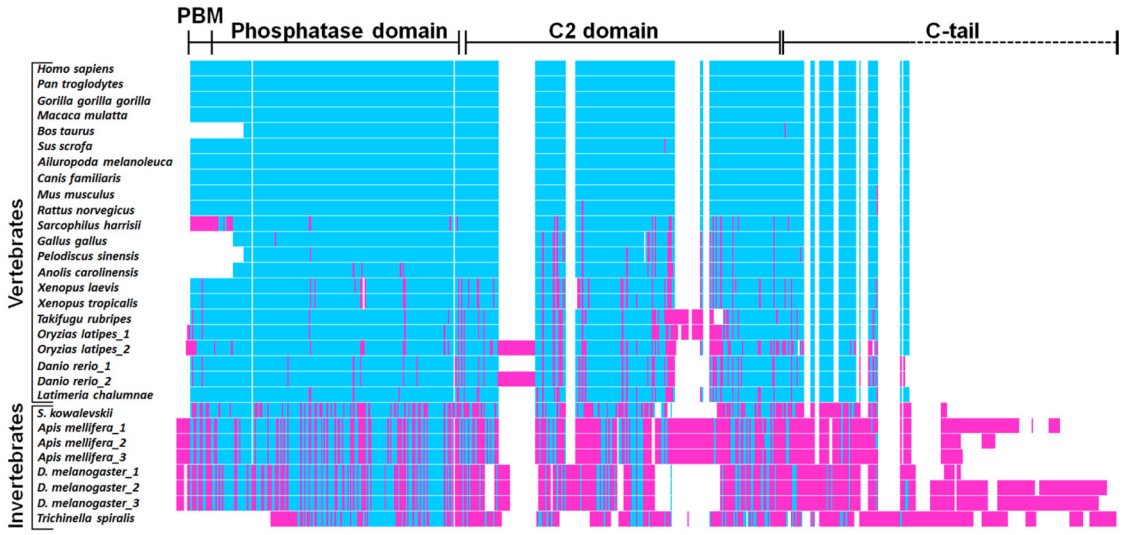


Figure 11. Evolutionary conservation of the PTEN IDPR. Sequence alignment for PTEN proteins from different species reveals that the disordered PTEN C-tail is conserved only in vertebrates. This points to the recent emergence of the functional implications of the C-tail in PTEN function in evolutionary history. Adapted with permission from ref 37. Copyright 2013 Nature Publishing Group.

The PTEN gene emerged early in evolution, being present in primitive organisms including slime molds, *C. elegans*, and *D. melanogaster*. PTEN phosphatase and C2 domains are largely conserved across different vertebrate and invertebrate species, with 100% conservation of the dual specificity catalytic motif. However, C-tail IDPR conservation is observed only in vertebrates, indicating that this IDPR emerged late in evolution with unique functional consequences (Figure 11).[37] Most of PTEN's functional versatility seen in higher eukaryotes is executed by its intrinsically disordered C-tail.[37]

A longer variant of PTEN, PTEN Long, containing an additional 173 amino acids at its N-terminus (referred to herein as the N-173 region) (Figure 12A), is a product of alternative translation initiation at a noncanonical start site (CTG), which is 519 bp upstream of the canonical ATG.[243] 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.[243] A polyalanine stretch within the N-173 region is essential to its secretion, while a polyarginine stretch is critical for uptake.[243] Furthermore, the N-173 region is enriched in nonpolar and positively charged polar amino acids, facilitating its transduction across negatively charged membranes.[241]

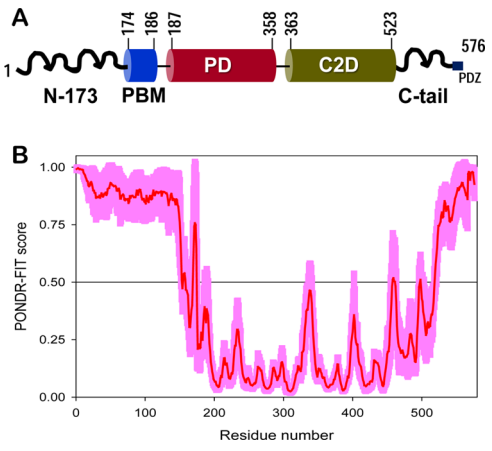


Figure 12. Intrinsic disorder and PTEN Long. (A) Diagrammatic representation of PTEN Long, a translational variant of PTEN that contains an extra 173 amino acids at its N-terminus (N-173). The peculiarity of PTEN Long is that it can be secreted from and taken up into cells at distant locations in the body. This secretory property of PTEN Long is conferred by critical polyalanine and polyarginine stretches in the N-173 region. (B) Disorder prediction for PTEN Long. PONDR-FIT was used to make disorder predictions for the PTEN Long protein. The additional N-173 amino acids in this PTEN translational variant are largely disordered. Reprinted with permission from ref 241. Copyright 2013 Royal Society of Chemistry.

Consistent with its amino acid bias, the N-173 region is largely disordered (Figure 12B).[241] Sequence analysis of PTEN Long across different species revealed that the N-173 region emerged late in evolution (Figure 13A).[241] PTEN Long functions just like PTEN. Thus, increased levels of PTEN Long observed in the stroma of breast tumors suggest a regulatory role in tumor maintenance.[243] Presence of PTEN Long in human plasma and serum samples indicates that circulating PTEN is a check on cell growth, shape, and motility as required.[243] However, the mechanism inducing PTEN Long production and its mechanism of distribution throughout the body and to tumors is not known and remains an active area of research. In summary, the N-173 region, by virtue of its cell-penetrating properties, can be utilized as a drug delivery system not only for PTEN Long but also for other proteins, opening up avenues for therapeutic intervention in cancers (Figure 13B).[241,243]

5.2.2. Regulation of PTEN. Tight regulation at the genomic, transcriptional, and translational levels is a hallmark of IDPs.[19] Consistent with this model, cells employ a variety of genomic and nongenomic mechanisms to regulate PTEN function to enable signaling fidelity (Figures 14 and 15). These inherent regulatory mechanisms ensure that appropriate amounts of PTEN are present within the cell, eliminating the possibility for formation of any nonspecific interactions or functional associations that may cause aberrant signaling. Dysregulation in any one of these regulatory processes is associated with cancers and myriad pathologies in multiple organs.[237] The genomic and nongenomic mechanisms regulating PTEN levels and activity in cells are outlined below.

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX−XXX

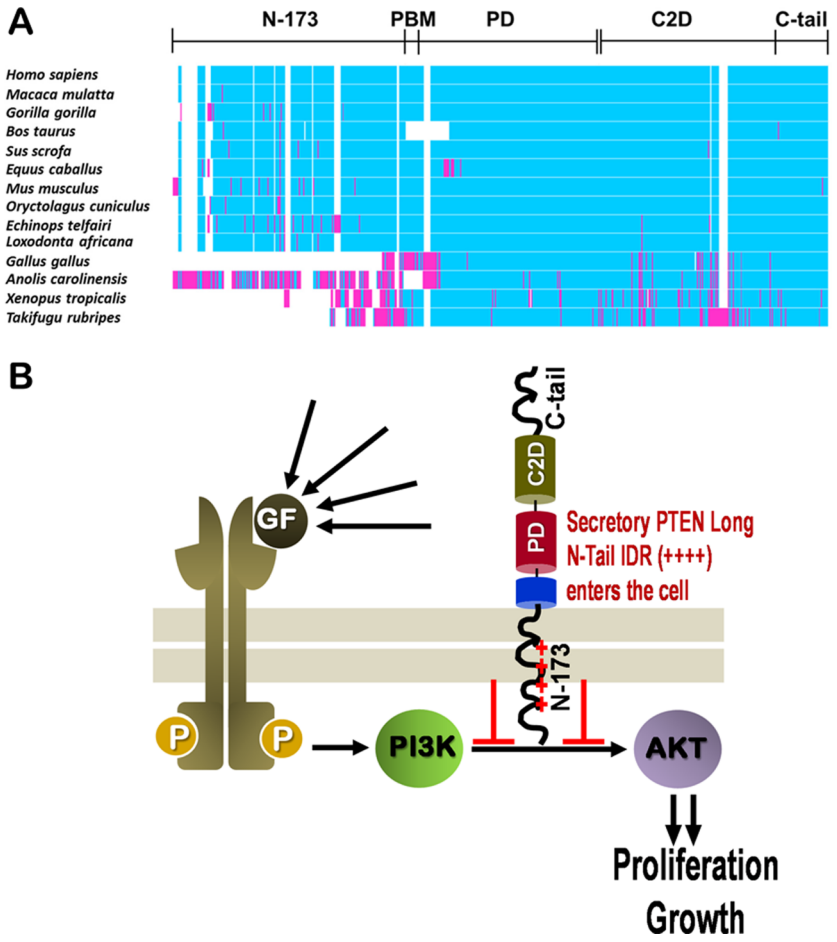


Figure 13. Sequence analysis and clinical implications of PTEN Long. (A) Sequence analysis. Alignments for the PTEN Long protein reveal the emergence of the additional N-173 region in higher terrestrial vertebrates. (B) Therapeutic relevance. Given the ability of PTEN Long to be taken up into cells, recombinant forms of the PTEN Long protein represent a novel chemotherapeutic modality. Recombinant PTEN Long may be taken up into cells where it can then abrogate proliferative PI3K/AKT/mTOR signaling to arrest tumor growth. Reprinted with permission from ref 241. Copyright 2013 Royal Society of Chemistry.

5.2.2.1. Genomic Mechanisms. Genomic mechanisms regulating PTEN function include mutations and epigenetic and transcriptional phenomena at the PTEN gene. Except for exon 9, which encodes the disordered C-tail,[37] PTEN mutations span the entire gene (Figure 14A). Complete allelic deletions of the PTEN gene and mutations in the PTEN promoter also occur frequently.[244] Epigenetic regulation mainly occurs via promoter methylation, which prevents PTEN transcription, reducing its protein levels, which is frequently observed in several types of cancers (Figure 14B),[244] indicating that PTEN deficiency either drives or exacerbates the cancer phenotype. PTEN expression is also regulated by several transcription factors (Figure 14C).[237,245] Sal-like protein 4 (SALL4), EMT transcription factor SNAIL, inhibitor of DNA binding 1 (ID1), ecotropic virus integration site 1 protein (EVI1), BMI1, and c-Jun repress PTEN transcription either directly or indirectly by competitively preventing the binding of transcriptional activators to the PTEN promoter.[237] Mitogen-activated protein kinase kinase 4 (MKK4), 17β-estradiol, and insulin-like growth factor 1 (IGF1) suppress PTEN transcription via the NF-κB transcription factor,[246] while transforming growth factor β (TGF-β) suppresses PTEN transcription by recruiting the SMAD transcription factors.[245,246] Several of these inhibitory transcription factors are dysregulated in cancers, downregulating PTEN, which is a critical step in oncogenic cellular transformations.[237] Early growth-response protein 1 (EGR1), insulin-like growth factor 2 (IGF2), peroxisome proliferator-activated receptor γ (PPARγ), p53 tumor suppressor protein, sprouty homologue 2 (SPRY2), and resistin, a cytokine, increase PTEN expression at the RNA level (Figure 14C). Interestingly, several naturally occurring compounds like phytoestrogens (found in soy), resveratrol (in red wine), and quercetin and indole-3-carbinol (in broccoli) upregulate PTEN mRNA expression[245] and are currently under investigation for use as chemopreventive agents.

Recent evidence suggests that PTEN mRNA levels are regulated by a host of noncoding RNAs (Figure 14D). Several miRNAs, such as miR-17−92, miR-19, miR-19b, miR-21, miR-22, miR-23b, miR-26a, miR-29b, miR-92a, miR-106b~25, miR-155, miR-214, miR-216a, miR-217, miR-221, miR-222, miR-301, miR-486, and miR494 among others, regulate PTEN expression and are frequently altered in cancers, autoimmune disorders, and cardiovascular and metabolic diseases.[237,247] In fact, several oncogenic molecules upregulate miRNA expression to keep a check on cellular PTEN protein levels during tumorigenesis.[237] The PTENP1 pseudogene protects PTEN mRNA from miRNA inhibitory effects. PTENP1 mRNA shares sequence homology with PTEN mRNA, works as a decoy, and acts as a sponge, sequestering all miRNAs that suppress PTEN function (Figure 14D).[237,248] Consistent with the role of

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX−XXX

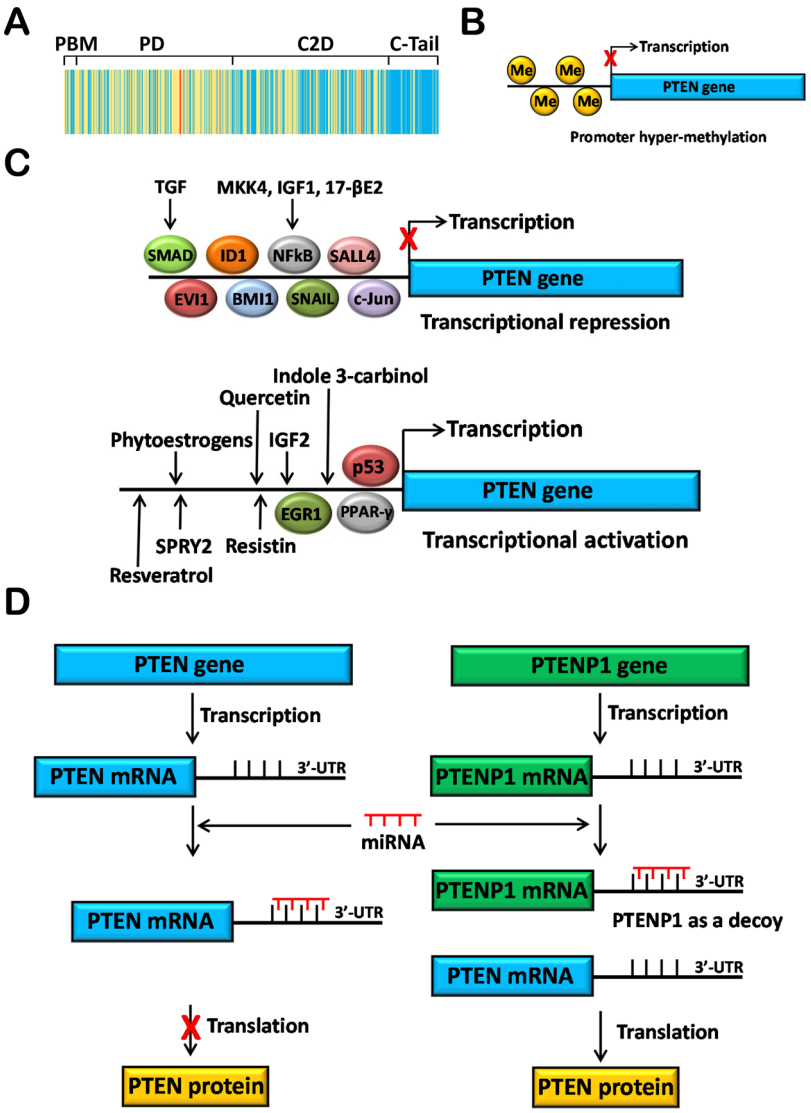


Figure 14. Genomic mechanisms of PTEN regulation. (A) Mutations. PTEN mutations are seen in several diseases like Cowden disease, hamartoma syndrome, and cancers. Mutations are found all along the length of the PTEN gene with the exception of exon 9, which codes for the disordered C-tail region. Mutational hotspots are indicated in shades of yellow, while areas with fewer observed mutations are indicated in blue. Data are taken from ref 379. (B) Epigenetic regulation. PTEN promoter methylation is frequently observed in cancers, resulting in suppression of its transcription. (C) Transcriptional regulation of PTEN. Several transcription factors, ligands, and dietary compounds modulate (increase or decrease) PTEN transcription. (D) Regulation by noncoding RNAs. Several miRNAs bind to the 3′-untranslated region (UTR) of the PTEN mRNA, thereby preventing translation of the mRNA into PTEN protein. PTENP1, a PTEN pseudogene, shares homology with the PTEN gene. The PTENP1 mRNA acts as a pseudosubstrate for the miRNAs and sequesters them, thereby allowing production of the PTEN protein.

PTENP1 in regulating PTEN expression levels, loss of PTENP1 is observed in breast and colon cancers.[237,248]

5.2.2.2. Nongenomic Mechanisms. Subcellular localization, degradation, and turnover of PTEN protein affect cell physiology; aberrations in these processes lead to pathological conditions. PTEN localizes to the plasma membrane, cytosol, nucleus, and exosomes, while PTEN Long is secreted out (Figure 15A). Cytosolic and membrane localization of PTEN is dictated by phosphorylation modifications in its disordered C-tail region. Several PTEN-binding proteins, such as MAGI2, MAGI3, myosin V, caveolin, and MEK1, enhance its membrane binding.[237,249,250] Nuclear localization of PTEN occurs via monoubiquitination modifications at Lys13 and Lys289 by the E3 ligase NEDD4-1 (neural precursor cell expressed, developmentally downregulated 4-1),[245] while the role of E3 ligase XIAP is less clear.[251] Once in the nucleus, PTEN maintains chromosomal integrity and stability and controls cell cycle progression.[237] Lack of nuclear PTEN is associated with aggressive cancers.[237]

Monoubiquitinated cytoplasmic PTEN has several possible fates within the cell, including further ubiquitination and degradation, shuttling between the nucleus and the cytoplasm, or deubiquitination, which retains it in the nucleus.[245] HAUSP (herpesvirus-associated ubiquitin-specific protease) deubiquitinates PTEN in the nucleus. Loss of HAUSP function is frequently seen in leukemias.[237] p63 is a transcription factor that belongs to the p53 gene family (see section 5.1). A p63 splice variant, ΔNp63α, inhibits the nuclear translocation of PTEN by inhibiting NEDD4-mediated monoubiquitination of PTEN.[252] Lysine residues Lys13 and Lys289 are frequently mutated in Cowden syndrome, resulting in the nuclear exclusion of PTEN. Given the disease-associated aberrations

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX–XXX

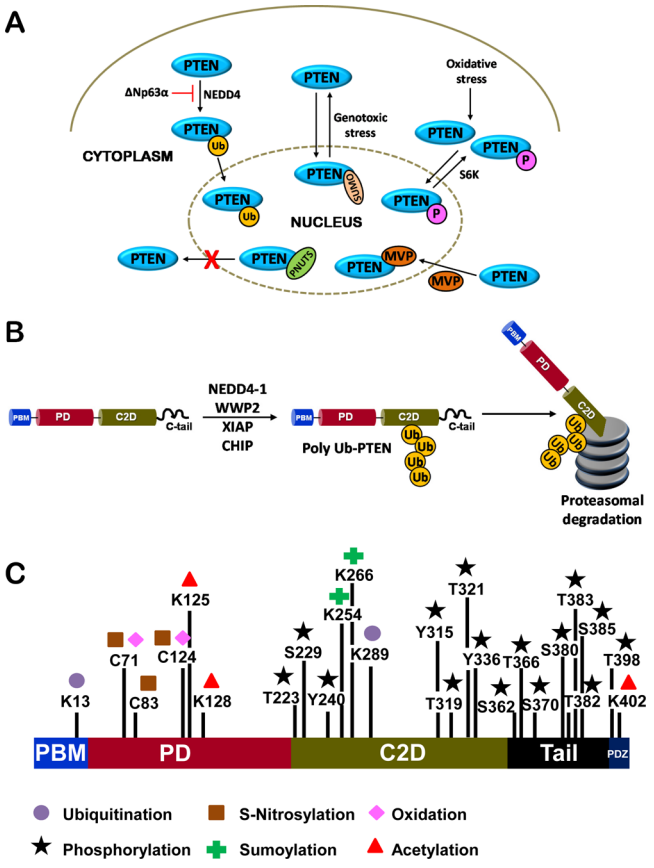


Figure 15. Nongenomic mechanisms of PTEN regulation. (A) Subcellular localization of PTEN. Monoubiquitination and sumoylation modifications on the PTEN protein regulate its nuclear trafficking. Oxidative stress is also known to modulate nuclear PTEN levels. (B) Proteasomal degradation of PTEN. Several E3 ligases like WWP2, CHIP, XIAP and NEDD4-1 cause polyubiquitination and subsequent proteasomal degradation of PTEN. The NEDD4 and CHIP E3 ligases are frequently upregulated in human cancers and negatively correlate with PTEN protein levels. (C) Posttranslational modifications of PTEN. The PTEN protein, like most IDPs, undergoes a vast repertoire of PTMs, each with different functional implications. Of the various modifications, the phosphorylation modifications are best studied and are critical to regulating PTEN function and stability in the cell.

in PTEN nuclear import and localization, nuclear PTEN levels have potential prognostic value in disease progression.[237] PTEN sumoylation at Lys254 is required for its nuclear translocation and retention, an event that is dependent on PTEN phosphorylation by protein kinase ataxia telangiectasia mutated (ATM).[253]

Genotoxic stress and DNA damaging agents cause nuclear export of sumoylated PTEN,[253] while oxidative stress inhibits nuclear export. The increase in nuclear PTEN levels in response to oxidative stress subsequently results in elevated p53 levels as a measure to limit oxidative DNA damage.[254] PTEN nuclear accumulation upon oxidative stress depends on the phosphorylation status of Ser380 within the disordered C-tail.[254] A recent study showed that phosphorylation of Ser380 by S6K (S6 kinase) in human endothelial cells triggers PTEN deubiquitination and nuclear export.[255] Protein phosphatase 1 nuclear targeting subunit (PNUTS) is an oncogene that allows tumor proliferation by sequestering PTEN in the nucleus. PNUTS-mediated nuclear localization of PTEN is independent of monoubiquitination.[256] Major-vault protein (MVP) is also known to mediate nuclear translocation of PTEN.[245]

Degradation and turnover of PTEN occurs via polyubiquitination catalyzed by E3 ligases NEDD4 (neural precursor cell-expressed developmentally downregulated 4-1), XIAP (X-linked inhibitor of apoptosis protein), WWP2 (WW-domain containing protein 2), and CHIP (carboxyl terminus of Hsc70-interacting protein) (Figure 15B). Upregulation of NEDD4 is observed in lung and breast cancer tissue samples and correlates inversely with PTEN protein levels.[251] p34SE-1 increases NEDD4 expression at both the mRNA and protein level, resulting in increased degradation of PTEN.[257] NEDD4-mediated degradation of PTEN is inhibited by Rak through phosphorylation at Tyr336. Consistent with the role of Rak in regulating PTEN stability, loss of heterozygosity of the Rak gene has been observed in breast cancers.[251] Similarly, the levels of CHIP negatively correlate with PTEN levels in human prostate cancer samples.[258]

5.2.2.3. Posttranslational Modification of PTEN. PTEN phosphorylation in the C2 domain and the C-terminal tail involves serine, threonine, and tyrosine residues (Figure 15C). Casein kinase II phosphorylates Ser370, Ser380, Thr382, Thr383, and Ser385, increasing PTEN stability but decreasing its activity (discussed below).[254] Ser380 is also phosphorylated by S6K, resulting in the nuclear exclusion of PTEN.[255] Ser362 and Thr366 are phosphorylated by glycogen synthase kinase 3-β (GSK3-β).[254] In addition, Plk3 (polo-like kinase 3) phosphorylates PTEN at Thr366 and Ser370.[259] Thr398 is phosphorylated by ATM and regulates nuclear transport of PTEN in conjunction with sumoylation.[253]

RhoA-associated protein kinase (ROCK) phosphorylates PTEN at Thr223, Ser229, T319, and T321, thereby modulating its intracellular localization during cell migration and chemotaxis.[237] A tyrosine kinase, Rak, phosphorylates PTEN on Tyr336 and downregulates ubiquitination by the E3 ligase NEDD4-1.[237] Src kinase phosphorylates PTEN at Tyr240, Tyr336, and Tyr315. Phosphorylation at these sites upregulates the PI3K pathway through destabilization of PTEN.[254] Therefore, the stability of PTEN is controlled differentially by the cytoplasmic level of various kinases involved in its phosphorylation. Several molecules modulate PTEN phosphorylation. Leptin enhances phosphorylation at Ser380, Thr382, and Thr383,[254] while prostaglandin E2 (PGE2) and TGF-β inhibit PTEN activity through phosphorylation at serine residues.[254] PICT-1, a product of the GLTSCR2 (glioma tumor suppressor candidate region gene-2 protein) enhances phosphorylation of PTEN at Ser380, thereby protecting it from degradation.[254] Tissue ischemia alters PTEN phosphorylation, causing its inactivation and triggering a pro-survival response via the PI3K/Akt pathway.[254]

More than 50% of phosphorylation modifications are clustered in the disordered C-tail region. Phosphorylation events at Ser380, Thr382, Thr383, and Ser385 within the disordered C-tail region regulate PTEN activity and stability (Figure 16). Upon phosphorylation, the C-tail folds onto the PTEN molecule and produces a "closed" conformation.[260−262] Due to occlusion of its membrane-binding C2 domain, the closed form of PTEN shows reduced membrane binding and hence reduced phosphatase activity. However, the "closed" conformation of PTEN represents a more stable form of the protein.[261−263]

Acetylation and ubiquitination alters PTEN interaction and subcellular localization. PCAF, a histone acetyltransferase

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX−XXX

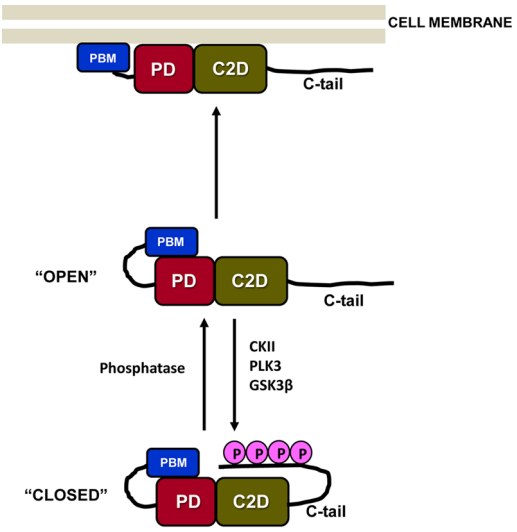


Figure 16. Intramolecular interactions of PTEN. The PTEN molecule forms intramolecular interactions, an event that is dependent on phosphorylation modifications within the disordered C-tail region. Phosphorylation at a serine−threonine cluster (Ser380, Thr382, Thr383, and Ser385) in the C-tail region causes the phosphorylated tail region to fold onto the rest of the PTEN molecule, forming a "closed" conformation. This intramolecular association masks the C2 domain, disallowing interaction of the PTEN molecule with the cell membrane (its site of enzyme action). As a result, the "closed" form of PTEN is enzymatically inactive. However, this "closed" form is relatively resistant to proteasomal degradation. Dephosphorylation of the serine−threonine cluster reverses this intramolecular association, resulting in the "open" conformation of PTEN. "Open" PTEN can now bind to the plasma membrane and dephosphorylate PIP3 to PIP2, thereby keeping the proliferative PI3K/AKT/mTOR pathway in check.

protein (HAT), acetylates PTEN at Lys125 and Lys128, inhibiting the phosphatase activity, which results in increased PI3K activity (Figure 15C).[254] CBP, another HAT, acetylates PTEN at Lys402, increasing its interaction with the PDZ domain containing protein MAGI-2, which results in increased translocation of PTEN to cell junctions.[254] Mono- and polyubiquitination of PTEN causes nuclear translocation and proteasomal degradation, respectively. Ret finger protein (RFP) is a newly identified PTEN E3 ligase responsible for atypical polyubquitination of PTEN at multiple lysine residues within the C2 domain,[264] inhibiting its phosphatase activity and thereby potentiating PI3K activity.[264]

Oxidative stress inactivates PTEN by triggering formation of a disulfide bridge between Cys71 and Cys124 within its catalytic pocket (Figure 15C).[254] PTEN shields itself from oxidative inactivation through a peroxidase enzyme called peroxiredoxin,[254] while thioredoxin regenerates catalytically active PTEN once it has been oxidized. S-Nitrosylation of PTEN occurs at residues Cys83, Cys71, and Cys124, enhancing its degradation via ubiquitination; this is observed in early Alzheimer's disease and animal models of cerebral ischemia.[254]

5.2.3. Protein−Protein Interactions of PTEN and Their Implication in Disease. Several protein−protein interactions modulate cellular sublocalization, stability and phosphatase activity of PTEN.[237,245,249,250,256,264−266] Most of these interactions (PPIs) target the disordered C-tail region,[37] reinforcing the critical and versatile role the PTEN C-tail IDPR plays in cellular homeostasis. PTEN interacting proteins and their effects on phosphatase activity and stability are listed in Table 1.

Table 1. Regulation of PTEN through Protein−Protein Interactions

(A) Protein−Protein Interactions Regulating PTEN Catalytic Activity	
PTEN-binding protein	effect on PTEN lipid phosphatase activity
MAGI2 (membrane-associated guanylate kinase inverted 2)	increase
MAGI3 (membrane-associated guanylate kinase inverted 3)	increase
NHERF (Na^+/H^+ exchanger regulator factor)	increase
β-arrestin	increase
myosin V	increase
DLG1 (disks large homologue)	increase
p85 (regulatory subunit of PI3K)	increase
MEK1 (dual-specificity mitogen-activated protein kinase kinase 1)	increase
RFP (Ret finger protein)	decrease
PNUTS (protein phosphatase 1 nuclear targeting subunit)	decrease
BMI1 (polycomb complex protein BMI-1)	decrease
DJ-1 (PARK7, Parkinson protein 7)	decrease
PREX2a (phosphatidylinositol 3,4,5-trisphosphate-dependent RAC exchanger factor 2a)	decrease
sharpin	decrease
MAN2C1 (α-mannosidase 2C1)	decrease
(B) Protein−Protein Interactions Regulating PTEN Protein Stability	
PTEN-binding protein	effect on PTEN stability
MAST1 (microtubule-associated serine/threonine kinase 1)	increase
MAST3 (microtubule-associated serine/threonine kinase 3)	increase
PICT1 (protein interacting with carboxyl terminus 1)	increase
ROCK (RhoA-associated protein kinase)	increase
DLG1 (disks large homologue)	increase
MC1R (melanocortin-1 receptor)	increase

Approximately 400 proteins were identified in the primary interactome of PTEN (Figure 17A), of which 40 proteins directly interact with known regions in the PTEN molecule, including 25 that associate with the C-tail region.[37] A disease enrichment analysis of PTEN PPIs comprising ~400 proteins (both mapped and unmapped) identifies cancer, infectious diseases, and neurological diseases as the top three diseases associated with the PTEN primary interactome (Figure 17B). Approximately 65% of the primary interactome is associated with cancer (Figure 18A), which is consistent with PTEN's role as a major tumor suppressor gene.

However, a significant number of proteins that interact with PTEN also feature in other diseases, including infectious diseases (111 proteins) (Figure 18B), neurological diseases (113 proteins) (Figure 18C), and skeletal and muscular disorders (115 proteins), thus expanding the range of disorders associated with PTEN and its interactome.

5.2.4. PTEN as a Therapeutic Target. Dysregulation/loss of PTEN function, associated with aberrant PI3K/AKT/mTOR signaling, is observed in several malignancies, neurological disorders, and cardiopulmonary diseases. Augmentation/restoration of PTEN function is beneficial in these pathological conditions. Currently, two strategies are used to target PTEN function. First, kinase inhibitors against various components of

dx.doi.org/10.1021/cr400713r | Chem. Rev. XXXX, XXX, XXX−XXX

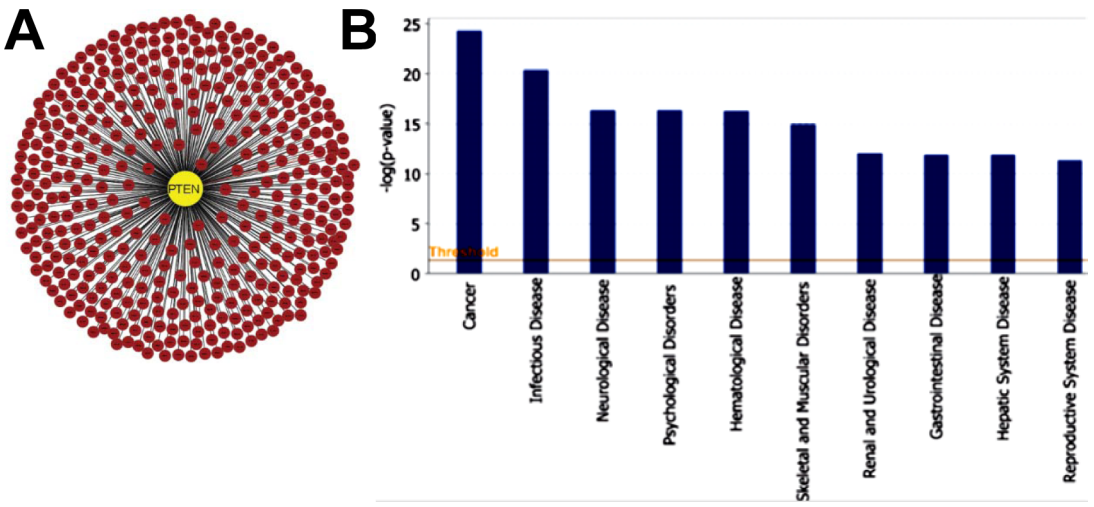


Figure 17. PTEN primary interactome. (A) 395 proteins form the primary interactome of PTEN. (Visualization tool: Cytoscape). (B) Functional analysis of the PTEN primary interactome. UniProt IDs from all the PTEN primary interactome were used as the input for the functional analysis performed by use of the Core Analysis function from the IPA suite (Ingenuity Systems, Mountain View, CA; www.ingenuity.com). The significance calculated for each function returned in Functional Analysis is a measurement of the likelihood that the function is associated with the data set by random chance. On the *y*-axis of the diagram, the significance is expressed as the negative exponent of the *p*-value calculated for each function. Taller histogram bars are more significant than shorter bars. Threshold represents $P < 0.05$.

the PI3K/AKT/mTOR cascade are popular and are extensively being used for the pharmacotherapy of these hyperproliferative diseases in the clinical setting.[267−271] Second, kinase inhibitors that block PTEN phosphorylation, thereby keeping it active, are also gaining popularity.[272−275] Several analogues of the natural compound curcumin have been found to increase PTEN expression levels by targeting miRNAs or other molecules that downregulate PTEN function.[276−280] Other potential approaches that can be developed in the future include inhibition of E3 ligases that cause proteasomal degradation of PTEN, inhibition of miRNAs that downregulate PTEN transcript levels, and inhibition of protein binding partners that negatively regulate PTEN stability or enzymatic activity. Furthermore, the recently identified isoform PTEN Long has secretory properties and represents a novel therapeutic target for pathological conditions associated with aberrant PI3K/PTEN/AKT signaling (Figure 13B).[241,243]

6. INTRINSICALLY DISORDERED PROTEINS IN PROTEIN−PROTEIN INTERACTION INHIBITION

Proteins exist along a continuum from highly ordered and minimally dynamic to fully disordered and extended with maximum dynamics. In the range between the extremes, proteins exist with varying amounts and amplitudes of disorder. Traditionally, small-molecule effectors, and drug discovery, began with a focus on enzymes. These proteins tend to fall on the structured end of the continuum. Membrane receptors provided another class of small-molecule targets and are also biased toward structure. Both of these classes have been extremely fruitful as targets and account for the bulk of approved drugs. As knowledge and understanding have progressed, the capability also to target the "high-hanging fruit",[281] protein−protein interactions (PPIs), has been developed. Given the high participation of IDPRs in protein−protein interactions, this has led to a concomitant increase in the targeting of IDP interactions as either one, or both, of the interaction partners.

As proteins themselves exist on a continuum of structure−disorder,[282] there is also a continuum of the degree to which protein disorder is a component of small-molecule targets of protein function. On one end are many active-site inhibitors of enzymes, with generally (though not universally) low participation of disorder (Figure 19). Along the scale of increasing dynamics would be allosteric effectors, where dynamics and disorder have been increasingly recognized both as a mechanism for transmitting information[283] and a means of enzyme inhibition by trapping species in more dynamic states.[284] Bimolecular, protein−protein targets provide for greater participation of disorder; the binding of a helix mimetic or other small molecules to the interaction site on a structured partner can displace an IDPR that no longer undergoes coupled binding and folding and hence shifts the population to greater disorder. Furthest toward disorder is the direct targeting of disordered regions. Here, protein−protein interactions are disrupted by binding to a disordered region, thereby disfavoring a coupled binding and folding interaction by stabilizing an ensemble of conformations that are incompatible with binding. The foci here are those interactions where disorder is a dominant feature of at least one component of the targeted interaction.

6.1. Targeting Structure to Enhance Disorder

Given the abundance and importance of IDPR-mediated protein−protein interactions in the interactome, as protein−protein interactions have gained traction as viable targets, interactions where one binding partner is disordered have naturally emerged. The widely known MDM2−p53 interaction, which has been discussed in detail in section 5.1, provides a paradigm for this type of inhibition. To briefly recapitulate, the binding of p53 by the MDM2 protein prevents the tumor suppressor activity and targets p53 for ubiquitination and degradation. The binding site on p53 is localized to residues 15−29, which are intrinsically disordered and fold into a helix to bind in a groove on MDM2.[145] Structure-based drug design was employed to guide discovery of molecules (nutlins) that would bind to the MDM2 groove and displace the p53 helix,

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX−XXX

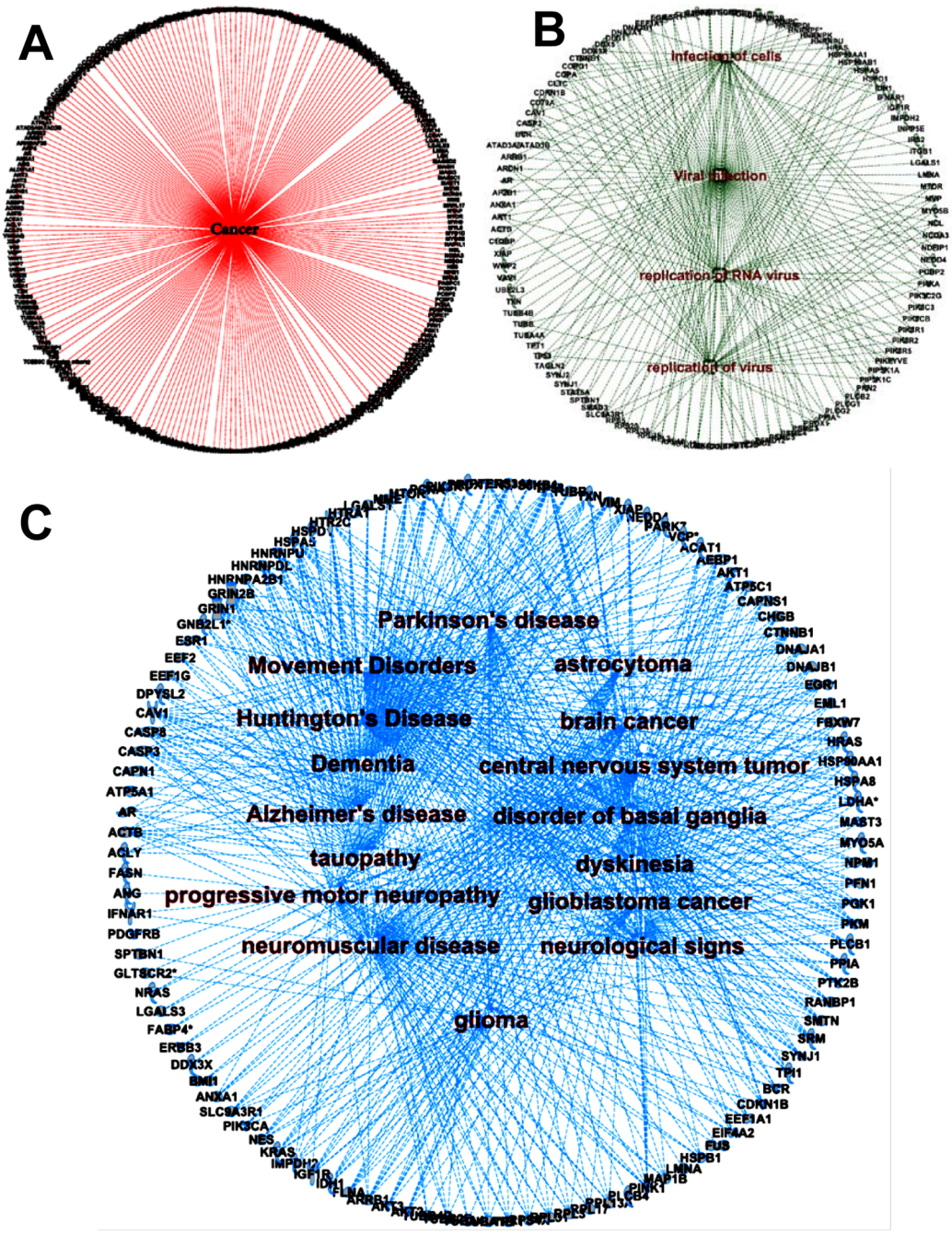


Figure 18. Top enriched networks of PTEN. UniProt IDs of all primary PTEN-interacting proteins were imported into the Ingenuity Pathway Analysis (IPA) Software (Ingenuity Systems, Mountain View, CA; www.ingenuity.com). By use of the Core Analysis function, the top diseases were identified, and individual networks were visualized by use of the "Display Networks" option. (A) Network depicting all primary PTEN-interacting proteins involved in various types of cancers (P value 5.00×10^{-25}). (B) Network depicting all primary PTEN-interacting proteins involved in infectious diseases (P value 4.97×10^{-21}). (C) Network depicting all primary PTEN-interacting proteins involved in neurological diseases (P value 4.57×10^{-17}).

releasing it back to its disordered state.[215,216] Further development led to the compound RG7112, which entered clinical trials.[285] Cheng et al.[286] laid out a generalized scheme for inhibiting interactions of this type where recognition of an IDPR mediates a crucial interaction. Their system of using bioinformatics to predict important IDPR recognition sequences and then to use these as a starting point to find interactions and design mimetics allows for access into pathways and interactions that, unlike p53−MDM2, are not fully characterized structurally or even fully mapped.[286]

The use of a small molecule to displace a disordered binding segment can also take place intramolecularly. Kaposi's sarcoma-associated herpesvirus protease (KSHV Pr) undergoes dimerization, with concomitant folding of helices 5 and 6, to form the active protease and is inactive as the monomer.[287] Craik and co-workers[288] therefore targeted the dimer interface in a search for inhibitors. As with p53, a helix mimetic strategy

dx.doi.org/10.1021/cr400713r | Chem. Rev. XXXX, XXX, XXX−XXX

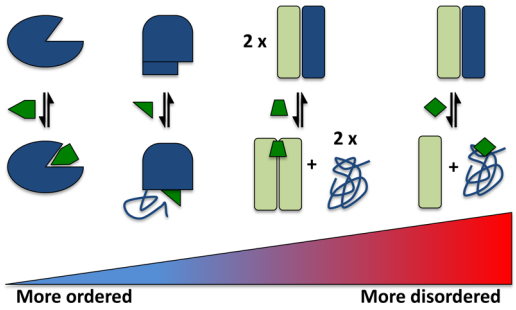


Figure 19. Schematic examples of the continuum of disorder in the binding of small molecules (green) to proteins. On the left is a highly ordered protein (e.g., an enzyme) binding to a small molecule, and on the far right is a small molecule binding directly to a disordered protein and stabilizing the disordered ensemble. In the center are examples of small molecules inducing protein disorder by preventing binding of a disordered region to an ordered partner.

led to an initial hit that was improved to yield DD2. DD2 binds to KSHV Pr at structured surface residues and, in doing so, stabilizes the monomeric form in which helices 5 and 6 remain disordered.[288,289]

The c-Myc oncoprotein (hereafter Myc) is dysregulated in the majority of human cancers, and its expression is correlated with cell proliferation and poor prognosis.[290−292] Myc is a bHLHZip transcription factor that is disordered in its monomeric form and undergoes coupled binding and folding in the HLH-Zip region when binding to its obligate heterodimerization partner Max. While Myc does not homodimerize, Max does, and Jiang et al.[293] targeted Myc activity by finding molecules that could stabilize structured Max dimers, leaving Myc in its monomeric, disordered state. Using virtual ligand screening, they identified several molecules with potential specificity for Max−Max dimers over Myc−Max dimers. Experimental validation and follow-up showed compound NSC13728 to be a specific stabilizer of Max dimers that kept Myc disordered and inhibited Myc-driven transcription and transformation.

Several groups have targeted inhibition of bZip proteins (which are unstructured in the zipper region before dimerization and unstructured in the basic region before DNA binding) by binding to the dimeric form while inhibiting the DNA binding function and leaving the basic region largely disordered. Shiozawa and co-workers[294,295] targeted the basic regions of the c-Fos/c-Jun dimer. Using a cyclic peptide inhibitor as a starting point, they developed a pharmacophore model and synthesized several small molecules capable of inhibiting coupled binding and folding of the basic region. Vinson and co-workers[296] identified a stibonic acid inhibitor (NSC13778) of C/EBP that recognized residues at the junction between the basic region and the leucine zipper and stabilized the dimer while inhibiting DNA binding and folding of the basic region. Further screening of stibonic acids identified additional very active but promiscuous inhibitors of bZip (CREB) and bHLHZip (USF and Mitf) proteins.[297,298] Rudenko and co-workers[299] screened 54 498 molecules for inhibitors of ΔFosB DNA binding. They identified two, C2 and C6, that were active in follow-up assays. Testing of structural analogues produced several related, active compounds. Interestingly, C2 did not perturb the degree of helicity in the protein, whereas C6 increased it. Although specific binding sites have not been identified, the authors' model involves C2 binding to the disordered basic region and maintaining it in a disordered state incompatible with DNA binding. The C6 compound increases helicity, inducing folding upon binding but in a conformation that includes the small molecule that is also incompatible with DNA binding.

6.2. Direct Targeting of Small Molecules to IDPs

The examples of direct targeting of IDPs by small molecules (though still few) have clustered around two targets: transcription factors and amyloidogenic proteins. Here, direct targeting means the binding of small molecules to an IDPR involved in a PPI such that the resulting small-molecule complex remains substantially dynamic, and this stabilized ensemble is incompatible with binding to the regular protein partner or partners. This interaction can be seen as a recapitulation, with a small-molecule partner, of features found in IDPR complexes with other proteins, such as the ability to adopt different structures in complex with different partners and the ability to form "fuzzy" complexes in which an ensemble of conformations is responsible for binding.[300,301]

An early indication of the capacity of small molecules to recognize IDPRs came from the ability of peptide sequences displayed on phage to bind small molecules.[302] Two separate experiments found paclitaxel binding peptides that were mapped onto known proteins (though not previously known targets of paclitaxel).[302,303] Peptides were also found in screens that used camptothecin, NK109, and trimannoside.[304−306] A difficulty that emerges when an IDPR is targeted in the context of a PPI is determining the actual small-molecule binding site on a disordered target. In structured targets the location and extent of the binding sites is generally clear, even if the energetic contributions of each component are not.[307] For IDPRs the specific binding site is often unknown at first, potentially occurring at any point along the length of the disordered interaction sequence. Although the binding sites for IDPRs have so far been localized to short linear sequences,[308,309] in disordered proteins adjacent sequences can modulate the conformational ensemble[310] and long-range contacts still occur.[311] These types of interactions may further complicate the precise determination of small-molecule binding sites in IDPs.

6.2.1. Targeting Transcription Factors. As indicated above, Myc is an important target in cancer therapy. With its obligate heterodimerization partner Max, Myc binds to its cognate E-box sequence in promoter regions.[312,313] The accumulation of Myc in cells was recently shown to cause accumulation at additional, lower-affinity binding sites in active promoter and enhancer regions and thereby increase overall transcript output.[314] This activity is in line with IDPs generally in that their levels are normally tightly regulated and their overexpression leads to negative effects at a greater rate than the proteome generally.[19,69] These characteristics also indicate that the activity of IDPs is likely to be modulated by small molecules that affect their interaction energy.

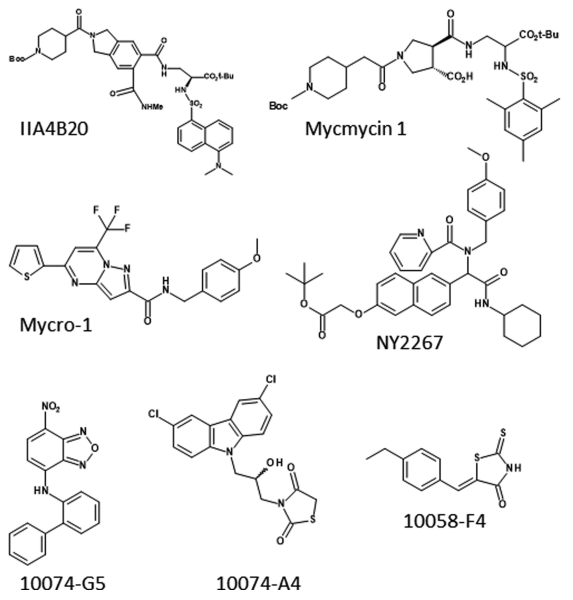
In the direct targeting of Myc, the Myc−Max dimer does not provide a simple binding site for small molecules;[315] moreover, binding to the dimer could stabilize this undesired dimeric species.[293,308,316] The desired disruption of the Myc−Max dimer would drive the proteins back into their monomeric, disordered state to which, to satisfy thermodynamic requirements, the small molecules must bind. The direct targeting of intrinsically disordered sequences was not the motivation for the original discovery of various Myc−Max inhibitors; however,

dx.doi.org/10.1021/cr400713r | Chem. Rev. XXXX, XXX, XXX−XXX

the system became a hub of work on targeting IDPRs as multiple groups brought various methods and strategies to bear.

The first Myc binders were found in a screen of a 7000-compound peptidomimetic combinatorial library by Boger, Vogt, and co-workers.[317] Two molecules, IIA4B20 and IIA6B17, were active in both the initial dimer disruption assay and follow-up assays (Scheme 1). While the molecules

Scheme 1



inhibited Myc-driven cell growth, they also inhibited growth driven by Jun. This report demonstrated Myc as a potentially tractable target, although the mechanism and site of binding were unknown. A second generation of inhibitors was synthesized from the same combinatorial substituents but with a smaller pyrrolidine versus isoindoline core.[318] Two resulting molecules, Mycmycin-1 and -2, showed improved Myc-induced transformation inhibition while also eliminating inhibition of Jun-induced transformation, clearly showing an amenability of these systems to medicinal chemistry optimization.

A screen for Myc inhibition using a "credit-card" library with a naphthyl core designed to mimic the largely flat, hydrophobic interface of some protein−protein interactions yielded two molecules (NY2267 and NY2280) that were active in all of the assays.[319] However, these molecules showed transcriptional inhibition of both Myc and Jun, but not NF-κB, in luciferase reporter assays. Berg and co-workers[320] found a pyrazolopyrimidine-based Myc inhibitor in a 17 000-compound diversity library. The compound and close analogues, Mycro1 and 2, inhibited Myc−Max dimerization and E-box binding and slowed growth of several transformed cell lines but not Myc-independent PC-12 cells. In a focused screen of over 1400 pyrazolopyrimidine compounds, they were able to eliminate AP-1-dependent transcription inhibition present in the original Mycro compounds while maintaining cell growth inhibition selectivity.[321] In a screen against a limited library (1990 compounds), Mo and Henriksson[322] found a small molecule (Myra-A) that interfered with Myc activity but did not affect dimerization. Instead, it interfered with DNA binding of Myc−Max. However, the compound also interfered with Max−Max and Mnt−Max binding (though not the Ebox binding of USF). A recent screen of cytotoxic drugs that affect the Myc pathway found that doxorubicin had a similar DNA-binding inhibition profile as that of Myra-A.[323] The protein target (or targets) in the Myc system, the binding site, and the structural consequences for Myra-A and doxorubicin are not yet known.

A diversity library (10 000 compounds) was screened against Myc activity by Prochownik and co-workers[324] in a yeast two-hybrid system. The initial screen included an HLH protein dimer, Id2−E47, with this second interaction designed as a first check on compound specificity. The screen yielded seven compounds that inhibited Myc−Max, 10 that inhibited Id2−E47, and 28 that inhibited both. Thus, this screen illustrated clearly what was also seen in the other screens: an initial hit is likely to have substantial activity against several targets. To further define the specificity of the initial hits, they screened the specific binders and a subset of the dual Myc−Max and Id2−E47 inhibitors against a panel of 32 other bHLH, bHLHZip, and bZip transcription factors. The specific inhibitors averaged strong inhibition of less than one other protein pair, whereas dual-specific compounds averaged more than three. Structure−activity relationship (SAR) studies have been conducted on two of the initial Myc inhibitors, 10058-F4 and 10074-G5, and in both cases compounds with moderate gains in activity were identified.[325,326] On the basis of SAR data for 10058-F4, a three-dimensional pharmacophore model was developed and used to predict nine new potential binders, of which four had experimentally demonstrated inhibition of Myc−Max.[327] While the new compounds did not show improved affinity, their diverse chemotypes demonstrated both that pharmacophore modeling was viable in targeting IDPRs and that lead-hopping was readily achievable. In the 10074-G5 screen, a carboxylic acid derivative, JY-3-094, was found with improved in vitro inhibition activity but poor activity in cell assays. Further work showed that a strategy of delivering the molecule in an esterified form increased cell penetration and, through the activity of intracellular esterases, liberated the active molecule.[328] As with other small molecules, absorption, distribution, metabolism, and excretion (ADME) properties are a crucial but difficult component of moving forward in preclinical testing.[329] Pharmacokinetics studies on 10058-F4 and 10074-G5 demonstrate that both compounds had relatively short plasma half-lives and poor tumor penetration in a xenograft mouse model.[330,331]

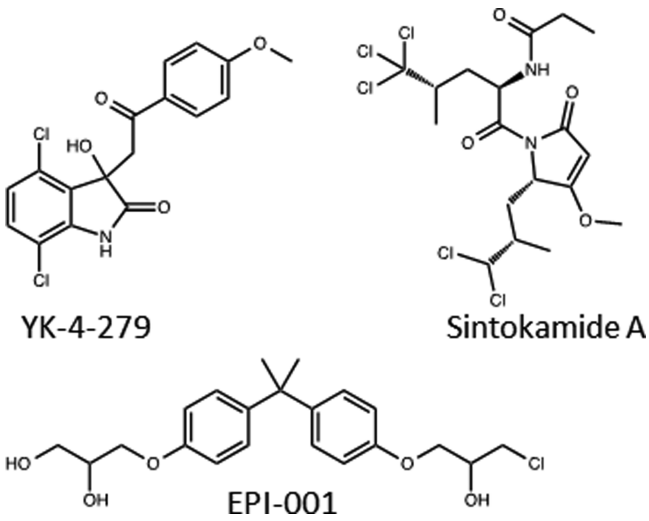
Dimerization of Myc with Max buries 3200 Å^2 of surface. Despite this large contact area, a variety of compounds can disrupt the complex and return the proteins to a disordered state. To better understand small-molecule IDPR binding, the binding sites for two of the molecules from the screen by Prochownik were determined. The entire contact area between Myc and Max was a potential binding site, and binding any portion could shift conformational ensembles away from α-helical and disrupt dimerization. Random mutants were generated in the Myc bHLHZip region, and intrinsic fluorescence of two compounds, 10074-G5 and 10058-F4, was used to probe binding to these mutants and to truncations.[308] The molecules bound to two distinct regions within the Myc bHLHZip, and NMR and circular dichroism (CD) with short peptides indicated that binding was localized to short linear stretches that remained dynamic even in the complex. By use of competition experiments, the binding sites of four other molecules from the Prochownik screen were determined, with three binding the 10058-F4 site and one binding at the 10074-G5 site. The final molecule, 10074-A4,

dx.doi.org/10.1021/cr400713r | Chem. Rev. XXXX, XXX, XXX−XXX

was found to bind independently at a third site just C-terminal to the 10074-G5 site.[309] The presence of three distinct sites within the 85 amino acid domain indicated the potential prevalence of small-molecule binding sites within intrinsically disordered recognition regions, while the presence of non-conserved residues within the binding sites provided a potential basis for specificity.[309] In an interesting application, drift-time ion mobility mass spectroscopy (DT IM-MS) was used to probe binding of 10058-F4 to Myc−Max leucine zippers.[332] The presence of the compound caused the peptides to shift to a more compact ensemble, consistent with loss of helical zipper conformation, although the direct interaction of compound with Myc was not observed.

Outside of their DNA-binding domains, transcription factors typically contain substantial regions of disorder, particularly in their transactivation domains.[333,334] These regions present attractive potential targets for the modulation of transcription factor activity, but the disordered and often repetitive nature of the sequence in these regions makes identification of specific binding sites difficult. The oncoprotein EWS-FLI1 is a translocation-generated fusion protein containing the EWS activation domain in the N-terminus and the ETS DNA-binding domain in the C-terminus.[335] Outside of the ETS domain, the protein is substantially disordered and depends on sequence composition, not order, for transactivation.[35,336,337] By use of surface plasmon resonance (SPR), potential inhibitors were screened for binding to immobilized EWS-FLI1. An initial hit was improved to yield the compound YK-4-279 (Scheme 2).[338] The small molecule bound in a manner

Scheme 2



that interfered with the functionally important RNA helicase A binding to EWS-FLI1, although this activity was not part of the original screen. Resolution of the YK-4-279 enantiomers demonstrated enantiospecific EWS-FLI1 activity both in vitro and in vivo.[339] However, the exact nature of the interaction of YK-4-279 with EWS-FLI1 has not been determined, as the binding site (or potentially sites) has not yet been defined.

Androgen receptor (AR) is the primary target for prostate cancer chemotherapy. Suppression of AR-driven transcription leads to tumor regression; however, relapse can occur in the form of castration-resistant prostate cancer, CRPC, for which treatments are generally ineffective.[340] The AR has three domains: N-terminal domain (NTD), DNA-binding domain (DBD), and ligand-binding domain (LBD). The DBD and the LBD are predominantly helical, structured domains, whereas the NTD is disordered with some residual helical structure.[65] Sadar, Andersen, and co-workers[341] isolated chlorinated peptide sintokamides from marine sponge (*Dysidea* sp.) and screened them for AR inhibition. Although the specific binding site was not identified, sintokamide A bound to the disordered NTD and disrupted interactions required for transcriptional control. Extracts from the marine sponge *Niphates digitalis* yielded glycerol ethers containing an enone moiety that were shown to have activity against AR. Of these, niphatenone B was the most active natural product and was found to covalently bind the NTD via the Michael acceptor enone.[342]

During the screening of other marine sponge compounds, EPI-001 was found to inhibit AR NTD function.[343] Interestingly, this and related compounds were apparently anthropogenic, as they are derivatives of the widely used diepoxide cross-linker bisphenol A diglycidic ether (BADGE). Previously BADGE, the dihydroxy derivative (BADGE·2H_2O), and the chlorohydroxy derivative (BADGE·2HCl) were shown to have AR antagonist activity by Satoh et al.,[344] with BADGE·2HCl having the greatest antagonist effect of the three. Sadar and co-workers[343] tested a range of BADGE derivatives and found EPI-001 (BADGE·HCl·H_2O) to be the most potent, while BADGE·H_2O (their compound 185-9-1) showed no activity. Satoh et al.[344] were using the entire AR, versus the isolated NTD, and did not test EPI-001/BADGE·HCl·H_2O. The binding of EPI-001 to the NTD reduced interaction with CBP and AR-dependent cell proliferation but did not affect AR-independent proliferation. Significantly, EPI-001 inhibited constitutively active AR that lacked a LBD, indicating its potential against CRPS. A series of EPI-001 derivatives, including the four stereoisomers, showed only mild differences in activity, with one isomer, EPI-002, showing good tumor reduction while maintaining animal weight in a mouse model.[345] EPI-001 and its specific isomers bound covalently to the NTD. The authors proposed an initial binding event, followed by formation of an epoxide (catalyzed by functional groups from the protein) and then covalent binding via a NTD protein nucleophile.[345] A variety of FDA-approved covalent inhibitors exist that target noncatalytic sites in proteins.[346] Given the modest affinity of small molecules targeted to IDPRs (low micromolar), targeted covalent inhibitors provide a potential means of increasing potency when specificity can be maintained.

6.2.2. Targeting Amyloid Forming Proteins. A variety of IDPs (such as α-synuclein, Aβ42, and τ) are implicated in neurodegenerative disease and fibrillar pathologies. Oligomers of these proteins, not just fibrils, may be toxic, and binding to monomers to inhibit their aggregation may be a viable treatment strategy.[347] Consequently, there have been a large number of small molecules reported to inhibit either fibril formation or oligomerization.[348,349] However, many molecules do not yet have well-defined targets (sequence or oligomeric state) or modes of action. Further complicating the analysis is the ability of small molecules to display inhibition of amyloid polymerization through action as nanoaggregates or detergent-like entities.[350,351]

Previously, the molecules fenofibrate and flurbiprofen were identified as γ-secretase modulators that, instead of being directed to the enzyme, bound to the substrate amyloid precursor protein (APP) and modulated the length of the Aβ peptide that was proteolytically produced.[352] Subsequent work

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX−XXX

with flurbiprofen and sulindac sulfide showed that the ability of the compounds to modulate cleavage by γ-secretase was dependent on the sequence of the substrate and in particular a GxxxG motif.[353,354] However, other researchers applying NMR techniques saw no signs of specific interactions between Aβ peptides and flurbiprofen or sulindac sulfide. Rather, they saw aggregate formation and postulated that compound or peptide aggregates influenced the observed enzyme activity.[355,356] Whether these molecules truly represent substrate targeted inhibitors[357] is unclear. Even when there is an explicit, specific issue of aggregation and addressing the concern is part of the work, the results from different laboratories still may not agree. The ability of nanoaggregates (soluble colloids) to bind to and sequester target proteins is a problem in the broad context of small-molecule discovery.[350] Given the generally micromolar affinity of small molecules binding to IDPs, the potential for IDPs to adopt various binding conformations, and the nontrivial nature of monitoring complexes between IDPs and small molecules, IDPs may be particularly susceptible to binding by promiscuous inhibitors.

Despite the number of reported fibril inhibitors, well-characterized examples of noncovalent specific binders to monomeric disordered species are lacking. Instead, in characterized systems, a common mode of action for molecules binding to protein monomers has emerged whereby they redirect proteins toward aggregates that are not on the pathway to fibril formation.[358] The polyphenol (−)-epigallocatechin gallate (EGCG) has shown antifibril activity against a variety of targets.[349] Well-controlled experiments with EGCG and Aβ42, Aβ40, α-synuclein, IAPP, and Sup35NM have all demonstrated the ability of EGCG to intercept aggregation-prone sequences and direct them into spherical aggregates that are not prone to fibril formation.[358,359] A similar phenomenon was observed for the pthalocyanine tetrasulfate (PcTS) interaction with τ. A range of techniques [small-angle X-ray scattering (SAXS), NMR, and electron paramagnetic resonance (EPR)] was used to establish that PcTS interacted with aromatic residues of τ and directed the protein into compact oligomers that differ from toxic β-structures containing small oligomers.[360] A screen of β_2-microglobulin fibril inhibitors found that the antibiotic rifamycin SV bound unfolded protein monomers and shifted them onto a path toward spherical aggregates rather than fibrils.[361] The ability of carnosine to inhibit amyloid growth without modifying the conformational features of Aβ42 was shown by scanning force microscopy, circular dichroism, and thioflavin T fluorescence.[362] In the same study, molecular dynamics (MD) simulation analysis revealed that carnosine interacted transiently with monomeric Aβ42 by salt bridges with charged side chains and by van der Waals contacts with residues in and around the central hydrophobic cluster $_{17}$LVFFA$_{21}$.[362] In NMR experiments, carnosine was shown to interfere with the local propensity of the peptide to form backbone hydrogen bonds close to the central hydrophobic cluster (residues E22, S26, and N27).[362] On the basis of these and other observations, the authors concluded that although carnosine did not form stable contacts with Aβ, this molecule was able to block the pathway toward toxic aggregates via perturbation of the hydrogen-bond network near residues that play some key roles in Aβ fibrillation.[362]

6.3. Simulations, Predictions, and Specificity of Small-Molecule IDP Interactions

The techniques of MD simulation and of binding-site prediction hold tremendous promise, especially as they become increasingly refined for the particular challenges of highly dynamic complexes. The discovery of small-molecule binders of IDPRs has relied on library screening, followed, in some cases, by defining binding regions. While many of these screens have yielded hits from relatively small sets of compounds, identifying binding sites and understanding the specificity of interactions is labor-intensive and has lagged for most complexes. Increased computational guidance for identifying binding sites and understanding the specific nature of these interactions will be a substantial advance.

Many protein−protein interactions are mediated by relatively short sequence segments binding to a protein partner interface, and these sequences exist disproportionately in disordered regions.[363] While initial efforts involved identifying known interaction regions, various groups have used combinations of properties such as sequence composition, conservation, physiochemical properties, and intrachain energy calculations to generate predictors such as ANCHOR, SLiMSearch, and MoRFpred (among others) for finding disordered, linear protein interaction regions.[364−366] Although additional, well-defined small-molecule IDPR binding sites are needed to generate sufficient data sets, correlations between the characteristics of those sites and the predicted protein interaction sites likely exist and will begin to be exploited.[300,309]

The availability of defined binding sites for small-molecule IDPR complexes has made possible simulations of these interactions that can be correlated with experimental observations. Sufficient conformational sampling is challenging in IDP systems; Michel and Cuchillo[367] used a bias-exchange metadynamics technique to improve conformational sampling of the peptide Myc$_{402-412}$ alone and in complex with 10058-F4. In the simulations, both free and bound forms of the peptide existed as heterogeneous ensembles of conformations, and 10058-F4 interacted with a range of different conformations with no clearly dominant structure. This result correlated with experimental indications that the peptide small-molecule complexes remained dynamic.[308] Using a replica-exchange molecular dynamics (REMD) approach with implicit solvent, Liu and co-workers[368] simulated another binding interaction on Myc, 10074-A4 binding to Myc$_{370-409}$. As in the simulation with 10058-F4, the peptide remained dynamic even in the bound form. Further, 10074-A4 was seen as highly mobile along the site, prompting the authors to call it a "ligand cloud". In comparison with Myc sequences from outside the binding site (the Myc leucine zipper) 10074-A4 interactions with its binding site could clearly be distinguished from interactions outside the binding site.

In an example of the potential for combined computational and experimental work, Herrera et al.[369] used computational models of α-synuclein (AS)−dopamine interactions to direct production of specific mutants. Dopamine has been reported to inhibit AS fibrillization, with some debate as to whether a covalent adduct from the oxidized form is the relevant species[370] or not.[371] Herrera et al. used NMR along with limited MD simulations and dopamine docking to investigate binding modes on AS ensembles. Independent of starting structures, they saw interactions with the AS C-terminal region as anticipated from experiments. They also observed an electrostatic contribution from the more N-terminal residue

dx.doi.org/10.1021/cr400713r | Chem. Rev. XXXX, XXX, XXX−XXX

Glu83. Testing dopamine binding to Ala point mutants indicated that the C-terminal residues were relatively insensitive to mutations, likely indicative of dynamic binding interactions there, while mutation of Glu83 eliminated dopamine's antifibrillization effects, consistent with the presence of electrostatic interactions seen in the models.

Using REMD, Zhu et al.[372] generated an ensemble of Aβ42 structures from which they selected clusters of conformations in order to identify potential binding pockets in the monomer. Forty-five clusters were generated with populations from 0.05% to 2%, and their correlation with experimental NMR data was validated against chemical shift, ^3J coupling, and residual dipolar coupling. Ten small organic fragments ($\leq$6 heavy atoms) were docked to clusters to identify potential hot spots that could be combined into binding regions. The authors point out the significance of this method in reducing computational costs while maintaining coverage of conformations in the docking.

Experiment and simulation indicate that small-molecule IDPR interactions can be both specific and dynamic, and multiple different chemotypes may bind to the same IDPR. Reconciling the specificity of binding of a given molecule with the ability to bind multiple molecules can be seen as analogous to an IDP's ability to interact with multiple, specific protein partners.[373] This mode of binding often involves adopting different conformations to interact with different targets. Unlike a structured sequence, a disordered sequence may present a broad array of viable interfaces displaying its chemical groups in various constellations. Therefore, the number of pairwise potential combinations with a given set of chemotypes could be greatly increased along with the probability of finding a match. While the probability to find a specific interaction is increased, the ability to bind nonspecifically is also increased. The large number of physiologically relevant IDPRs represents an attractive pool of potential targets for those looking to inhibit them, as well as a deep reservoir of potential off-target interactions, especially for molecules reaching micromolar concentrations.

7. CONCLUDING REMARKS

Many proteins involved in the pathogenesis of human diseases contain intrinsically disordered regions. This conclusion follows from the numerous illustrative examples of well-characterized disease-related proteins (e.g., p53, α-synuclein, PTEN), as well as from results of comprehensive bioinformatics studies. High degrees of association between intrinsic disorder and diseases are determined by the unique structural and functional properties of IDPs and IDPRs, which are frequently found to serve as major cellular regulators, recognizers, and signal transducers. Their normal functionality is tightly controlled and modulated via a wide spectrum of PTMs and alternative splicing. 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 point mutations, altered splicing, chromosomal translocations, or exposure to internal or external toxins. Formation of pathologic conformations can also be triggered by impaired PTMs, oxidative damage, increased degradation propensities, impaired trafficking, or loss of binding partners. All these factors can act independently, additively, or synergistically to generate pathogenic transitions in proteins.

Being extensively involved in crucial protein−protein interactions and being intimately linked to the pathogenesis of various human diseases, IDPs/IDPRs represent novel, attractive drug targets. While ordered proteins, such as enzymes and cell-surface receptors, can be targeted relatively easily with small molecules mimicking natural substrates, a new emerging field is the development of protein−protein interaction inhibitors. This is where intrinsic disorder is coming into play because disorder-based interactions are common in signaling, regulation, and recognition and are abundantly found in pathological interactions associated with various human diseases. Although targeting disorder-based protein−protein interactions is a relatively new field, it is clearly moving beyond the proof-of-concept stage. The variety of currently available approaches for finding small molecules affecting functions of IDPs/IDPRs is promising because it clearly shows that intrinsic disorder-based interactions are druggable and can be modulated by small molecules.

AUTHOR INFORMATION

Corresponding Authors

*Phone 1-813-974-5816; fax 1-813-974-7357; e-mail vuversky@health.usf.edu.
*Phone +44-1223-267290; e-mail acj2@mrc-lmb.cam.ac.uk.

Notes

The authors declare no competing financial interest.

Biographies

Vladimir Uversky obtained a Ph.D. in biophysics from Moscow Institute of Physics and Technology (1991) and a D.Sc. in biophysics from Institute of Experimental and Theoretical Biophysics, Russian Academy of Sciences (1998). He spent his early career working on protein folding at the Institute of Protein Research and the Institute for Biological Instrumentation (Russian Academy of Sciences). In 1998, he moved to the University of California Santa Cruz to work on protein folding, misfolding, and protein intrinsic disorder. In 2004, he moved to the Center for Computational Biology and Bioinformatics, Indiana University−Purdue University Indianapolis, to work on intrinsically disordered proteins. Since 2010, he has been with the Department of Molecular Biology at the University of South Florida.

dx.doi.org/10.1021/cr400713r | Chem. Rev. XXXX, XXX, XXX−XXX

Vrushank Davé obtained his bachelor's degree in biochemistry (1985) from St. Xavier's College, India. After obtaining a Department of Biotechnology Fellowship from the Government of India, he completed his master's degree in biotechnology (1987) from the MS University of Baroda, India. Later, he obtained the esteemed GATE (Graduate Aptitude Test in Engineering) fellowship from IIT (Indian Institute of Technology) and the Department of Biotechnology graduate fellowship, which allowed him to complete his graduate studies in molecular biology from the Jawaharlal Nehru University, New Delhi, India, in 1994. During his graduate studies he developed a keen interest in allosteric binding of transcription factors to different DNA sites required for target gene selection in early molecular events in making of an embryo. He maintained this interest and subsequently published a series of papers at the University of Cincinnati, utilizing genetic, biochemical, and biophysical techniques including NMR that highlighted the role of protein flexibility in achieving a high degree of DNA recognition specificity and fidelity during tissue specific transcription during development. In 2001, Dr. Davé joined the faculty as an instructor at The Perinatal Institute at the University of Cincinnati Children's Research Foundation, where he continued to work on protein−protein interactions forming multiprotein complexes driving tissue-specific gene regulation during development and organogenesis. He identified a host of novel transcription factors involved in lung branching morphogenesis and maturation. Since 2011, he has been a tenure-track assistant professor with the Department of Pathology and Cell Biology at the University of South Florida, where he collaborates with Dr. Uversky and Dr. Haura in the area of systems-level structural informatics focused on critical tumor suppressors and oncogenes. Over the last 10 years Dr. Davé has been invited to present his work at Cold Spring Harbor Laboratory, Keystone Symposia, FASEB conferences, and American Heart Association and American Thoracic Society international meetings. He also serves on the editorial board of the *American Journal of Physiology: Lung Cellular and Molecular Physiology*.

Lilia Iakoucheva obtained a B.S. in genetics from Kiev State University (Kiev, Ukraine) and a Ph.D. in molecular biology and immunology from the Institute of Immunology (Moscow, Russia). After completing postdoctoral training in protein biochemistry at Pacific Northwest National Laboratory (Richland, WA), she joined the group of Professor Keith Dunker to study intrinsically disordered proteins. During that time and with her active participation, the group made a series of fundamental discoveries about disordered proteins, including their involvement in cell signaling and cancer and the importance of disorder for posttranslational modifications and for interactions with other proteins and ligands. In 2003, Dr. Iakoucheva joined the Rockefeller University (New York) as a research assistant professor, where she continued to investigate functional properties of disordered proteins, at the same time gradually shifting her interests into disease-oriented studies. Rapid advancement in disease gene discovery in the postgenomic era opened new avenues and opportunities for more detailed investigation of protein interaction networks and pathways underlying many human diseases. Dr. Iakoucheva became especially interested in the molecular basis of psychiatric diseases, which she began to explore using systems biology-oriented approaches. She joined the Psychiatry Department of the University of California San Diego (La Jolla, CA) as an assistant professor in 2010, where she continues to apply her experience in protein structure and protein−protein interaction analyses toward investigation of neuropsychiatric disorders. Dr. Iakoucheva has been the principal investigator on research grants from the National Science Foundation, National Cancer Institute, National Institute of Child Health and Human Development, and National Institute of Mental Health. She also serves as an Associate Editor of the PLOS Computational Biology journal.

Prerna Malaney obtained her bachelor's degree in pharmacy (B.Pharm.) from the University of Mumbai, India, in 2011. In August 2011, she joined Dr. Vrushank Davé's lab as a graduate student in the Department of Pathology and Cell Biology at the University of South Florida. She is currently a third-year Ph.D. student in medical sciences. Her dissertation project primarily focuses on structure−function relationships of the tumor suppressor gene PTEN. Ms. Malaney is using several bioinformatic and experimental approaches at the systems level to define the role of intrinsic disorder in PTEN function. As a proponent of network medicine, she believes that a global systems-level approach allows for a more holistic understanding of complex signaling cascades in pathological conditions. Her work in network medicine was recognized by two travel grants awarded by the Graduate School at the University of South Florida and by the International Conference on Intelligent Biology and Medicine (ICIBM) held at Vanderbilt University in 2013. Ms. Malaney currently has four peer-reviewed publications and hopes to develop her career as a systems pharmacologist.

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX−XXX

Steven J. Metallo obtained his Ph.D. in chemistry from Yale University (with Alanna Schepartz), where he studied protein−DNA interactions. He then conducted postdoctoral research at Harvard University where he studied multivalency and surface chemistry with George M. Whitesides. In 2001, he joined the faculty at Georgetown University in the Department of Chemistry where he is currently an associate professor. He is an associate member of the Lombardi Comprehensive Cancer Center and a founding member of Georgetown's Institute for Soft Matter Synthesis and Metrology. Currently his work focuses on the binding of disordered proteins, in particular the binding of small molecules to disordered protein regions.

Dr. Ravi Ramesh Pathak obtained his Ph.D. in plant molecular biology from G.G.S. Indraprastha University in New Delhi, India, in 2010. His Ph.D. work was focused on whole-genome profiling of plant model systems by high-density microarray and systems biology approaches. His research work was the only one selected for an oral presentation from India across all disciplines at the prestigious Cold Spring Harbor Asia Conference held in 2010. Since then, he has worked extensively on cancer biology at the Mount Sinai Medical Center, New York, and the University of South Florida. Dr. Pathak has successfully combined elements of bioinformatics and systems biology with classical transcriptional biology and complex mouse models to successfully identify novel mechanisms in lung cancers. His most recent publication in the *Journal of Biochemistry* in 2013 was recommended by the faculty of F1000 prime as being among the most significant findings in the field. He has routinely published his research in peer-reviewed international journals and has contributed a number of book chapters in leading publications. In addition to this, his work has also been presented at a number of prestigious national and international conferences. Dr. Pathak has served as assistant editor on the editorial board of *Physiology and Molecular Biology of Plants*, an international plant science journal published by Springer.

Andreas C. Joerger studied chemistry at the University of Freiburg (Germany), specializing in biochemistry, and received his Ph.D. in 2000 for crystallographic and functional studies on fuculose-1-phosphate aldolase in the group of Professor Georg E. Schulz. He then did a postdoc with Professor Sir Alan Fersht at the MRC Centre for Protein Engineering in Cambridge (United Kingdom), working on protein design and structural studies of the tumor suppressor p53. He determined the first crystal structures of p53 cancer mutants and discovered a unique druggable surface crevice in one of these mutants. He plays a leading role in the ongoing efforts to develop potent small-molecule stabilizers of mutant p53. Dr. Joerger is currently a senior scientist at the MRC Laboratory of Molecular Biology, Cambridge. His main areas of research are the structural biology and evolution of the p53 pathway and related proteins, molecular interactions, the structural basis of disease mutations, and p53 drug discovery.

ABBREVIATIONS

Aβ	amyloid-β
AD	Alzheimer's disease
ALL	acute lymphoblastic leukemia
ALK	anaplastic lymphoma kinase
AML	acute myelogenous leukemia
APC	adenomatous polyposis coli
APP	amyloid precursor protein
AR	androgen receptor
AS	α-synuclein
ATM	ataxia telangiectasia mutated
BADGE	diepoxide cross-linker bisphenol A diglycidic ether
CHIP	carboxyl terminus of Hsc70 interacting protein
CML	chronic myelogenous leukemia
CRC	colorectal cancer
CTA	cancer/testis antigen
CTR	C-terminal regulatory domain
CVD	cardiovascular disease
C-tail	C-terminal tail
DBD	DNA-binding domain
DM	disease-associated mutations
D → O	disorder to order transition
EGCG	polyphenol (−)-epigallocatechin gallate
EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
EGR1	early growth-response protein 1
ELM	eukaryotic linear motif
EVI1	ecotropic virus integration site 1 protein
EWS	Ewing sarcoma
GLTSCR2	glioma tumor suppressor candidate region gene-2 protein
GSK3-β	glycogen synthase kinase 3-β

dx.doi.org/10.1021/cr400713r | *Chem. Rev.* XXXX, XXX, XXX−XXX

HAUSP	herpesvirus-associated ubiquitin-specific protease
HAT	histone acetyltransferase
HDAC	histone deacetylase
Htt	huntingtin protein
IDP	intrinsically disordered protein
IDPR	intrinsically disordered protein region
ID1	inhibitor of DNA binding 1
IGF	insulin-like growth factor
JAK2	Janus tyrosine kinase 2
KSVH Pr	Kaposi's sarcoma-associated herpesvirus protease
LBD	ligand-binding domain
MBD	methyl DNA binding domain
MECP2	methyl-CpG-binding protein 2
miRNA	microRNA
MKK4	mitogen-activated protein kinase kinase 4
MoRE	molecular recognition element
MoRF	molecular recognition feature
MVP	Major-vault protein
NCBD	nuclear coactivator binding domain
NEDD4	neural precursor cell-expressed developmentally downregulated 4−1
NES	neutral evolutionary substitutions
NPM	nucleolar phosphoprotein nucleophosmin
NTD	N-terminal domain
OR	ordered region
O → D	order to disorder transition
PBM	phosphatidylinositol (4,5)-bisphosphate binding module
PcTS	phthalocyanine tetrasulfate
PD	phosphatase domain
PGE2	prostaglandin E2
PIP2	phosphatidylinositol 4,5-bisphosphate
PIP3	phosphatidylinositol 3,4,5-trisphosphate
PI3K	phosphoinositide 3-kinase
PLCb	phospholipase Cb
Plk3	polo-like kinase 3
PNUTS	protein phosphatase 1 nuclear targeting subunit
Poly	polymorphism
PPARγ	peroxisome proliferator-activated receptor γ
PPI	protein−protein interaction
PR	proline-rich region
PTEN	phosphatase and tensin homologue deleted on chromosome 10
PTHS	PTEN tumor hamartoma syndrome
PTM	posttranslational modification
RDC	residual dipolar coupling
REMD	replica-exchange molecular dynamics
ROCK	RhoA-associated protein kinase
RTK	receptor tyrosine kinases
RTT	Rett syndrome
SALL4	Sal-like protein 4
SAM	sterile α motif
SMA	spinal muscular atrophy
S6K	S6 kinase
snRNP	small nuclear ribonucleic particle
SPRY2	sprouty homologue 2
TAD	transactivation domain
TEL	ETS translocation variant 6
TET	tetramerization domain
TFG	TRK-fused gene
TGF-β	transforming growth factor β
TI	transactivation inhibitory region
WT	wild type
WWP2	WW domain containing protein 2
XIAP	X-linked inhibitor of apoptosis protein

REFERENCES

(1) Wright, P. E.; Dyson, H. J. *J. Mol. Biol.* **1999**, *293*, 321.
(2) Uversky, V. N.; Gillespie, J. R.; Fink, A. L. *Proteins: Struct., Funct., Genet.* **2000**, *41*, 415.
(3) Dunker, A. K.; Lawson, J. D.; Brown, C. J.; Williams, R. M.; Romero, P.; Oh, J. S.; Oldfield, C. J.; Campen, A. M.; Ratliff, C. M.; Hipps, K. W.; Ausio, J.; Nissen, M. S.; Reeves, R.; Kang, C.; Kissinger, C. R.; Bailey, R. W.; Griswold, M. D.; Chiu, W.; Garner, E. C.; Obradovic, Z. *J. Mol. Graphics Modell.* **2001**, *19*, 26.
(4) Tompa, P. *Trends Biochem. Sci.* **2002**, *27*, 527.
(5) Uversky, V. N.; Dunker, A. K. *Biochim. Biophys. Acta* **2010**, *1804*, 1231.
(6) Choy, W. Y.; Forman-Kay, J. D. *J. Mol. Biol.* **2001**, *308*, 1011.
(7) Huang, A.; Stultz, C. M. *PLoS Comput. Biol.* **2008**, *4*, No. e1000155.
(8) Romero, P. R.; Zaidi, S.; Fang, Y. Y.; Uversky, V. N.; Radivojac, P.; Oldfield, C. J.; Cortese, M. S.; Sickmeier, M.; LeGall, T.; Obradovic, Z.; Dunker, A. K. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8390.
(9) Uversky, V. N. *Curr. Pharm. Des.* **2013**, *19*, 4191.
(10) Oldfield, C. J.; Meng, J.; Yang, J. Y.; Yang, M. Q.; Uversky, V. N.; Dunker, A. K. *BMC Genomics* **2008**, *9* (Suppl.1), S1.
(11) Dunker, A. K.; Obradovic, Z.; Romero, P.; Garner, E. C.; Brown, C. J. *Genome Inf. Ser.* **2000**, *11*, 161.
(12) Ward, J. J.; Sodhi, J. S.; McGuffin, L. J.; Buxton, B. F.; Jones, D. T. *J. Mol. Biol.* **2004**, *337*, 635.
(13) Tokuriki, N.; Oldfield, C. J.; Uversky, V. N.; Berezovsky, I. N.; Tawfik, D. S. *Trends Biochem. Sci.* **2009**, *34*, 53.
(14) Xue, B.; Williams, R. W.; Oldfield, C. J.; Dunker, A. K.; Uversky, V. N. *BMC Syst. Biol.* **2010**, *4* (Suppl. 1), S1.
(15) Uversky, V. N. *J. Biomed. Biotechnol.* **2010**, *2010*, No. 568068.
(16) Xue, B.; Dunker, A. K.; Uversky, V. N. *J. Biomol. Struct. Dyn.* **2012**, *30*, 137.
(17) Oldfield, C. J.; Cheng, Y.; Cortese, M. S.; Brown, C. J.; Uversky, V. N.; Dunker, A. K. *Biochemistry* **2005**, *44*, 1989.
(18) Dunker, A. K.; Brown, C. J.; Lawson, J. D.; Iakoucheva, L. M.; Obradovic, Z. *Biochemistry* **2002**, *41*, 6573.
(19) Gsponer, J.; Futschik, M. E.; Teichmann, S. A.; Babu, M. M. *Science* **2008**, *322*, 1365.
(20) Iakoucheva, L. M.; Radivojac, P.; Brown, C. J.; O'Connor, T. R.; Sikes, J. G.; Obradovic, Z.; Dunker, A. K. *Nucleic Acids Res.* **2004**, *32*, 1037.
(21) Tompa, P. *FEBS Lett.* **2005**, *579*, 3346.
(22) Xie, H.; Vucetic, S.; Iakoucheva, L. M.; Oldfield, C. J.; Dunker, A. K.; Obradovic, Z.; Uversky, V. N. *J. Proteome Res.* **2007**, *6*, 1917.
(23) Grimmler, M.; Wang, Y.; Mund, T.; Cilensek, Z.; Keidel, E. M.; Waddell, M. B.; Jakel, H.; Kullmann, M.; Kriwacki, R. W.; Hengst, L. *Cell* **2007**, *128*, 269.
(24) Uversky, V. N.; Dunker, A. K. *Science* **2008**, *322*, 1340.
(25) Uversky, V. N.; Oldfield, C. J.; Midic, U.; Xie, H.; Xue, B.; Vucetic, S.; Iakoucheva, L. M.; Obradovic, Z.; Dunker, A. K. *BMC Genomics* **2009**, *10* (Suppl.1), S7.
(26) Uversky, V. N. *Expert Rev. Proteomics* **2010**, *7*, 543.
(27) Uversky, V. N. *Expert Opin. Drug Discovery* **2012**, *7*, 475.
(28) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. *Annu. Rev. Biophys.* **2008**, *37*, 215.
(29) Midic, U.; Oldfield, C. J.; Dunker, A. K.; Obradovic, Z.; Uversky, V. N. *BMC Genomics* **2009**, *10* (Suppl1), S12.
(30) Midic, U.; Oldfield, C. J.; Dunker, A. K.; Obradovic, Z.; Uversky, V. N. *Protein Pept. Lett.* **2009**, *16*, 1533.
(31) Uversky, V. N. *Front. Biosci., Landmark Ed.* **2009**, *14*, 5188.
(32) Iakoucheva, L. M.; Brown, C. J.; Lawson, J. D.; Obradovic, Z.; Dunker, A. K. *J. Mol. Biol.* **2002**, *323*, 573.
(33) Wells, M.; Tidow, H.; Rutherford, T. J.; Markwick, P.; Jensen, M. R.; Mylonas, E.; Svergun, D. I.; Blackledge, M.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 5762.

(34) Mark, W. Y.; Liao, J. C.; Lu, Y.; Ayed, A.; Laister, R.; Szymczyna, B.; Chakrabartty, A.; Arrowsmith, C. H. *J. Mol. Biol.* **2005**, *345*, 275.
(35) Ng, K. P.; Potikyan, G.; Savene, R. O.; Denny, C. T.; Uversky, V. N.; Lee, K. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 479.
(36) Uversky, V. N.; Roman, A.; Oldfield, C. J.; Dunker, A. K. *J. Proteome Res.* **2006**, *5*, 1829.
(37) Malaney, P.; Pathak, R. R.; Xue, B.; Uversky, V. N.; Dave, V. *Sci. Rep.* **2013**, *3*, 2035.
(38) Rajagopalan, K.; Mooney, S. M.; Parekh, N.; Getzenberg, R. H.; Kulkarni, P. *J. Cell. Biochem.* **2011**, *112*, 3256.
(39) Uversky, V. N. *J. Biomol. Struct. Dyn.* **2003**, *21*, 211.
(40) Cheng, Y.; LeGall, T.; Oldfield, C. J.; Dunker, A. K.; Uversky, V. N. *Biochemistry* **2006**, *45*, 10448.
(41) Casu, F.; Duggan, B. M.; Hennig, M. *Biophys. J.* **2013**, *105*, 1004.
(42) Baker, J. M.; Hudson, R. P.; Kanelis, V.; Choy, W. Y.; Thibodeau, P. H.; Thomas, P. J.; Forman-Kay, J. D. *Nat. Struct. Mol. Biol.* **2007**, *14*, 738.
(43) Goh, K. I.; Cusick, M. E.; Valle, D.; Childs, B.; Vidal, M.; Barabasi, A. L. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 8685.
(44) Uversky, V. N. *Curr. Alzheimer Res.* **2008**, *5*, 260.
(45) Mohan, A.; Sullivan, W. J., Jr.; Radivojac, P.; Dunker, A. K.; Uversky, V. N. *Mol. BioSyst.* **2008**, *4*, 328.
(46) Xie, H.; Vucetic, S.; Iakoucheva, L. M.; Oldfield, C. J.; Dunker, A. K.; Uversky, V. N.; Obradovic, Z. *J. Proteome Res.* **2007**, *6*, 1882.
(47) Vucetic, S.; Xie, H.; Iakoucheva, L. M.; Oldfield, C. J.; Dunker, A. K.; Obradovic, Z.; Uversky, V. N. *J. Proteome Res.* **2007**, *6*, 1899.
(48) Hegyi, H.; Buday, L.; Tompa, P. *PLoS Comput. Biol.* **2009**, *5*, No. e1000552.
(49) Roccato, E.; Pagliardini, S.; Cleris, L.; Canevari, S.; Formelli, F.; Pierotti, M. A.; Greco, A. *Oncogene* **2003**, *22*, 807.
(50) Lacronique, V.; Boureux, A.; Valle, V. D.; Poirel, H.; Quang, C. T.; Mauchauffe, M.; Berthou, C.; Lessard, M.; Berger, R.; Ghysdael, J.; Bernard, O. A. *Science* **1997**, *278*, 1309.
(51) Bischof, D.; Pulford, K.; Mason, D. Y.; Morris, S. W. *Mol. Cell. Biol.* **1997**, *17*, 2312.
(52) Stamm, S.; Ben-Ari, S.; Rafalska, I.; Tang, Y.; Zhang, Z.; Toiber, D.; Thanaraj, T. A.; Soreq, H. *Gene* **2005**, *344*, 1.
(53) Weatheritt, R. J.; Gibson, T. J. *Trends Biochem. Sci.* **2012**, *37*, 333.
(54) Buljan, M.; Chalancon, G.; Eustermann, S.; Wagner, G. P.; Fuxreiter, M.; Bateman, A.; Babu, M. M. *Mol. Cell* **2012**, *46*, 871.
(55) Buljan, M.; Chalancon, G.; Dunker, A. K.; Bateman, A.; Balaji, S.; Fuxreiter, M.; Babu, M. M. *Curr. Opin. Struct. Biol.* **2013**, *23*, 443.
(56) David, C. J.; Manley, J. L. *Genes Dev.* **2010**, *24*, 2343.
(57) Chen, G. L.; Miller, G. M. *PLoS One* **2013**, *8*, e62217.
(58) Pal, S.; Gupta, R.; Davuluri, R. V. *Pharmacol. Ther.* **2012**, *136*, 283.
(59) Lee, H. K.; Choi, Y. S.; Park, Y. A.; Jeong, S. *Cancer Res.* **2006**, *66*, 10560.
(60) Sette, C. *Int. J. Cell. Biol.* **2013**, No. 458727.
(61) Adamia, S.; Haibe-Kains, B.; Pilarski, P. M.; Bar-Natan, M.; Pevzner, S.; Avet-Loiseau, H.; Lode, L.; Verselis, S.; Fox, E. A.; Burke, J.; Galinsky, I.; Dagogo-Jack, I.; Wadleigh, M.; Steensma, D. P.; Motyckova, G.; Deangelo, D. J.; Quackenbush, J.; Stone, R. M.; Griffin, J. D. *Clin. Cancer Res.* **2014**, *20*, 1135.
(62) Peng, Z.; Xue, B.; Kurgan, L.; Uversky, V. N. *Cell Death Differ.* **2013**, *20*, 1257.
(63) Korneta, I.; Bujnicki, J. M. *PLoS Comput. Biol.* **2012**, *8*, No. e1002641.
(64) Coelho Ribeiro, M. d. L.; Espinosa, J.; Islam, S.; Martinez, O.; Thanki, J. J.; Mazariegos, S.; Nguyen, T.; Larina, M.; Xue, B.; Uversky, V. N. *PeerJ* **2013**, *1*, No. e2.
(65) McEwan, I. J. *Mol. BioSyst.* **2012**, *8*, 82.
(66) Hu, Y.; Dobi, A.; Sreenath, T.; Cook, C.; Tadase, A. Y.; Ravindranath, L.; Cullen, J.; Furusato, B.; Chen, Y.; Thangapazham, R. L.; Mohamed, A.; Sun, C.; Sesterhenn, I. A.; McLeod, D. G.; Petrovics, G.; Srivastava, S. *Clin. Cancer Res.* **2008**, *14*, 4719.
(67) Shkreta, L.; Bell, B.; Revil, T.; Venables, J. P.; Prinos, P.; Elela, S. A.; Chabot, B. *Cancer Treat. Res.* **2013**, *158*, 41.
(68) Philips, A. V.; Cooper, T. A. *Cell. Mol. Life Sci.* **2000**, *57*, 235.
(69) Vavouri, T.; Semple, J. I.; Garcia-Verdugo, R.; Lehner, B. *Cell* **2009**, *138*, 198.
(70) Babu, M. M.; van der Lee, R.; de Groot, N. S.; Gsponer, J. *Curr. Opin. Struct. Biol.* **2011**, *21*, 432.
(71) Markiv, A.; Rambaruth, N. D.; Dwek, M. V. *Curr. Opin. Pharmacol.* **2012**, *12*, 408.
(72) Park, J. J.; Lee, M. *Gut Liver* **2013**, *7*, 629.
(73) Peng, Z.; Mizianty, M. J.; Xue, B.; Kurgan, L.; Uversky, V. N. *Mol. BioSyst.* **2012**, *8*, 1886.
(74) Campbell, M. J.; Turner, B. M. *Adv. Exp. Med. Biol.* **2013**, *754*, 81.
(75) Gargalionis, A. N.; Piperi, C.; Adamopoulos, C.; Papavassiliou, A. G. *Int. J. Biochem. Cell Biol.* **2012**, *44*, 1276.
(76) Chatterjee, S.; Senapati, P.; Kundu, T. K. *Essays Biochem.* **2012**, *52*, 93.
(77) Gioeli, D.; Paschal, B. M. *Mol. Cell. Endocrinol.* **2012**, *352*, 70.
(78) Di Gennaro, E.; Bruzzese, F.; Caraglia, M.; Abruzzese, A.; Budillon, A. *Amino Acids* **2004**, *26*, 435.
(79) Moumne, L.; Betuing, S.; Caboche, J. *Front. Neurol.* **2013**, *4*, 127.
(80) Ehrnhoefer, D. E.; Sutton, L.; Hayden, M. R. *Neuroscientist* **2011**, *17*, 475.
(81) Lee, L.; Sakurai, M.; Matsuzaki, S.; Arancio, O.; Fraser, P. *Neuromol. Med.* **2013**, *15*, 720.
(82) Hernandez, F.; Avila, J. *Cell. Mol. Life Sci.* **2007**, *64*, 2219.
(83) Wang, J. Z.; Xia, Y. Y.; Grundke-Iqbal, I.; Iqbal, K. *J. Alzheimer's Dis.* **2013**, *33* (Suppl. 1), S123.
(84) Stefl, S.; Nishi, H.; Petukh, M.; Panchenko, A. R.; Alexov, E. *J. Mol. Biol.* **2013**, *425*, 3919.
(85) Yue, P.; Moult, J. *J. Mol. Biol.* **2006**, *356*, 1263.
(86) Kumar, P.; Henikoff, S.; Ng, P. C. *Nat. Protoc.* **2009**, *4*, 1073.
(87) Adzhubei, I. A.; Schmidt, S.; Peshkin, L.; Ramensky, V. E.; Gerasimova, A.; Bork, P.; Kondrashov, A. S.; Sunyaev, S. R. *Nat. Methods* **2010**, *7*, 248.
(88) Teng, S.; Srivastava, A. K.; Wang, L. *BMC Genomics* **2010**, *11* (Suppl.2), S5.
(89) Bromberg, Y.; Rost, B. *Nucleic Acids Res.* **2007**, *35*, 3823.
(90) Thusberg, J.; Olatubosun, A.; Vihinen, M. *Hum. Mutat.* **2011**, *32*, 358.
(91) Li, B.; Krishnan, V. G.; Mort, M. E.; Xin, F.; Kamati, K. K.; Cooper, D. N.; Mooney, S. D.; Radivojac, P. *Bioinformatics* **2009**, *25*, 2744.
(92) Vacic, V.; Iakoucheva, L. M. *Mol. BioSyst.* **2012**, *8*, 27.
(93) Vacic, V.; Markwick, P. R.; Oldfield, C. J.; Zhao, X.; Haynes, C.; Uversky, V. N.; Iakoucheva, L. M. *PLoS Comput. Biol.* **2012**, *8*, No. e1002709.
(94) Li, S.; Iakoucheva, L. M.; Mooney, S. D.; Radivojac, P. *Pac. Symp. Biocomput.* **2010**, 337.
(95) Radivojac, P.; Vacic, V.; Haynes, C.; Cocklin, R. R.; Mohan, A.; Heyen, J. W.; Goebl, M. G.; Iakoucheva, L. M. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 365.
(96) Petroski, M. D. *BMC Biochem.* **2008**, *9* (Suppl. 1), S7.
(97) Tan, C. S.; Bodenmiller, B.; Pasculescu, A.; Jovanovic, M.; Hengartner, M. O.; Jorgensen, C.; Bader, G. D.; Aebersold, R.; Pawson, T.; Linding, R. *Sci. Signal.* **2009**, *2*, ra39.
(98) Radivojac, P.; Baenziger, P. H.; Kann, M. G.; Mort, M. E.; Hahn, M. W.; Mooney, S. D. *Bioinformatics* **2008**, *24*, i241.
(99) Oldfield, C. J.; Cheng, Y.; Cortese, M. S.; Romero, P.; Uversky, V. N.; Dunker, A. K. *Biochemistry* **2005**, *44*, 12454.
(100) Fong, J. H.; Shoemaker, B. A.; Garbuzynskiy, S. O.; Lobanov, M. Y.; Galzitskaya, O. V.; Panchenko, A. R. *PLoS Comput. Biol.* **2009**, *5*, No. e1000316.
(101) Gunasekaran, K.; Tsai, C. J.; Kumar, S.; Zanuy, D.; Nussinov, R. *Trends Biochem. Sci.* **2003**, *28*, 81.
(102) Cumberworth, A.; Lamour, G.; Babu, M. M.; Gsponer, J. *Biochem. J.* **2013**, *454*, 361.
(103) Zhong, Q.; Simonis, N.; Li, Q. R.; Charloteaux, B.; Heuze, F.; Klitgord, N.; Tam, S.; Yu, H.; Venkatesan, K.; Mou, D.; Swearingen,

V.; Yildirim, M. A.; Yan, H.; Dricot, A.; Szeto, D.; Lin, C.; Hao, T.; Fan, C.; Milstein, S.; Dupuy, D.; Brasseur, R.; Hill, D. E.; Cusick, M. E.; Vidal, M. *Mol. Syst. Biol.* **2009**, *5*, 321.
(104) Wang, X.; Wei, X.; Thijssen, B.; Das, J.; Lipkin, S. M.; Yu, H. *Nat. Biotechnol.* **2012**, *30*, 159.
(105) Uversky, V. N.; Iakoucheva, L. M.; Dunker, A. K. In *eLS (Encyclopedia of Life Sciences)*; John Wiley & Sons: Chichester, U.K., 2012.
(106) Huang, S. F.; Liu, H. P.; Li, L. H.; Ku, Y. C.; Fu, Y. N.; Tsai, H. Y.; Chen, Y. T.; Lin, Y. F.; Chang, W. C.; Kuo, H. P.; Wu, Y. C.; Chen, Y. R.; Tsai, S. F. *Clin. Cancer Res.* **2004**, *10*, 8195.
(107) Shan, Y.; Eastwood, M. P.; Zhang, X.; Kim, E. T.; Arkhipov, A.; Dror, R. O.; Jumper, J.; Kuriyan, J.; Shaw, D. E. *Cell* **2012**, *149*, 860.
(108) Minde, D. P.; Anvarian, Z.; Rudiger, S. G.; Maurice, M. M. *Mol. Cancer* **2011**, *10*, 101.
(109) Amir, R. E.; Van den Veyver, I. B.; Wan, M.; Tran, C. Q.; Francke, U.; Zoghbi, H. Y. *Nat. Genet.* **1999**, *23*, 185.
(110) Hansen, J. C.; Wexler, B. B.; Rogers, D. J.; Hite, K. C.; Panchenko, T.; Ajith, S.; Black, B. E. *J. Biol. Chem.* **2011**, *286*, 18938.
(111) Bezsonova, I.; Singer, A.; Choy, W. Y.; Tollinger, M.; Forman-Kay, J. D. *Biochemistry* **2005**, *44*, 15550.
(112) Lee, S.; Mahler, B.; Toward, J.; Jones, B.; Wyatt, K.; Dong, L.; Wistow, G.; Wu, Z. *J. Mol. Biol.* **2010**, *399*, 320.
(113) Lane, D. P. *Nature* **1992**, *358*, 15.
(114) Vousden, K. H.; Prives, C. *Cell* **2009**, *137*, 413.
(115) Vousden, K. H.; Ryan, K. M. *Nat. Rev. Cancer* **2009**, *9*, 691.
(116) Maddocks, O. D.; Berkers, C. R.; Mason, S. M.; Zheng, L.; Blyth, K.; Gottlieb, E.; Vousden, K. H. *Nature* **2013**, *493*, 542.
(117) Dotsch, V.; Bernassola, F.; Coutandin, D.; Candi, E.; Melino, G. *Cold Spring Harbor Perspect. Biol.* **2010**, *2*, No. a004887.
(118) Bourdon, J. C. *Curr. Pharm. Biotechnol.* **2007**, *8*, 332.
(119) Joerger, A. C.; Fersht, A. R. *Annu. Rev. Biochem.* **2008**, *77*, 557.
(120) Bell, S.; Klein, C.; Muller, L.; Hansen, S.; Buchner, J. *J. Mol. Biol.* **2002**, *322*, 917.
(121) Cho, Y.; Gorina, S.; Jeffrey, P. D.; Pavletich, N. P. *Science* **1994**, *265*, 346.
(122) Chen, Y.; Dey, R.; Chen, L. *Structure* **2010**, *18*, 246.
(123) Kitayner, M.; Rozenberg, H.; Rohs, R.; Suad, O.; Rabinovich, D.; Honig, B.; Shakked, Z. *Nat. Struct. Mol. Biol.* **2010**, *17*, 423.
(124) Meek, D. W.; Anderson, C. W. *Cold Spring Harbor Perspect. Biol.* **2009**, *1*, No. a000950.
(125) Dai, C.; Gu, W. *Trends Mol. Med.* **2010**, *16*, 528.
(126) Deutsch, G. B.; Zielonka, E. M.; Coutandin, D.; Weber, T. A.; Schafer, B.; Hannewald, J.; Luh, L. M.; Durst, F. G.; Ibrahim, M.; Hoffmann, J.; Niesen, F. H.; Senturk, A.; Kunkel, H.; Brutschy, B.; Schleiff, E.; Knapp, S.; Acker-Palmer, A.; Grez, M.; McKeon, F.; Dotsch, V. *Cell* **2011**, *144*, 566.
(127) Chi, S. W.; Ayed, A.; Arrowsmith, C. H. *EMBO J.* **1999**, *18*, 4438.
(128) Sathyamurthy, A.; Freund, S. M.; Johnson, C. M.; Allen, M. D.; Bycroft, M. *FEBS J.* **2011**, *278*, 2680.
(129) Serber, Z.; Lai, H. C.; Yang, A.; Ou, H. D.; Sigal, M. S.; Kelly, A. E.; Darimont, B. D.; Duijf, P. H.; Van Bokhoven, H.; McKeon, F.; Dotsch, V. *Mol. Cell. Biol.* **2002**, *22*, 8601.
(130) Tidow, H.; Melero, R.; Mylonas, E.; Freund, S. M.; Grossmann, J. G.; Carazo, J. M.; Svergun, D. I.; Valle, M.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 12324.
(131) Melero, R.; Rajagopalan, S.; Lazaro, M.; Joerger, A. C.; Brandt, T.; Veprintsev, D. B.; Lasso, G.; Gil, D.; Scheres, S. H.; Carazo, J. M.; Fersht, A. R.; Valle, M. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 557.
(132) Bista, M.; Freund, S. M.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 15752.
(133) Lv, L.; Wang, P.; Zhou, X.; Sun, B. *Tumour Biol.* **2013**, *34*, 1451.
(134) Bergamaschi, D.; Samuels, Y.; Sullivan, A.; Zvelebil, M.; Breyssens, H.; Bisso, A.; Del Sal, G.; Syed, N.; Smith, P.; Gasco, M.; Crook, T.; Lu, X. *Nat. Genet.* **2006**, *38*, 1133.
(135) Wade, M.; Wang, Y. V.; Wahl, G. M. *Trends Cell. Biol.* **2010**, *20*, 299.
(136) Stad, R.; Little, N. A.; Xirodimas, D. P.; Frenk, R.; van der Eb, A. J.; Lane, D. P.; Saville, M. K.; Jochemsen, A. G. *EMBO Rep.* **2001**, *2*, 1029.
(137) Linares, L. K.; Hengstermann, A.; Ciechanover, A.; Muller, S.; Scheffner, M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12009.
(138) Linke, K.; Mace, P. D.; Smith, C. A.; Vaux, D. L.; Silke, J.; Day, C. L. *Cell Death Differ.* **2008**, *15*, 841.
(139) Tang, Y.; Zhao, W.; Chen, Y.; Zhao, Y.; Gu, W. *Cell* **2008**, *133*, 612.
(140) Jenkins, L. M. M.; Durell, S. R.; Mazur, S. J.; Appella, E. *Carcinogenesis* **2012**, *33*, 1441.
(141) Toledo, F.; Wahl, G. M. *Nat. Rev. Cancer* **2006**, *6*, 909.
(142) Teufel, D. P.; Freund, S. M.; Bycroft, M.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 7009.
(143) Lee, C. W.; Arai, M.; Martinez-Yamout, M. A.; Dyson, H. J.; Wright, P. E. *Biochemistry* **2009**, *48*, 2115.
(144) Venot, C.; Maratrat, M.; Sierra, V.; Conseiller, E.; Debussche, L. *Oncogene* **1999**, *18*, 2405.
(145) Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P. *Science* **1996**, *274*, 948.
(146) Popowicz, G. M.; Czarna, A.; Holak, T. A. *Cell Cycle* **2008**, *7*, 2441.
(147) Feng, H.; Jenkins, L. M.; Durell, S. R.; Hayashi, R.; Mazur, S. J.; Cherry, S.; Tropea, J. E.; Miller, M.; Wlodawer, A.; Appella, E.; Bai, Y. *Structure* **2009**, *17*, 202.
(148) Di Lello, P.; Jenkins, L. M.; Jones, T. N.; Nguyen, B. D.; Hara, T.; Yamaguchi, H.; Dikeakos, J. D.; Appella, E.; Legault, P.; Omichinski, J. G. *Mol. Cell* **2006**, *22*, 731.
(149) Bochkareva, E.; Kaustov, L.; Ayed, A.; Yi, G. S.; Lu, Y.; Pineda-Lucena, A.; Liao, J. C.; Okorokov, A. L.; Milner, J.; Arrowsmith, C. H.; Bochkarev, A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15412.
(150) Vise, P. D.; Baral, B.; Latos, A. J.; Daughdrill, G. W. *Nucleic Acids Res.* **2005**, *33*, 2061.
(151) Rajagopalan, S.; Andreeva, A.; Rutherford, T. J.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 8587.
(152) Rajagopalan, S.; Andreeva, A.; Teufel, D. P.; Freund, S. M.; Fersht, A. R. *J. Biol. Chem.* **2009**, *284*, 21728.
(153) Wong, T. S.; Rajagopalan, S.; Townsley, F. M.; Freund, S. M.; Petrovich, M.; Loakes, D.; Fersht, A. R. *Nucleic Acids Res.* **2009**, *37*, 568.
(154) Lee, C. W.; Martinez-Yamout, M. A.; Dyson, H. J.; Wright, P. E. *Biochemistry* **2010**, *49*, 9964.
(155) Nishi, H.; Fong, J. H.; Chang, C.; Teichmann, S. A.; Panchenko, A. R. *Mol. BioSyst.* **2013**, *9*, 1620.
(156) Sakaguchi, K.; Saito, S.; Higashimoto, Y.; Roy, S.; Anderson, C. W.; Appella, E. *J. Biol. Chem.* **2000**, *275*, 9278.
(157) Schon, O.; Friedler, A.; Bycroft, M.; Freund, S. M.; Fersht, A. R. *J. Mol. Biol.* **2002**, *323*, 491.
(158) Teufel, D. P.; Bycroft, M.; Fersht, A. R. *Oncogene* **2009**, *28*, 2112.
(159) Brown, C. J.; Srinivasan, D.; Jun, L. H.; Coomber, D.; Verma, C. S.; Lane, D. P. *Cell Cycle* **2008**, *7*, 608.
(160) Lee, C. W.; Ferreon, J. C.; Ferreon, A. C.; Arai, M.; Wright, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 19290.
(161) Burge, S.; Teufel, D. P.; Townsley, F. M.; Freund, S. M.; Bycroft, M.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 3142.
(162) Lee, W.; Harvey, T. S.; Yin, Y.; Yau, P.; Litchfield, D.; Arrowsmith, C. H. *Nat. Struct. Biol.* **1994**, *1*, 877.
(163) Clore, G. M.; Ernst, J.; Clubb, R.; Omichinski, J. G.; Kennedy, W. M.; Sakaguchi, K.; Appella, E.; Gronenborn, A. M. *Nat. Struct. Biol.* **1995**, *2*, 321.
(164) Jeffrey, P. D.; Gorina, S.; Pavletich, N. P. *Science* **1995**, *267*, 1498.
(165) Stommel, J. M.; Marchenko, N. D.; Jimenez, G. S.; Moll, U. M.; Hope, T. J.; Wahl, G. M. *EMBO J.* **1999**, *18*, 1660.
(166) Nicholls, C. D.; McLure, K. G.; Shields, M. A.; Lee, P. W. *J. Biol. Chem.* **2002**, *277*, 12937.
(167) Rajagopalan, S.; Huang, F.; Fersht, A. R. *Nucleic Acids Res.* **2011**, *39*, 2294.

(168) Joerger, A. C.; Rajagopalan, S.; Natan, E.; Veprintsev, D. B.; Robinson, C. V.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 17705.
(169) Coutandin, D.; Lohr, F.; Niesen, F. H.; Ikeya, T.; Weber, T. A.; Schafer, B.; Zielonka, E. M.; Bullock, A. N.; Yang, A.; Guntert, P.; Knapp, S.; McKeon, F.; Ou, H. D.; Dotsch, V. *Cell Death Differ.* **2009**, *16*, 1582.
(170) Natan, E.; Joerger, A. C. *J. Mol. Biol.* **2012**, *415*, 503.
(171) Mateu, M. G.; Sanchez Del Pino, M. M.; Fersht, A. R. *Nat. Struct. Biol.* **1999**, *6*, 191.
(172) Gaglia, G.; Guan, Y.; Shah, J. V.; Lahav, G. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 15497.
(173) van Dieck, J.; Brandt, T.; Teufel, D. P.; Veprintsev, D. B.; Joerger, A. C.; Fersht, A. R. *Oncogene* **2010**, *29*, 2024.
(174) van Dieck, J.; Fernandez-Fernandez, M. R.; Veprintsev, D. B.; Fersht, A. R. *J. Biol. Chem.* **2009**, *284*, 13804.
(175) Rajagopalan, S.; Jaulent, A. M.; Wells, M.; Veprintsev, D. B.; Fersht, A. R. *Nucleic Acids Res.* **2008**, *36*, 5983.
(176) Schumacher, B.; Mondry, J.; Thiel, P.; Weyand, M.; Ottmann, C. *FEBS Lett.* **2010**, *584*, 1443.
(177) Sakaguchi, K.; Sakamoto, H.; Lewis, M. S.; Anderson, C. W.; Erickson, J. W.; Appella, E.; Xie, D. *Biochemistry* **1997**, *36*, 10117.
(178) Retzlaff, M.; Rohrberg, J.; Kupper, N. J.; Lagleder, S.; Bepperling, A.; Manzenrieder, F.; Peschek, J.; Kessler, H.; Buchner, J. *J. Mol. Biol.* **2013**, *425*, 144.
(179) Suh, E. K.; Yang, A.; Kettenbach, A.; Bamberger, C.; Michaelis, A. H.; Zhu, Z.; Elvin, J. A.; Bronson, R. T.; Crum, C. P.; McKeon, F. *Nature* **2006**, *444*, 624.
(180) Luh, L. M.; Kehrloesser, S.; Deutsch, G. B.; Gebel, J.; Coutandin, D.; Schafer, B.; Agostini, M.; Melino, G.; Dotsch, V. *Cell Death Differ.* **2013**, *20*, 1008.
(181) Joerger, A. C.; Fersht, A. R. *Cold Spring Harbor Perspect. Biol.* **2010**, *2*, No. a000919.
(182) Andreeva, A. *Methods Mol. Biol.* **2012**, *857*, 1.
(183) Mezei, M. *Protein Eng.* **1998**, *11*, 411.
(184) Tidow, H.; Lauber, T.; Vitzithum, K.; Sommerhoff, C. P.; Rosch, P.; Marx, U. C. *Biochemistry* **2004**, *43*, 11238.
(185) Rustandi, R. R.; Baldisseri, D. M.; Weber, D. J. *Nat. Struct. Biol.* **2000**, *7*, 570.
(186) Lowe, E. D.; Tews, I.; Cheng, K. Y.; Brown, N. R.; Gul, S.; Noble, M. E.; Gamblin, S. J.; Johnson, L. N. *Biochemistry* **2002**, *41*, 15625.
(187) Avalos, J. L.; Celic, I.; Muhammad, S.; Cosgrove, M. S.; Boeke, J. D.; Wolberger, C. *Mol. Cell* **2002**, *10*, 523.
(188) Mujtaba, S.; He, Y.; Zeng, L.; Yan, S.; Plotnikova, O.; Sachchidanand; Sanchez, R.; Zeleznik-Le, N. J.; Ronai, Z.; Zhou, M. M. *Mol. Cell* **2004**, *13*, 251.
(189) Friedler, A.; Veprintsev, D. B.; Freund, S. M.; von Glos, K. I.; Fersht, A. R. *Structure* **2005**, *13*, 629.
(190) Weinberg, R. L.; Freund, S. M.; Veprintsev, D. B.; Bycroft, M.; Fersht, A. R. *J. Mol. Biol.* **2004**, *342*, 801.
(191) Tafvizi, A.; Huang, F.; Fersht, A. R.; Mirny, L. A.; van Oijen, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 563.
(192) McKinney, K.; Mattia, M.; Gottifredi, V.; Prives, C. *Mol. Cell* **2004**, *16*, 413.
(193) King, N.; Westbrook, M. J.; Young, S. L.; Kuo, A.; Abedin, M.; Chapman, J.; Fairclough, S.; Hellsten, U.; Isogai, Y.; Letunic, I.; Marr, M.; Pincus, D.; Putnam, N.; Rokas, A.; Wright, K. J.; Zuzow, R.; Dirks, W.; Good, M.; Goodstein, D.; Lemons, D.; Li, W.; Lyons, J. B.; Morris, A.; Nichols, S.; Richter, D. J.; Salamov, A.; Sequencing, J. G.; Bork, P.; Lim, W. A.; Manning, G.; Miller, W. T.; McGinnis, W.; Shapiro, H.; Tjian, R.; Grigoriev, I. V.; Rokhsar, D. *Nature* **2008**, *451*, 783.
(194) Lane, D. P.; Cheok, C. F.; Brown, C.; Madhumalar, A.; Ghadessy, F. J.; Verma, C. *Cell Cycle* **2010**, *9*, 540.
(195) Pankow, S.; Bamberger, C. *PLoS One* **2007**, *2*, No. e782.
(196) Belyi, V. A.; Ak, P.; Markert, E.; Wang, H.; Hu, W.; Puzio-Kuter, A.; Levine, A. J. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a001198.
(197) Lane, D. P.; Madhumalar, A.; Lee, A. P.; Tay, B. H.; Verma, C.; Brenner, S.; Venkatesh, B. *Cell Cycle* **2011**, *10*, 4272.
(198) Nedelcu, A. M.; Tan, C. *Dev. Genes Evol.* **2007**, *217*, 801.
(199) Brandt, T.; Petrovich, M.; Joerger, A. C.; Veprintsev, D. B. *BMC Genomics* **2009**, *10*, 628.
(200) Xue, B.; Brown, C. J.; Dunker, A. K.; Uversky, V. N. *Biochim. Biophys. Acta* **2013**, *1834*, 725.
(201) Momand, J.; Villegas, A.; Belyi, V. A. *Gene* **2011**, *486*, 23.
(202) Bista, M.; Petrovich, M.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 17814.
(203) Brown, C. J.; Lain, S.; Verma, C. S.; Fersht, A. R.; Lane, D. P. *Nat. Rev. Cancer* **2009**, *9*, 862.
(204) Bullock, A. N.; Henckel, J.; Fersht, A. R. *Oncogene* **2000**, *19*, 1245.
(205) Joerger, A. C.; Fersht, A. R. *Oncogene* **2007**, *26*, 2226.
(206) Petitjean, A.; Mathe, E.; Kato, S.; Ishioka, C.; Tavtigian, S. V.; Hainaut, P.; Olivier, M. *Hum. Mutat.* **2007**, *28*, 622.
(207) Khoo, K. H.; Andreeva, A.; Fersht, A. R. *J. Mol. Biol.* **2009**, *393*, 161.
(208) Joerger, A. C.; Ang, H. C.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 15056.
(209) Basse, N.; Kaar, J. L.; Settanni, G.; Joerger, A. C.; Rutherford, T. J.; Fersht, A. R. *Chem. Biol.* **2010**, *17*, 46.
(210) Boeckler, F. M.; Joerger, A. C.; Jaggi, G.; Rutherford, T. J.; Veprintsev, D. B.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 10360.
(211) Liu, X.; Wilcken, R.; Joerger, A. C.; Chuckowree, I. S.; Amin, J.; Spencer, J.; Fersht, A. R. *Nucleic Acids Res.* **2013**, *41*, 6034.
(212) Wilcken, R.; Liu, X.; Zimmermann, M. O.; Rutherford, T. J.; Fersht, A. R.; Joerger, A. C.; Boeckler, F. M. *J. Am. Chem. Soc.* **2012**, *134*, 6810.
(213) Wassman, C. D.; Baronio, R.; Demir, O.; Wallentine, B. D.; Chen, C. K.; Hall, L. V.; Salehi, F.; Lin, D. W.; Chung, B. P.; Hatfield, G. W.; Chamberlin, A. R.; Luecke, H.; Lathrop, R. H.; Kaiser, P.; Amaro, R. E. *Nat. Commun.* **2013**, *4*, 1407.
(214) Wade, M.; Li, Y. C.; Wahl, G. M. *Nat. Rev. Cancer* **2013**, *13*, 83.
(215) Vassilev, L. T.; Vu, B. T.; Graves, B.; Carvajal, D.; Podlaski, F.; Filipovic, Z.; Kong, N.; Kammlott, U.; Lukacs, C.; Klein, C.; Fotouhi, N.; Liu, E. A. *Science* **2004**, *303*, 844.
(216) Vassilev, L. T. *Cell Cycle* **2004**, *3*, 419.
(217) Sullivan, K. D.; Padilla-Just, N.; Henry, R. E.; Porter, C. C.; Kim, J.; Tentler, J. J.; Eckhardt, S. G.; Tan, A. C.; DeGregori, J.; Espinosa, J. M. *Nat. Chem. Biol.* **2012**, *8*, 646.
(218) Van Maerken, T.; Rihani, A.; Dreidax, D.; De Clercq, S.; Yigit, N.; Marine, J. C.; Westermann, F.; De Paepe, A.; Vandesompele, J.; Speleman, F. *Mol. Cancer Ther.* **2011**, *10*, 983.
(219) Yun, M. H.; Gates, P. B.; Brockes, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 17392.
(220) Grasberger, B. L.; Lu, T.; Schubert, C.; Parks, D. J.; Carver, T. E.; Koblish, H. K.; Cummings, M. D.; LaFrance, L. V.; Milkiewicz, K. L.; Calvo, R. R.; Maguire, D.; Lattanze, J.; Franks, C. F.; Zhao, S.; Ramachandren, K.; Bylebyl, G. R.; Zhang, M.; Manthey, C. L.; Petrella, E. C.; Pantoliano, M. W.; Deckman, I. C.; Spurlino, J. C.; Maroney, A. C.; Tomczuk, B. E.; Molloy, C. J.; Bone, R. F. *J. Med. Chem.* **2005**, *48*, 909.
(221) Koblish, H. K.; Zhao, S.; Franks, C. F.; Donatelli, R. R.; Tominovich, R. M.; LaFrance, L. V.; Leonard, K. A.; Gushue, J. M.; Parks, D. J.; Calvo, R. R.; Milkiewicz, K. L.; Marugan, J. J.; Raboisson, P.; Cummings, M. D.; Grasberger, B. L.; Johnson, D. L.; Lu, T.; Molloy, C. J.; Maroney, A. C. *Mol. Cancer Ther.* **2006**, *5*, 160.
(222) Zhao, Y.; Yu, S.; Sun, W.; Liu, L.; Lu, J.; McEachern, D.; Shargary, S.; Bernard, D.; Li, X.; Zhao, T.; Zou, P.; Sun, D.; Wang, S. *J. Med. Chem.* **2013**, *56*, 5553.
(223) Popowicz, G. M.; Czarna, A.; Wolf, S.; Wang, K.; Wang, W.; Domling, A.; Holak, T. A. *Cell Cycle* **2010**, *9*, 1104.
(224) Hardcastle, I. R.; Liu, J.; Valeur, E.; Watson, A.; Ahmed, S. U.; Blackburn, T. J.; Bennaceur, K.; Clegg, W.; Drummond, C.; Endicott, J. A.; Golding, B. T.; Griffin, R. J.; Gruber, J.; Haggerty, K.; Harrington, R. W.; Hutton, C.; Kemp, S.; Lu, X.; McDonnell, J. M.; Newell, D. R.;

Noble, M. E.; Payne, S. L.; Revill, C. H.; Riedinger, C.; Xu, Q.; Lunec, J. *J. Med. Chem.* **2011**, *54*, 1233.
(225) Vogel, S. M.; Bauer, M. R.; Joerger, A. C.; Wilcken, R.; Brandt, T.; Veprintsev, D. B.; Rutherford, T. J.; Fersht, A. R.; Boeckler, F. M. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 16906.
(226) Michelsen, K.; Jordan, J. B.; Lewis, J.; Long, A. M.; Yang, E.; Rew, Y.; Zhou, J.; Yakowec, P.; Schnier, P. D.; Huang, X.; Poppe, L. *J. Am. Chem. Soc.* **2012**, *134*, 17059.
(227) Bista, M.; Wolf, S.; Khoury, K.; Kowalska, K.; Huang, Y.; Wrona, E.; Arciniega, M.; Popowicz, G. M.; Holak, T. A.; Dömling, A. *Structure* **2013**, *21*, 2143.
(228) Bernal, F.; Tyler, A. F.; Korsmeyer, S. J.; Walensky, L. D.; Verdine, G. L. *J. Am. Chem. Soc.* **2007**, *129*, 2456.
(229) Brown, C. J.; Quah, S. T.; Jong, J.; Goh, A. M.; Chiam, P. C.; Khoo, K. H.; Choong, M. L.; Lee, M. A.; Yurlova, L.; Zolghadr, K.; Joseph, T. L.; Verma, C. S.; Lane, D. P. *ACS Chem. Biol.* **2013**, *8*, 506.
(230) Chang, Y. S.; Graves, B.; Guerlavais, V.; Tovar, C.; Packman, K.; To, K. H.; Olson, K. A.; Kesavan, K.; Gangurde, P.; Mukherjee, A.; Baker, T.; Darlak, K.; Elkin, C.; Filipovic, Z.; Qureshi, F. Z.; Cai, H.; Berry, P.; Feyfant, E.; Shi, X. E.; Horstick, J.; Annis, D. A.; Manning, A. M.; Fotouhi, N.; Nash, H.; Vassilev, L. T.; Sawyer, T. K. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, E3445.
(231) Verdine, G. L.; Hilinski, G. J. *Methods Enzymol.* **2012**, *503*, 3.
(232) Bernal, F.; Wade, M.; Godes, M.; Davis, T. N.; Whitehead, D. G.; Kung, A. L.; Wahl, G. M.; Walensky, L. D. *Cancer Cell* **2010**, *18*, 411.
(233) Baek, S.; Kutchukian, P. S.; Verdine, G. L.; Huber, R.; Holak, T. A.; Lee, K. W.; Popowicz, G. M. *J. Am. Chem. Soc.* **2012**, *134*, 103.
(234) Zhao, Y.; Bernard, D.; Wang, S. *BioDiscovery* **2013**, *8*, 4.
(235) Tai, E.; Benchimol, S. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 11431.
(236) Piccinin, S.; Tonin, E.; Sessa, S.; Demontis, S.; Rossi, S.; Pecciarini, L.; Zanatta, L.; Pivetta, F.; Grizzo, A.; Sonego, M.; Rosano, C.; Dei Tos, A. P.; Doglioni, C.; Maestro, R. *Cancer Cell* **2012**, *22*, 404.
(237) Song, M. S.; Salmena, L.; Pandolfi, P. P. *Nat. Rev. Mol. Cell. Biol.* **2012**, *13*, 283.
(238) Maehama, T.; Dixon, J. E. *J. Biol. Chem.* **1998**, *273*, 13375.
(239) Sun, H.; Lesche, R.; Li, D. M.; Liliental, J.; Zhang, H.; Gao, J.; Gavrilova, N.; Mueller, B.; Liu, X.; Wu, H. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6199.
(240) Blanco-Aparicio, C.; Renner, O.; Leal, J. F. M.; Carnero, A. *Carcinogenesis* **2007**, *28*, 1379.
(241) Malaney, P.; Uversky, V. N.; Dave, V. *Mol. BioSyst.* **2013**, *9*, 2877.
(242) Lee, J. O.; Yang, H.; Georgescu, M. M.; Di Cristofano, A.; Maehama, T.; Shi, Y.; Dixon, J. E.; Pandolfi, P.; Pavletich, N. P. *Cell* **1999**, *99*, 323.
(243) Hopkins, B. D.; Fine, B.; Steinbach, N.; Dendy, M.; Rapp, Z.; Shaw, J.; Pappas, K.; Yu, J. S.; Hodakoski, C.; Mense, S.; Klein, J.; Pegno, S.; Sulis, M. L.; Goldsteini, H.; Amendolara, B.; Lei, L.; Maurer, M.; Bruce, J.; Canoll, P.; Hibshoosh, H.; Parsons, R. *Science* **2013**, *341*, 399.
(244) Hollander, M. C.; Blumenthal, G. M.; Dennis, P. A. *Nat. Rev. Cancer* **2011**, *11*, 289.
(245) Tamguney, T.; Stokoe, D. *J. Cell Sci.* **2007**, *120*, 4071.
(246) Zhang, P.; Chen, J. H.; Guo, X. L. *Biomed. Pharmacother.* **2012**, *66*, 485.
(247) Tian, L.; Fang, Y. X.; Xue, J. L.; Chen, J. Z. *PLoS One* **2013**, *8*, No. e75885.
(248) Poliseno, L.; Salmena, L.; Zhang, J.; Carver, B.; Haveman, W. J.; Pandolfi, P. P. *Nature* **2010**, *465*, 1033.
(249) Xia, H.; Khalil, W.; Kahm, J.; Jessurun, J.; Kleidon, J.; Henke, C. A. *Am. J. Pathol.* **2010**, *176*, 2626.
(250) Zmajkovicova, K.; Jesenberger, V.; Catalanotti, F.; Baumgartner, C.; Reyes, G.; Baccarini, M. *Mol. Cell* **2013**, *50*, 43.
(251) Fata, J. E.; Debnath, S.; Jenkins, E. C., Jr.; Fournier, M. V. *Int. J. Cell. Biol.* **2012**, *2012*, No. 379685.
(252) Leonard, M. K.; Kommagani, R.; Payal, V.; Mayo, L. D.; Shamma, H. N.; Kadakia, M. P. *Cell Death Differ.* **2011**, *18*, 1924.
(253) Bassi, C.; Ho, J.; Srikumar, T.; Dowling, R. J.; Gorrini, C.; Miller, S. J.; Mak, T. W.; Neel, B. G.; Raught, B.; Stambolic, V. *Science* **2013**, *341*, 395.
(254) Singh, G.; Chan, A. M. *Curr. Cancer Drug Targets* **2011**, *11*, 536.
(255) Wu, Y.; Zhou, H.; Wu, K.; Lee, S.; Li, R.; Liu, X. *Antioxid. Redox Signaling* **2014**, *20*, 1382.
(256) Kavela, S.; Shinde, S. R.; Ratheesh, R.; Viswakalyan, K.; Bashyam, M. D.; Gowrishankar, S.; Vamsy, M.; Pattnaik, S.; Rao, S.; Sastry, R. A.; Srinivasulu, M.; Chen, J. J.; Maddika, S. *Cancer Res.* **2013**, *73*, 205.
(257) Jung, S.; Li, C.; Jeong, D.; Lee, S.; Ohk, J.; Park, M.; Han, S.; Duan, J.; Kim, C.; Yang, Y.; Kim, K.; Lim, J.; Kang, Y.; Lee, M. *Int. J. Oncol.* **2013**, *43*, 1587.
(258) Ahmed, S. F.; Deb, S.; Paul, I.; Chatterjee, A.; Mandal, T.; Chatterjee, U.; Ghosh, M. K. *J. Biol. Chem.* **2012**, *287*, 15996.
(259) Xu, D.; Yao, Y.; Jiang, X.; Lu, L.; Dai, W. *J. Biol. Chem.* **2010**, *285*, 39935.
(260) Bolduc, D.; Rahdar, M.; Tu-Sekine, B.; Sivakumaren, S. C.; Raben, D.; Amzel, L. M.; Devreotes, P.; Gabelli, S. B.; Cole, P. *Elife* **2013**, *2*, No. e00691.
(261) Odriozola, L.; Singh, G.; Hoang, T.; Chan, A. M. *J. Biol. Chem.* **2007**, *282*, 23306.
(262) Rahdar, M.; Inoue, T.; Meyer, T.; Zhang, J.; Vazquez, F.; Devreotes, P. N. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 480.
(263) Vazquez, F.; Ramaswamy, S.; Nakamura, N.; Sellers, W. R. *Mol. Cell. Biol.* **2000**, *20*, 5010.
(264) Lee, J. T.; Shan, J.; Zhong, J.; Li, M.; Zhou, B.; Zhou, A.; Parsons, R.; Gu, W. *Cell Res.* **2013**, *23*, 552.
(265) Cao, J. X.; Wan, L. X.; Hacker, E.; Dai, X. P.; Lenna, S.; Jimenez-Cervantes, C.; Wang, Y. J.; Leslie, N. R.; Xu, G. X.; Widlund, H. R.; Ryu, B.; Alani, R. M.; Dutton-Regester, K.; Goding, C. R.; Hayward, N. K.; Wei, W. Y.; Cui, R. T. *Mol. Cell* **2013**, *51*, 409.
(266) Fan, C.; He, L. Z.; Kapoor, A.; Rybak, A. P.; De Melo, J.; Cutz, J. C.; Tang, D. *Mol. Cancer* **2009**, *8*, 98.
(267) Wu, P.; Hu, Y. Z. *Curr. Med. Chem.* **2010**, *17*, 4326.
(268) Torres, V. E.; Boletta, A.; Chapman, A.; Gattone, V.; Pei, Y.; Qian, Q.; Wallace, D. P.; Weimbs, T.; Wüthrich, R. P. *Clin. J. Am. Soc. Nephrol.* **2010**, *5*, 1312.
(269) Wong, M. *Biomed J.* **2013**, *36*, 40.
(270) Eisenreich, A.; Rauch, U. *Cardiovasc. Ther.* **2011**, *29*, 29.
(271) Matsuda, S.; Nakanishi, A.; Wada, Y.; Kitagishi, Y. *Open Med. Chem. J.* **2013**, 23.
(272) Shehata, M.; Schnabl, S.; Demirtas, D.; Hilgarth, M.; Hubmann, R.; Ponath, E.; Badrnya, S.; Lehner, C.; Hoelbl, A.; Duechler, M.; Gaiger, A.; Zielinski, C.; Schwarzmeier, J. D.; Jaeger, U. *Blood* **2010**, *116*, 2513.
(273) Silva, A.; Yunes, J. A.; Cardoso, B. A.; Martins, L. R.; Jotta, P. Y.; Abecasis, M.; Nowill, A. E.; Leslie, N. R.; Cardoso, A. A.; Barata, J. T. *J. Clin. Invest.* **2008**, *118*, 3762.
(274) Buontempo, F.; Orsini, E.; Martins, L. R.; Antunes, I.; Lonetti, A.; Chiarini, F.; Tabellini, G.; Evangelisti, C.; Melchionda, F.; Pession, A.; Bertaina, A.; Locatelli, F.; McCubrey, J. A.; Cappellini, A.; Barata, J. T.; Martelli, A. M. *Leukemia* **2014**, *28*, 543.
(275) Singh, G.; Chan, A. M. *Curr. Cancer Drug Targets* **2011**, *11*, 536.
(276) Zhang, W.; Bai, W. *Clin. Transl. Oncol.* **2013**.
(277) Roy, S.; Yu, Y.; Padhye, S. B.; Sarkar, F. H.; Majumdar, A. P. *PLoS One* **2013**, *8*, No. e68543.
(278) Tierney, B. J.; McCann, G. A.; Cohn, D. E.; Eisenhauer, E.; Sudhakar, M.; Kuppusamy, P.; Hideg, K.; Selvendiran, K. *Cancer Biol. Ther* **2012**, *13*, 766.
(279) Ravi, Y.; Karuppaiyah, S.; Naidu, S. K.; Meduru, S.; Citro, L.; Khan, M.; Rivera, B. K.; Sai-Sudhakar, C. B.; Kuppusamy, P. Abstract 16618: Pharmacological Modulation of PTEN Ameliorates the Progression of Pulmonary Hypertension in Heart Failure. *American Heart Association*, 2011, *124*: A16618.

(280) Selvendiran, K.; Kuppusamy, M. L.; Bratasz, A.; Tong, L. Y.; Rivera, B. K.; Rink, C.; Sen, C. K.; Kalai, T.; Hideg, K.; Kuppusamy, P. *J. Pharmacol. Exp. Ther.* **2009**, *329*, 959.
(281) Wells, J. A.; McClendon, C. L. *Nature* **2007**, *450*, 1001.
(282) Uversky, V. N. *Biochim. Biophys. Acta* **2013**, *1834*, 932.
(283) Hilser, V. J.; Thompson, E. B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 8311.
(284) Lee, G. M.; Craik, C. S. *Science* **2009**, *324*, 213.
(285) Fry, D. C.; Wartchow, C.; Graves, B.; Janson, C.; Lukacs, C.; Kammlott, U.; Belunis, C.; Palme, S.; Klein, C.; Vu, B. *Med. Chem. Lett.* **2013**, *4*, 95.
(286) Cheng, Y.; LeGall, T.; Oldfield, C. J.; Mueller, J. P.; Van, Y. Y.; Romero, P.; Cortese, M. S.; Uversky, V. N.; Dunker, A. K. *Trends Biotechnol.* **2006**, *24*, 435.
(287) Shimba, N.; Nomura, A. M.; Marnett, A. B.; Craik, C. S. *J. Virol.* **2004**, *78*, 6657.
(288) Shahian, T.; Lee, G. M.; Lazic, A.; Arnold, L. A.; Velusamy, P.; Roels, C. M.; Guy, R. K.; Craik, C. S. *Nat. Chem. Biol.* **2009**, *5*, 640.
(289) Lee, G. M.; Shahian, T.; Baharuddin, A.; Gable, J. E.; Craik, C. S. *J. Mol. Biol.* **2011**, *411*, 999.
(290) Adler, A. S.; Lin, M. H.; Horlings, H.; Nuyten, D. S. A.; van de Vijver, M. J.; Chang, H. Y. *Nat. Genet.* **2006**, *38*, 421.
(291) Dang, C. V. *Cell* **2012**, *149*, 22.
(292) Nesbit, C. E.; Tersak, J. M.; Prochownik, E. V. *Oncogene* **1999**, *18*, 3004.
(293) Jiang, H.; Bower, K. E.; Beuscher, A. E.; Zhou, B.; Bobkov, A. A.; Olson, A. J.; Vogt, P. K. *Mol. Pharmacol.* **2009**, *76*, 491.
(294) Aikawa, Y.; Morimoto, K.; Yamamoto, T.; Chaki, H.; Hashiramoto, A.; Narita, H.; Hirono, S.; Shiozawa, S. *Nat. Biotechnol.* **2008**, *26*, 817.
(295) Tsuchida, K.; Chaki, H.; Takakura, T.; Kotsubo, H.; Tanaka, T.; Aikawa, Y.; Shiozawa, S.; Hirono, S. *J. Med. Chem.* **2006**, *49*, 80.
(296) Rishi, V.; Potter, T.; Laudeman, J.; Reinhart, R.; Silvers, T.; Selby, M.; Stevenson, T.; Krosky, P.; Stephen, A. G.; Acharya, A.; Moll, J.; Oh, W. J.; Scudiero, D.; Shoemaker, R. H.; Vinson, C. *Anal. Biochem.* **2005**, *340*, 259.
(297) Rishi, V.; Oh, W. J.; Heyerdahl, S. L.; Zhao, J. F.; Scudiero, D.; Shoemaker, R. H.; Vinson, C. *J. Struct. Biol.* **2010**, *170*, 216.
(298) Zhao, J. F.; Stagno, J. R.; Varticovski, L.; Nimako, E.; Rishi, V.; McKinnon, K.; Akee, R.; Shoemaker, R. H.; Ji, X. H.; Vinson, C. *Mol. Pharmacol.* **2012**, *82*, 814.
(299) Wang, Y.; Cesena, T. I.; Ohnishi, Y.; Burger-Caplan, R.; Lam, V.; Kirchhoff, P. D.; Larsen, S. D.; Larsen, M. J.; Nestler, E. J.; Rudenko, G. *Chem. Neurosci.* **2012**, *3*, 546.
(300) Metallo, S. J. *Curr. Opin. Chem. Biol.* **2010**, *14*, 481.
(301) Rezaei-Ghaleh, N.; Blackledge, M.; Zweckstetter, M. *ChemBioChem* **2012**, *13*, 930.
(302) Rodi, D. J.; Janes, R. W.; Sanganee, H. J.; Holton, R. A.; Wallace, B. A.; Makowski, L. *J. Mol. Biol.* **1999**, *285*, 197.
(303) Aoki, S.; Morohashi, K.; Sunoki, T.; Kuramochi, K.; Kobayashi, S.; Sugawara, F. *Bioconjugate Chem.* **2007**, *18*, 1981.
(304) Morohashi, K.; Yoshino, A.; Yoshimori, A.; Saito, S.; Tanuma, S.; Sakaguchi, K.; Sugawara, F. *Biochem. Pharmacol.* **2005**, *70*, 37.
(305) Nishiyama, K.; Takakusagi, Y.; Kusayanagi, T.; Matsumoto, Y.; Habu, S.; Kuramochi, K.; Sugawara, F.; Sakaguchi, K.; Takahashi, H.; Natsugari, H.; Kobayashi, S. *Bioorg. Med. Chem.* **2009**, *17*, 195.
(306) Takakusagi, Y.; Kobayashi, S.; Sugawara, F. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 4850.
(307) Clackson, T.; Wells, J. A. *Science* **1995**, *267*, 383.
(308) Follis, A. V.; Hammoudeh, D. I.; Wang, H. B.; Prochownik, E. V.; Metallo, S. J. *Chem. Biol.* **2008**, *15*, 1149.
(309) Hammoudeh, D. I.; Follis, A. V.; Prochownik, E. V.; Metallo, S. J. *J. Am. Chem. Soc.* **2009**, *131*, 7390.
(310) Das, R. K.; Crick, S. L.; Pappu, R. V. *J. Mol. Biol.* **2012**, *416*, 287.
(311) Meng, W. L.; Lyle, N.; Luan, B. W.; Raleigh, D. P.; Pappu, R. V. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 2123.
(312) Amati, B.; Dalton, S.; Brooks, M. W.; Littlewood, T. D.; Evan, G. I.; Land, H. *Nature* **1992**, *359*, 423.
(313) Blackwood, E. M.; Eisenman, R. N. *Science* **1991**, *251*, 1211.
(314) Lin, C. Y.; Loven, J.; Rahl, P. B.; Paranal, R. M.; Burge, C. B.; Bradner, J. E.; Lee, T. I.; Young, R. A. *Cell* **2012**, *151*, 56.
(315) Nair, S. K.; Burley, S. K. *Cell* **2003**, *112*, 193.
(316) Follis, A. V.; Hammoudeh, D. I.; Daab, A. T.; Metallo, S. J. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 807.
(317) Berg, T.; Cohen, S. B.; Desharnais, J.; Sonderegger, C.; Maslyar, D. J.; Goldberg, J.; Boger, D. L.; Vogt, P. K. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 3830.
(318) Shi, J.; Stover, J. S.; Whitby, L. R.; Vogt, P. K.; Boger, D. L. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6038.
(319) Xu, Y.; Shi, J.; Yamamoto, N.; Moss, J. A.; Vogt, P. K.; Janda, K. D. *Bioorg. Med. Chem.* **2006**, *14*, 2660.
(320) Kiessling, A.; Sperl, B.; Hollis, A.; Eick, D.; Berg, T. *Chem. Biol.* **2006**, *13*, 745.
(321) Kiessling, A.; Wiesinger, R.; Sperl, B.; Berg, T. *ChemMedChem* **2007**, *2*, 627.
(322) Mo, H.; Henriksson, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 6344.
(323) Frenzel, A.; Zirath, H.; Vita, M.; Albihn, A.; Henriksson, M. A. *PLoS One* **2011**, *6*, No. e27988.
(324) Yin, X.; Giap, C.; Lazo, J. S.; Prochownik, E. V. *Oncogene* **2003**, *22*, 6151.
(325) Wang, H.; Hammoudeh, D. I.; Follis, A. V.; Reese, B. E.; Lazo, J. S.; Metallo, S. J.; Prochownik, E. V. *Mol. Cancer Ther.* **2007**, *6*, 2399.
(326) Yap, J. L.; Wang, H. B.; Hu, A.; Chauhan, J.; Jung, K. Y.; Gharavi, R. B.; Prochownik, E. V.; Fletcher, S. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 370.
(327) Mustata, G.; Follis, A. V.; Hammoudeh, D. I.; Metallo, S. J.; Wang, H. B.; Prochownik, E. V.; Lazo, J. S.; Bahar, I. *J. Med. Chem.* **2009**, *52*, 1247.
(328) Wang, H. B.; Chauhan, J.; Hu, A.; Pendleton, K.; Yap, J. L.; Sabato, P. E.; Jones, J. W.; Perri, M.; Yu, J. S.; Cione, E.; Kane, M. A.; Fletcher, S.; Prochownik, E. V. *Oncotarget* **2013**, *4*, 936.
(329) Whitty, A. *Future Med. Chem.* **2011**, *3*, 797.
(330) Clausen, D. M.; Guo, J. X.; Parise, R. A.; Beumer, J. H.; Egorin, M. J.; Lazo, J. S.; Prochownik, E. V.; Eiseman, J. L. *J. Pharmacol. Exp. Ther.* **2010**, *335*, 715.
(331) Guo, J. X.; Parise, R. A.; Joseph, E.; Egorin, M. J.; Lazo, J. S.; Prochownik, E. V.; Eiseman, J. L. *Cancer Chemother. Pharmacol.* **2009**, *63*, 615.
(332) Harvey, S. R.; Porrini, M.; Stachl, C.; MacMillan, D.; Zinzalla, G.; Barran, P. E. *J. Am. Chem. Soc.* **2012**, *134*, 19384.
(333) Fuxreiter, M.; Tompa, P.; Simon, I.; Uversky, V. N.; Hansen, J. C.; Asturias, F. J. *Nat. Chem. Biol.* **2008**, *4*, 728.
(334) Liu, J. G.; Perumal, N. B.; Oldfield, C. J.; Su, E. W.; Uversky, V. N.; Dunker, A. K. *Biochemistry* **2006**, *45*, 6873.
(335) Delattre, O.; Zucman, J.; Plougastel, B.; Desmaze, C.; Melot, T.; Peter, M.; Kovar, H.; Joubert, I.; Dejong, P.; Rouleau, G.; Aurias, A.; Thomas, G. *Nature* **1992**, *359*, 162.
(336) Song, J. H.; Ng, S. C.; Tompa, P.; Lee, K. A. W.; Chan, H. S. *PLoS Comput. Biol.* **2013**, *9*, No. e1003239.
(337) Uren, A.; Tcherkasskaya, O.; Toretsky, J. A. *Biochemistry* **2004**, *43*, 13579.
(338) Erkizan, H. V.; Kong, Y. L.; Merchant, M.; Schlottmann, S.; Barber-Rotenberg, J. S.; Yuan, L. S.; Abaan, O. D.; Chou, T. H.; Dakshanamurthy, S.; Brown, M. L.; Uren, A.; Toretsky, J. A. *Nat. Med.* **2009**, *15*, 750.
(339) Barber-Rotenberg, J. S.; Selvanathan, S. P.; Kong, Y. L.; Erkizan, H. V.; Snyder, T. M.; Hong, S. P.; Kobs, C. L.; South, N. L.; Summer, S.; Monroe, P. J.; Chruszcz, M.; Dobrev, V.; Tosso, P. N.; Scher, L. J.; Minor, W.; Brown, M. L.; Metallo, S. J.; Uren, A.; Toretsky, J. A. *Oncotarget* **2012**, *3*, 172.
(340) Yuan, X.; Cai, C.; Chen, S.; Chen, S.; Yu, Z.; Bal, S. P. *Oncogene* **2013**, in press.
(341) Sadar, M. D.; Williams, D. E.; Mawji, N. R.; Patrick, B. O.; Wikanta, T.; Chasanah, E.; Irianto, H. E.; Van Soest, R.; Andersen, R. J. *Org. Lett.* **2008**, *10*, 4947.

(342) Meimetis, L. G.; Williams, D. E.; Mawji, N. R.; Banuelos, C. A.; Lal, A. A.; Park, J. J.; Tien, A. H.; Fernandez, J. G.; de Voogd, N. J.; Sadar, M. D.; Andersen, R. J. *J. Med. Chem.* **2012**, *55*, 503.
(343) Andersen, R. J.; Mawji, N. R.; Wang, J.; Wang, G.; Haile, S.; Myung, J. K.; Watt, K.; Tam, T.; Yang, Y. C.; Banuelos, C. A.; Williams, D. E.; McEwan, I. J.; Wang, Y. Z.; Sadar, M. D. *Cancer Cell* **2010**, *17*, 535.
(344) Satoh, K.; Ohyama, K.; Aoki, N.; Iida, M.; Nagai, F. *Food Chem. Toxicol.* **2004**, *42*, 983.
(345) Myung, J. K.; Banuelos, C. A.; Fernandez, J. G.; Mawji, N. R.; Wang, J.; Tien, A. H.; Yang, Y. C.; Tavakoli, I.; Haile, S.; Watt, K.; McEwan, I. J.; Plymate, S.; Andersen, R. J.; Sadar, M. D. *J. Clin. Invest.* **2013**, *123*, 2948.
(346) Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. *Nat. Rev. Drug Discovery* **2011**, *10*, 307.
(347) Ghanta, J.; Shen, C. L.; Kiessling, L. L.; Murphy, R. M. *J. Biol. Chem.* **1996**, *271*, 29525.
(348) Bulic, B.; Pickhardt, M.; Mandelkow, E. M.; Mandelkow, E. *Neuropharmacology* **2010**, *59*, 276.
(349) Masuda, M.; Suzuki, N.; Taniguchi, S.; Oikawa, T.; Nonaka, T.; Iwatsubo, T.; Hisanaga, S.; Goedert, M.; Hasegawa, M. *Biochemistry* **2006**, *45*, 6085.
(350) Feng, B. Y.; Toyama, B. H.; Wille, H.; Colby, D. W.; Collins, S. R.; May, B. C. H.; Prusiner, S. B.; Weissman, J.; Shoichet, B. K. *Nat. Chem. Biol.* **2008**, *4*, 197.
(351) Lendel, C.; Bolognesi, B.; Wahlstrom, A.; Dobson, C. M.; Graslund, A. *Biochemistry* **2010**, *49*, 1358.
(352) Kukar, T. L.; Ladd, T. B.; Bann, M. A.; Fraering, P. C.; Narlawar, R.; Maharvi, G. M.; Healy, B.; Chapman, R.; Welzel, A. T.; Price, R. W.; Moore, B.; Rangachari, V.; Cusack, B.; Eriksen, J.; Jansen-West, K.; Verbeeck, C.; Yager, D.; Eckman, C.; Ye, W.; Sagi, S.; Cottrell, B. A.; Torpey, J.; Rosenberry, T. L.; Fauq, A.; Wolfe, M. S.; Schmidt, B.; Walsh, D. M.; Koo, E. H.; Golde, T. E. *Nature* **2008**, *453*, 925.
(353) Botev, A.; Munter, L. M.; Wenzel, R.; Richter, L.; Althoff, V.; Ismer, J.; Gerling, U.; Weise, C.; Koksch, B.; Hildebrand, P. W.; Bittl, R.; Multhaup, G. *Biochemistry* **2011**, *50*, 828.
(354) Sagi, S. A.; Lessard, C. B.; Winden, K. D.; Maruyama, H.; Koo, J. C.; Weggen, S.; Kukar, T. L.; Golde, T. E.; Koo, E. H. *J. Biol. Chem.* **2011**, *286*, 39794.
(355) Barrett, P. J.; Sanders, C. R.; Kaufman, S. A.; Michelsen, K.; Jordan, J. B. *Biochemistry* **2011**, *50*, 10328.
(356) Beel, A. J.; Barrett, P.; Schnier, P. D.; Hitchcock, S. A.; Bagal, D.; Sanders, C. R.; Jordan, J. B. *Biochemistry* **2009**, *48*, 11837.
(357) Kodadek, T. *Biopolymers* **2002**, *66*, 134.
(358) Ehrnhoefer, D. E.; Bieschke, J.; Boeddrich, A.; Herbst, M.; Masino, L.; Lurz, R.; Engemann, S.; Pastore, A.; Wanker, E. E. *Nat. Struct. Mol. Biol.* **2008**, *15*, 558.
(359) Palhano, F. L.; Lee, J.; Grimster, N. P.; Kelly, J. W. *J. Am. Chem. Soc.* **2013**, *135*, 7503.
(360) Akoury, E.; Gajda, M.; Pickhard, M.; Biernat, J.; Soraya, P.; Griesinger, C.; Mandelkow, E.; Zweckstetter, M. *J. Am. Chem. Soc.* **2013**, *135*, 2853.
(361) Woods, L. A.; Platt, G. W.; Hellewell, A. L.; Hewitt, E. W.; Homans, S. W.; Ashcroft, A. E.; Radford, S. E. *Nat. Chem. Biol.* **2011**, *7*, 730.
(362) Attanasio, F.; Convertino, M.; Magno, A.; Caflisch, A.; Corazza, A.; Haridas, H.; Esposito, G.; Cataldo, S.; Pignataro, B.; Milardi, D.; Rizzarelli, E. *ChemBioChem* **2013**, *14*, 583.
(363) Fuxreiter, M.; Tompa, P.; Simon, I. *Bioinformatics* **2007**, *23*, 950.
(364) Davey, N. E.; Haslam, N. J.; Shields, D. C.; Edwards, R. J. *Nucleic Acids Res.* **2011**, *39*, W56.
(365) Disfani, F. M.; Hsu, W. L.; Mizianty, M. J.; Oldfield, C. J.; Xue, B.; Dunker, A. K.; Uversky, V. N.; Kurgan, L. *Bioinformatics* **2012**, *28*, I75.
(366) Meszaros, B.; Simon, I.; Dosztanyi, Z. *PLoS Comput. Biol.* **2009**, *5*, No. e1000376.
(367) Michel, J.; Cuchillo, R. *PLoS One* **2012**, *7*, No. e41070.
(368) Jin, F.; Yu, C.; Lai, L.; Liu, Z. *PLoS Comput. Biol.* **2013**, *9*, No. e1003249.
(369) Herrera, F. E.; Chesi, A.; Paleologou, K. E.; Schmid, A.; Munoz, A.; Vendruscolo, M.; Gustincich, S.; Lashuel, H. A.; Carloni, P. *PLoS One* **2008**, *3*, No. e3394.
(370) Conway, K. A.; Rochet, J. C.; Bieganski, R. M.; Lansbury, P. T. *Science* **2001**, *294*, 1346.
(371) Norris, E. H.; Giasson, B. I.; Hodara, R.; Xu, S. H.; Trojanowski, J. Q.; Ischiropoulos, H.; Lee, V. M. Y. *J. Biol. Chem.* **2005**, *280*, 21212.
(372) Zhu, M.; De Simone, A.; Schenk, D.; Toth, G.; Dobson, C. M.; Vendruscolo, M. *J. Chem. Phys.* **2013**, *139*, No. 035101.
(373) Tompa, P.; Szasz, C.; Buday, L. *Trends Biochem. Sci.* **2005**, *30*, 484.
(374) Xue, B.; Dunbrack, R. L.; Williams, R. W.; Dunker, A. K.; Uversky, V. N. *Biochim. Biophys. Acta* **2010**, *1804*, 996.
(375) Natan, E.; Baloglu, C.; Pagel, K.; Freund, S. M.; Morgner, N.; Robinson, C. V.; Fersht, A. R.; Joerger, A. C. *J. Mol. Biol.* **2011**, *409*, 358.
(376) Edgar, R. C. *Nucleic Acids Res.* **2004**, *32*, 1792.
(377) Waterhouse, A. M.; Procter, J. B.; Martin, D. M.; Clamp, M.; Barton, G. J. *Bioinformatics* **2009**, *25*, 1189.
(378) Anil, B.; Riedinger, C.; Endicott, J. A.; Noble, M. E. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1358.
(379) Kuntzer, J.; Eggle, D.; Lenhof, H. P.; Burtscher, H.; Klostermann, S. *Hum. Mutat.* **2010**, *31*, 407.