602 *Biochemistry* **2011,** *50,* 602–611
DOI: 10.1021/bi101563v

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Current Topic

Multispecific Recognition: Mechanism, Evolution, and Design[†]

Ariel Erijman, Yonatan Aizner, and Julia M. Shifman*

Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

Received September 27, 2010; Revised Manuscript Received December 22, 2010

ABSTRACT: Accumulating evidence shows that many particular proteins have evolved to bind multiple targets, including other proteins, peptides, DNA, and small molecule substrates. Multispecific recognition might be not only common but also necessary for the robustness of signaling and metabolic networks in the cell. It is also important for the immune response and for regulation of transcription and translation. Multispecificity presents an apparent paradox: How can a protein encoded by a single sequence accommodate numerous targets? Analysis of sequences and structures of multispecific proteins revealed a number of mechanisms that achieve multispecificity. Interestingly, similar mechanisms appear in antibody−antigen, T-cell receptor−peptide, protein−DNA, enzyme−substrate, and protein−protein complexes. Directed evolution and protein design experiments with multispecific proteins offer some interesting insights into the evolution of such proteins and help in the dissection of molecular interactions that mediate multispecificity. Understanding the basic principles governing multispecificity could greatly assist in the unraveling of various complex processes in the cell. In addition, through manipulation of functional multispecificity, novel proteins could be created for use in various biotechnological and biomedical applications.

For many years, specificity has been considered essential to the functional mechanism of proteins. A widely accepted belief related a particular amino acid sequence to a particular three-dimensional (3D) protein structure, which in turn defined protein function. In the past decade, this hypothesis has been widely challenged by many counter examples of intrinsically disordered proteins and proteins that undergo large conformational changes during their life cycle. Accumulating experimental evidence shows that promiscuity and multispecificity are not rare phenomena in the protein world as previously thought. On the contrary, they are universal properties of molecular recognition and are manifested in protein−protein, protein−DNA, antibody−antigen, and T-cell receptor−peptide interactions (Figure 1). Multispecificity might be required to account for the robustness of signaling networks and for the efficiency of cellular metabolism and immune response. The necessity for multispecificity in molecular recognition changes our understanding of protein evolution and has far-reaching implications for protein and drug design.

Multispecificity raises a number of interesting questions. How can a protein encoded by a single sequence recognize multiple binding partners? What are the evolutionary compromises that allow binding promiscuity? What sequence and structural characteristics distinguish multispecific proteins from those that perform only a single function? What molecular interactions give rise to multispecificity? How can we incorporate the information described above when designing such multispecific proteins? These and some other questions are discussed in this review.

MULTISPECIFICITY IN PROTEIN−PROTEIN INTERACTIONS

Several proteins play a highly important role in the cell because they bind to and regulate a large number of partners. Some famous examples are p53, p21, p27, calmodulin, ubiquitin, thioredoxin, and Ras. In the past decade, several high-throughput experiments that aimed to identify all possible protein−protein (PP)[1] interactions in several organisms have been conducted (*1*−*8*). Although the obtained data are noisy and far from exhaustive, they instigated many theoretical studies of PP interaction networks. Interestingly, PP networks resemble many other complex networks such as the Web, power grids, and manuscript citations. The main characteristic of these "scale-free" networks is that they are not randomly connected (*9*). Most of the nodes in such networks have very few interactions, while only a small number of nodes have many interacting partners. These highly connected nodes in PP networks correspond to multispecific proteins, which are usually termed protein hubs in the interaction network language.

In the past, many groups have worked to identify the unique characteristics of protein hubs (reviewed in ref *10*). Protein hubs are essential to the network topology, so deleting them should result in the destruction of the network. Indeed, it was shown experimentally that deletion of a hub is lethal to the organism (*11*). From the network viewpoint, hub proteins should exhibit lower evolutionary rates. However, several

[†]This research is supported by the Israel Science Foundation (1372/10) and DFG Grant El 423/2-1.

*To whom correspondence should be addressed. Phone: 972-2-658-4078. Fax: 972-2-658-6448. E-mail: jshifman@cc.huji.ac.il.

[1]Abbreviations: PP, protein−protein; TCR, T-cell receptor; CDR, complementarity-determining region; MHC, major histocompatibility complex; VPS9, vacuolar protein sorting-associated protein; TSG101, tumor susceptibility gene 101 protein; CBP, CREB binding protein; HIF1α, hypoxia inducible factor 1α; NK receptor, natural killer receptor; PXR, human pregnane X receptor; HER2, Human epidermal growth factor receptor 2; VEGF, vascular endothelial growth factor; AbrB, antibiotic resistance protein B; BRCA1, breast cancer 1 susceptibility protein; HSL, homoserine lactone.

 Published on Web 01/03/2011

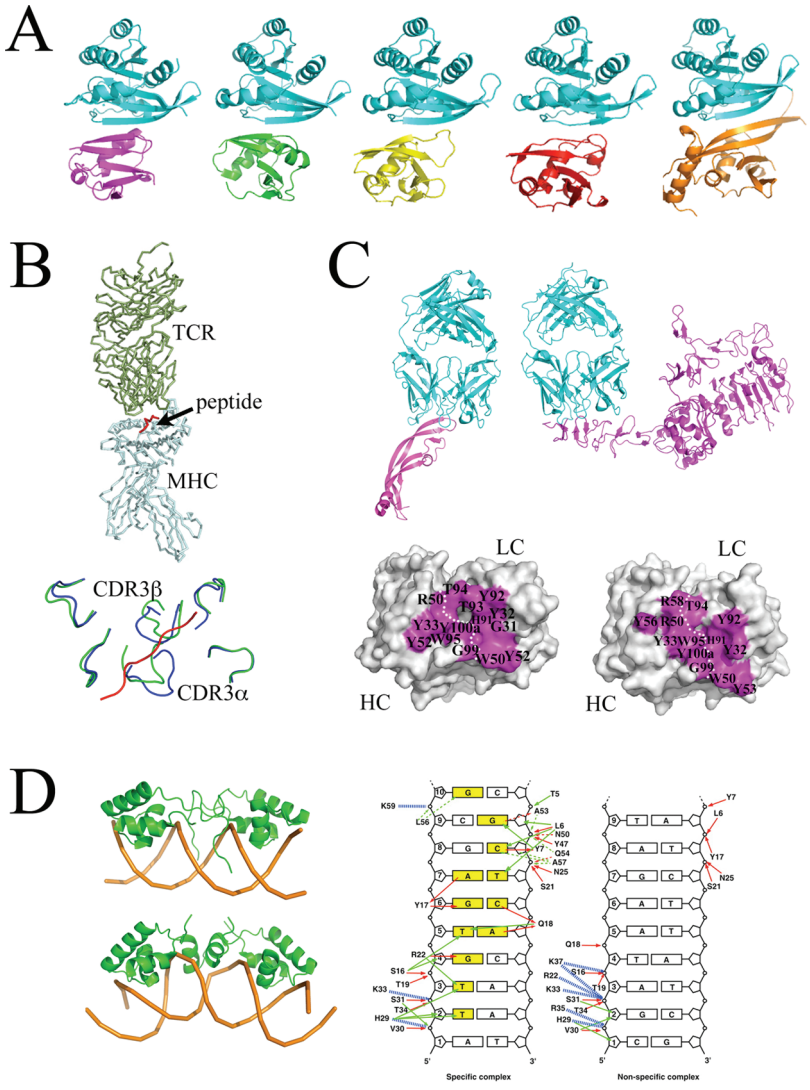

FIGURE 1: Examples of multispecific recognition. (A) Multispecific protein Ras interacts with many effectors through the same binding interface. Shown from left to right are complexes of Ras (cyan) with Ras-binding domains from Raf (magenta), Byr2 kinase (green), RalGDS (yellow), phospholipase Cε (red), and Nore1A (orange). (B) Multispecific recognition in TCRs. The top panel shows the structure of TCR in complex with the MHC−peptide complex (Protein Data Bank entry 2CKB). The bottom panel zooms into the TCR−peptide interface and superimposes CDR loops on BM3.3 TCR when interacting with two different peptides: VSV8 (RGYVYQGL) and pBM8 (SQYYYNSL). CDR loops are colored green for the TCR−VSV8 complex and blue for the TCR−pBM8 complex. Because the conformation of both peptides is very similar, only one peptide is shown in red. The MHC presenting the peptide is at the top (not shown). (C) The top panel shows structures of an engineered bispecific antibody (cyan) in complex with vascular endothelial growth factor (VEGF) (left) and human epidermal growth factor receptor 2 (HER2) (right) (*42*). The bottom panel shows the binding interface on the bispecific antibody when it interacts with VEGF (left) and HER2 (right). The residues contacting each antigen are colored magenta and labeled. (D) Comparison of nonspecific and specific protein−DNA interactions. The NMR structures of the lac repressor dimer (green) in complex with nonspecific DNA (top left) and with its specific DNA site (bottom left). The DNA is shown as an orange ribbon. In the nonspecific complex, the hinge region of the lac repressor is unfolded and the DNA remains in the canonical B form. In the specific complex, the hinge region folds up into an α-helix and the DNA is bent by ~36°. On the right, is the schematic diagram depicting molecular interactions in the specific and the nonspecific lac repressor−DNA complex. The bases that are specifically recognized by lac repressor are colored yellow. Red, green, and dashed blue lines indicate hydrogen bonding, hydrophobic, and electrostatic contacts, respectively. Part of panel D was reproduced from the work of Kalodimos et al. (*48*). Copyright 2004 American Association for the Advancement of Science. This figure was generated with PyMOL (*112*).

theoretical studies reached different conclusions (*12*−*14*). This controversy was resolved when structural information was incorporated into the study and revealed two possible scenarios for a hub's high connectivity. In the first scenario, a group of hubs can interact with many proteins at the same time by using multiple binding interfaces. These hubs are usually large multidomain proteins that participate in permanent interactions, such as in macromolecular assembly. In the second scenario, hubs termed "single-interface hubs" bind to many potential partners at different times through the same or highly overlapping binding interface, participating in transient interactions. The evolutionary mechanism of these two groups of hubs is likely to be different; thus, merging them could eliminate important signals. Indeed, Kim et al. (*15*) determined that the evolutionary rate of multi-interface hubs is significantly slower than that of the whole proteome, decreasing in proportion to the increasing surface area of the interacting regions. In contrast, the evolutionary rate of the single-interface hubs does not differ significantly from that of the whole proteome (*15*). These findings are not surprising because constraints that slow evolution are applied only to the binding interface area and not to the whole protein sequence (*16*).

MULTISPECIFICITY IN ENZYME CATALYSIS

For many years, enzymes have been considered proteins that evolved to catalyze a single reaction with high efficiency and high specificity. During the past decade, the dogma of "one enzyme−one catalytic reaction" has become obsolete because many enzymes have been shown to catalyze more than one reaction for more than one substrate. Several reviews on the subject of enzyme promiscuity have recently been written (*17*−*21*).

In contrast to PP interactions, the promiscuous functions of enzymes are difficult to identify in high-throughput experiments. One reason for this is that no available experimental technique can detect the whole range of different substrates and reactions. Another reason is that the efficiency of promiscuous reactions is usually low in comparison to that of the native reaction (*20*). Despite this, promiscuous reactions can be physiologically relevant under certain conditions. While at normal conditions promiscuous enzymes generate an undetectable quantity of their secondary products, with a sudden change in the substrate or enzyme concentration, alternative pathways could become more important and even preferred (*22*). Pathway promiscuity may be a form of adaptive plasticity for the generation of new metabolic capabilities with no genetic changes. This "underground" network is one reason why organisms can often be much more robust than expected after deletion of genes involved in major metabolic pathways. For example, Kim et al. found that 80 of 227 genes encoding metabolic enzymes involved in glucose metabolism are not essential for growth of *Escherichia coli* on glucose (*23*). Analysis of the data showed the emergence of alternative pathways in which the missing enzyme was substituted with a different enzyme. In another study, the authors screened 104 single-gene knockout strains and discovered that 20% of these auxotrophs were rescued by overexpression of at least one noncognate gene, which was usually unrelated to the deleted gene (*24*).

Because extensive experimental data on enzyme promiscuity are unavailable, computational predictions could help in understanding the scale of this phenomenon. Recently, several algorithms for predicting enzyme promiscuity were developed either on the basis of structural comparison between active sites and substrates of various enzymes (*25*) or on the basis of molecular docking of substrates (*26*, *27*) and transition-state analogues (*28*, *29*) into various enzymes. Computational predictions identified a large number of potential promiscuous enzymatic reactions and interesting bypasses and cross-links in metabolic pathways (*25*, *29*) that have yet to be verified experimentally.

MULTISPECIFICITY IN IMMUNE RESPONSE

While the extent of multispecificity in enzymes has not been appreciated until recently, the importance of multispecificity in the immune system was realized four decades ago (*30*, *31*). The immune system has to recognize an apparently limitless variety of antigenic determinants. A very large number of unique antibodies and TCRs could be generated through gene rearrangement and N region substitution. However, there are far more potentially immunogenic peptides compared to the number of T-cells in the animal (*32*). Similarly, the number of B-cells that produce antibodies appears to be insufficient to cover an enormous diversity of various antigens. This implies that antibodies, TCRs, and other immune system receptors must bind to multiple ligands (*33*).

Multispecificity in TCRs and antibodies has a distinct physiological role in immune response. TCRs recognize antigenic peptides presented by class I and II MHC proteins and hence contact both a linear antigenic peptide and the surface of the MHC molecule (Figure 1B). Binding of a TCR to the MHC−peptide complex is required for initiation and propagation of the cellular immune response and for generation and maintenance of the T-cell repertoire (*34*). TCR's ability to recognize various antigenic peptides is crucial to T-cell development (*32*, *35*). In addition, TCR's multispecificity has been implicated in several autoimmune pathologies and is the main cause of transplant rejection (*36*).

Multispecific antibodies serve as a first line of defense against infection by bacteria or viruses by recognizing various antigens without being exposed to them previously. Additional proposed biological roles for such antibodies include the clearance of cellular antigens after tissue damage or remodeling and prevention of tumor growth (*33*, *37*). In many cases, multispecific antibodies have not yet undergone somatic mutations that are associated with high-affinity antibodies (*38*, *39*). It has been established that multispecificity in antibodies is conveyed by the variable domain (*40*). However, no differences in physicochemical properties of multispecific and monospecific antibodies have been detected (*41*). Very recently, it was suggested that engineered multispecific antibodies could become promising therapeutical agents against various diseases (Figure 1C) (*42*).

A number of mechanisms have been proposed to explain multispecificity in immune recognition (reviewed in ref *43*). These mechanisms are only a subset of all possible ways available to proteins accommodating various binding partners.

MULTISPECIFICITY IN DNA-BINDING PROTEINS

Protein−DNA interactions are critical for regulation of key biological functions such as transcription, translation, replication, and recombination. Hence, the high specificity of protein−DNA interactions would be a likely strategy for avoiding errors in these vital processes. Nevertheless, a recent high-throughput experiment showed that approximately half of the 104 examined DNA-binding proteins recognize several distinctly different DNA sequence motifs (*44*). Furthermore, some transcriptional regulators are known to bind specifically to a set of DNA targets that show little or no obvious consensus sequence. Such recognition allows them to regulate the expression level of many genes simultaneously in response to an environmental change. One example of these regulators is the antibiotic resistance protein B (AbrB) from *Bacillus subtilis* (*45*) that controls the expression of more than 60 different genes. Some additional bacterial DNA-binding proteins with similar properties have been identified in the past few years (*46*, *47*). These findings challenge our general understanding of how proteins recognize their DNA-binding sites.

In addition to several examples of specific binding to more than one DNA site, most if not all DNA-binding proteins bind DNA nonspecifically with appreciable affinity (*48*). This nonspecific binding is essential for accelerating a protein's search for its target DNA site through one-dimensional diffusion (*49*, *50*). The nonspecific protein−DNA interactions are mostly electrostatic in nature, resulting in both fast association and fast dissociation rates and allowing the protein to hop along the DNA molecule. Recent high-resolution structures of nonspecific and specific protein−DNA complexes help to reveal the mechanisms for multispecificity in protein−DNA recognition (*48*, *51*, *52*) (Figure 1D).

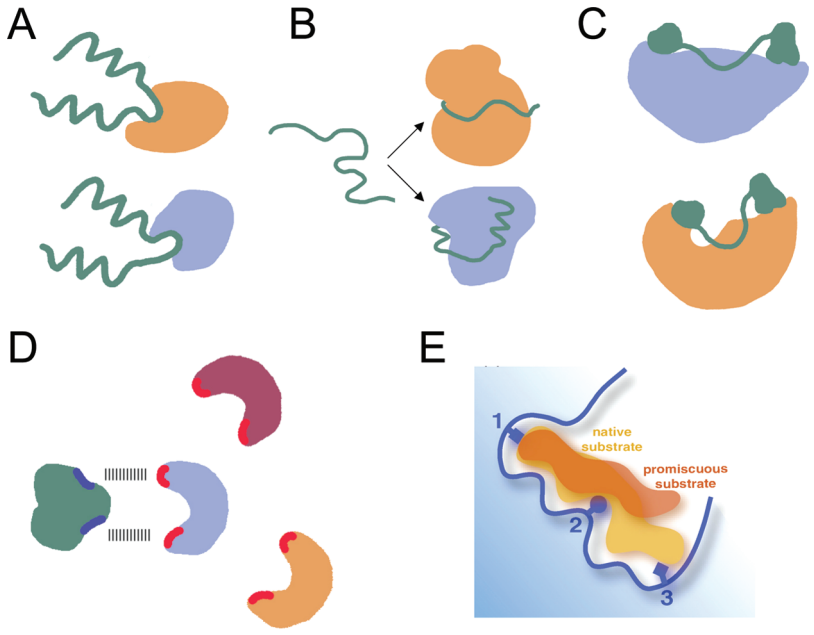

FIGURE 2: Mechanisms for multispecificity. (A) A different orientation of the binding loop provides an opportunity for binding to different partners colored orange and blue. (B) An intrinsically disordered protein can bind different targets by folding into different structures. It binds the orange partner in an open conformation and the blue partner in a helix−loop−helix conformation. (C) Some proteins use disordered linkers for changing relative orientations of the binding domains, thus accommodating different binding surfaces in various targets. (D) Charged residues give multispecific binders opportunity to make long-range nonspecific interactions with many partners before the specific interactions take place. (E) Partial ligand recognition [reproduced from the work of Babtie et al. (*20*); copyright 2010 Elsevier]. The binding site is optimized to bind one ligand (native substrate) with a few energetically important residues (indicated by 1−3). In most cases, promiscuous ligands show similarities with the native ligand and are able to make use of some of the same binding interactions and functional groups (1 and 2).

MECHANISMS FOR MULTISPECIFIC RECOGNITION

Multispecific recognition presents a certain paradox: "How has one protein evolved to accommodate many targets that are different in primary, secondary, and three-dimensional structure?" The simplest way to achieve multispecificity is to recognize each target through a distinct binding interface, providing simultaneous binding sites for various targets. The main strategy for such proteins is to use a large surface area and/or multiple binding domains. This mechanism has been observed for protein−protein, protein−DNA, and protein−small molecule interactions. For example, cytosine C5 DNA methyltransferase is able to bind to and methylate more than one DNA sequence because of the presence of several target-recognizing domains (*53*). The use of multiple domains is common in PP interactions, where the number of hub connections positively correlates with the number of domains present in the hub (*10*). Similarly, recognition of structurally different antigens by germline antibodies could be achieved by utilizing different binding interfaces (paratopes), giving rise to the "differential ligand positioning" mechanism (*38*).

Another group of mutispecific proteins utilize the same or overlapping residues for various target recognition (Figure 1A). Such proteins are frequently small or single-domain proteins; hence, they need to be more inventive when achieving multispecificity. Conformational flexibility is the most frequent explanation for multispecificity in PP and protein−small molecule recognition (*54*). Flexible proteins could easily adjust their structures to provide most favorable interactions with each of the interaction partners. Conformational flexibility could manifest itself at different levels starting with small changes such as reorientation of side chains, continuing with medium-range movement of loops and local alterations in the secondary structure, and finishing with large conformational changes, including disorder-to-order transitions. In PP interactions, all the levels of flexibility mentioned above are observed. In small molecule recognition, where binding is achieved through only a few residues, large-scale conformational changes do not usually occur.

Small conformational changes at the side chain level are frequently manifested in multispecific binding. It is thus not surprising that interfaces of hub proteins are enriched in flexible amino acids that can convey varying types of interactions (e.g., Arg, Tyr, His, and Met) (*55*). Arg, for example, can form salt bridges, hydrogen bonds, and cation−π interactions and can anchor into a binding groove of an interacting partner. Different positioning of the same Arg in ubiquitin permits this protein to recognize two targets: a vacuolar protein sorting-associated protein (VPS9) and tumor susceptibility gene 101 protein (TSG101) (*55*). The flexibility and mobility of Arg residues on the HoxD9 homeodomain are crucial for its interactions with nonspecific DNA, allowing the protein to slide along DNA without complete dissociation (*56*). Met is another example of a flexible amino acid that can pack differently in hydrophobic pockets, while complementing each particular target. This happens in calmodulin, which accommodates different targets by using a number of Met residues in its binding interface (*57*).

Movement of flexible loops is another evident strategy for accommodating various partners (Figure 2A). A recent NMR study of ubiquitin in solution revealed that this protein assumes an ensemble of conformations, differing by the orientation of several loops. The ensemble completely covers the structural heterogeneity observed in the X-ray structures of the ubiquitin−target complexes, suggesting that the conformational selection mechanism is responsible for target recognition (*58*). A similar mechanism is observed in antibody−antigen interactions, e.g., in dinitrophenyl antibody SPE7, which can exist in at least six different conformations in the unbound form. Two of these conformations

are selected upon binding to two different targets (*59*). In another example, adjustments of hypervariable loops are responsible for the binding of TCR molecules to different antigenic peptides. While CDR1 and CDR2 usually exhibit small rigid body movements, CDR3α and CDR3*β* exhibit larger changes to provide favorable interactions with each particular peptide (*60*) (Figure 1B).

Recent studies found that intrinsically disordered regions are frequent in protein hubs (*61*), and especially in single-interface hubs (*62*). Disordered regions in proteins could be readily predicted from protein sequences on the basis of amino acid composition, flexibility, hydropathy, charge, coordination number, and other factors (*63*−*65*). Dunker and colleagues analyzed sequences of hub proteins and divided them into several groups, including those that are entirely disordered, those that have medium-size disordered regions, and those that have little or no disorder (*66*). Interestingly, proteins in the latter group frequently bind to targets that are disordered.

Why are intrinsically disordered regions common to multispecific proteins? It turns out that disorder might provide certain advantages for the multispecific binding function. Disordered regions frequently fold upon target recognition and hence could take up different structures when binding to different targets (Figure 2B) (*67*, *68*). For example, hypoxia inducible factor 1α (HIF1α) is a mammalian transcription factor that lacks a 3D structure in its unbound state (*69*). It folds upon binding to the TAZ1 domain of the CREB binding protein (CBP) forming a helical structure and binds hydroxylase FIH using an open extended structure (*70*).

Disordered regions in protein hubs could be simple flexible linkers that allow the other binding domains to assume different orientations with respect to each other (Figure 2C). This mechanism is observed in E2 ubiquitin ligase (*71*) and in calmodulin (*72*). Breast cancer 1 susceptibility protein (BRCA1) has a central disordered domain of ∼1500 residues that binds to several proteins and DNA at the same time. This domain serves as a flexible linker between its N-terminal RING domain and the two C-terminal BRCT domains (*10*).

Disordered regions are useful in inducible reactions where binding could be turned on or off by an appropriate signal. For example, the KIX domain of the CBP interacts with many transcriptional activation domains, including the phosphorylated KID domain of CREB (*73*). Binding of KIX to the intrinsically disordered KID is usually impeded by the large entropic penalty associated with folding upon binding (*74*). However, when KID is phosphorylated at Ser 133, new hydrogen bonds and electrostatic interactions lead to a large negative binding enthalpy, making the binding thermodynamically favorable (*75*). In contrast to KID, the proto-oncogene c-Myb, which is structured in its unbound state, binds constitutively to KIX, without requiring activation by phosphorylation (*74*).

Structural disorder appears also to be useful for chaperones because their function requires them to bind to multiple partners. Disordered segments in chaperones provide a significant solubilizing effect upon binding to proteins [as was demonstrated in α-synuclein (*76*, *77*), α-casein (*78*), and α-crystallin (*79*, *80*)] and a long-range steric repulsive effect that prevents aggregation [as observed in HSP25 (*81*)]. Such chaperones utilize an ancient mechanism, termed "entropy transfer" (*82*), in which an intrinsically disordered chaperone rescues misfolded proteins by binding to them, inducing folding of the chaperone and unfolding of the protein.

Conformational changes in multispecific proteins are often caused by prior allosteric events such as binding to metal ions and/or cofactors, allowing recognition of one set of targets in one physiological state and another set of targets in another state. For example, calmodulin assumes different target-bound conformations depending on how many of its four calcium sites are occupied. With all four calciums bound, it recognizes one set of targets assuming a so-called "closed" conformation (*83*). It binds to another set of targets when two calcium sites are occupied in an extended conformation and binds to some targets without any calciums in yet another conformation (*83*).

Conformational flexibility and intrinsic disorder are important for binding multispecificity; however, they are not the only strategies for accommodating different targets. Most small hubs such as ubiquitin, ferredoxin, and Ras show low level of disordered residues. One apparent mechanism for such proteins to mediate interactions with many binding partners is a highly polar and charged surface (Figure 2D) (*55*). Such multispecific proteins, either positively or negatively charged, select targets that display the same charge, opposite to the multispecific protein (*84*, *85*). Here, long-range electrostatic forces play a major role in the formation of the initial encounter complex with any of the electrostatically suitable partners. Then, small conformational changes might occur to optimize the orientation of the receptor protein for a particular partner. Here, short-range interactions such as hydrogen bonding, van der Waals forces, and hydrophobic interactions become important. Finally, one of the two proteins can anchor a specific side chain in a groove of the binding partner to stabilize the nativelike complex (*86*, *87*). A similar mechanism of electrostatic attraction is observed in the initial nonspecific complexes of proteins with DNA. The residues that convey specificity in the complex with the target DNA sequence interact mostly with the negatively charged phosphate groups on DNA in nonspecific DNA complexes. This results in a change of protein conformation, facilitating its rapid search for the specific DNA site (*48*) (Figure 1D).

In addition to a high surface charge and conformational flexibility, multispecificity could be achieved through the "rigid adaptation" mechanism with the protein retaining the most important chemical interactions across the binding interface. For example, anti-lysozyme antibody D 1.3 binds to two different targets, lysozyme and idiotypic antibody E5.2, in the same conformation. In both complexes, the important hydrogen bonds are structurally conserved but are conveyed through different residues on each of the two target proteins (*88*). In another case, multispecific natural killer (NK) receptor NKG2D can recognize three distantly related targets (MICA, RAE-1b, and ULBP3) without undergoing backbone conformational changes. To bind to each target, the receptor uses the same hot-spot Tyr residues as well as a number of additional residues that differ in each complex (*89*).

Conformational changes are also not necessary for binding to multiple small molecule targets. Here, recognition can be partial, achieved through imperfect complementarity between the ligand and the binding site of the protein (Figure 2E). The human pregnane X receptor (PXR), for example, detects a broad range of structurally diverse compounds using a mostly rigid framework (*90*). The binding site of this receptor contains five hot-spot residues. Different PXR ligands interact with two, three, four, and five hot spots depending on their size and shape (*91*).

Lastly, Tsai et al. suggested post-translational modifications and alternative splicing as major explanations for multispecificity (*92*).

Post-translational modifications are frequently used to alter binding affinity for different targets, allowing modulation of PP interaction networks. p53, for example, undergoes an extensive phosphorylation on S46 and T55, which leads to substantially increased affinity of this protein for its binding partners, p62 and TFB1 (*92*). Alternative splicing could lead to deletion of some domains or parts of binding sites and to changes in domain architecture. This would most probably decrease the number of interacting partners. It remains unclear, however, if deletion or rearrangement of protein domains could give rise to new binding sites, which should be an important part of evolution in multispecific binders.

DIRECTED EVOLUTION EXPERIMENTS WITH MULTISPECIFIC PROTEINS

Many recent studies demonstrated that directed evolution could mimic the natural evolution of multispecific proteins. Directed evolution experiments demonstrated that only a few mutations can readily increase promiscuous activity of enzymes by 10−1000-fold (see ref *21* for a review). So, if a promiscuous function becomes more physiologically relevant during the natural evolution, it could become a dominant function through only a small number of mutations. Interestingly, substantial enhancement of the promiscuous activity could be achieved with only a small decrease in the native activity. For example, a single mutation in D-allose kinase improved its activity for the promiscuous substrate D-glucose 62-fold, decreasing its activity toward the native substrate D-allose by only 1.25-fold (*93*). This leads to a "generalist" intermediate exhibiting both the native and the promiscuous activities at relatively high levels. In some cases, this intermediate even gains additional activities that were never selected for (*94*). These results are consistent with the hypothesis that modern enzymes diverged from parent enzymes that displayed a range of activities at a low level (*95*).

In PP interactions, promiscuous low-level activity could be also enhanced by directed evolution. This was demonstrated in colicin−immunity protein interaction (*96*). In nature, there are several closely related pairs of colicins and their inhibitors, immunity proteins. An immunity protein binds with extremely high affinity to its cognate colicin and binds with low affinity to noncognate colicins. Directed evolution and selection experiments with an immunity protein led first to a generalist intermediate capable of efficient binding to both the original and the new colicin partner and further led to the full switch in specificity (*96*).

In another example, directed evolution was used to increase the affinity of a transcriptional activator LuxR for its promiscuous signal molecule, octanoyl-HSL (C8HSL) (*97*). Only three or fewer mutations were required to create generalist LuxR variants that bind strongly to several different homoserine lactones (HSLs), including C8HSL and the native ligand, 3-oxohexanoyl-HSL (3OC6HSL). One generalist LuxR variant was further changed through one mutation into a variant that binds to C8HSL with high specificity (*98*).

The small compromise in the native activity for generalist proteins could be explained by the fact that the acquired mutations are usually located not directly at the active site of the enzyme or at the binding interface but at their periphery (second or third shells). They are hence not part of the catalytic machinery or hot spots for binding. These mutated residues typically exhibit high conformational flexibility (*93*, *99*) that might allow the formation of the generalist intermediate (*100*).

Directed evolution could also be used to generate multispecific proteins from molecules that exhibit a single specificity. For example, a combination of rational and random mutagenesis converted a low-specificity dehydrogenase into an enzyme with broad substrate specificity. This new enzyme is capable of producing various D-amino acids via the reductive amination of the corresponding 2-keto acid with ammonia (*101*). Recent work by Fuh and colleagues (*42*) reports a dual-action antibody that was optimized by phage display to recognize with high affinity two targets, human epidermal growth factor receptor 2 (HER2) and vascular endothelial growth factor (VEGF) (Figure 1C). This antibody uses mostly overlapping binding interfaces to interact with each of the targets. However, it displays slightly different conformations when bound to each target and conveys the most energetically favorable interactions with each partner by a different set of residues. This work not only presents a proof of principle that antibodies could be generated for simultaneous recognition of multiple targets but also reveals the great potential of such antibodies as future drugs.

COMPUTATIONAL DESIGN OF MULTISPECIFIC PROTEINS

Protein design predictions can offer some interesting insights into evolution of multispecific proteins and help in the dissection of molecular interactions that mediate multispecificity in binding. Some recent studies introduced multiconstrained or multistate protein design protocols for the redesign of multispecific proteins (*57*, *102*−*105*). In these works, a binding interface sequence of a given multispecific protein is first designed considering interactions with each binding partner individually in a single-state design procedure and then considering interactions with several binding partners together in a multistate design procedure (Figure 3A). The sequences generated from the two procedures are then compared to each other as well as to the native binding interface sequence of the given protein. Humphris and Kortemme (*104*) designed and analyzed sequences of 20 multispecific proteins that have available X-ray structures with two to seven binding partners. Fromer et al. (*106*) used a more sophisticated algorithm to generate not one but 100 best sequences for three multispecific proteins also used by Humphris and Kortemme (*104*). Furthermore, Fromer and Shifman (*57*) utilized the same algorithm and 16 X-ray structures of the calmodulin−target complexes to generate 100 lowest-energy calmodulin sequences for almost 700 design scenarios: calmodulin interacting with each individual target, calmodulin interacting with all possible combinations of two and three targets, and calmodulin interacting with all 16 targets. All the above protein design studies reached similar major conclusions. Optimizing a protein for interaction with only one partner generates interface sequences with the lowest interaction energy for this particular partner and high interaction energies for alternative interaction partners. On the other hand, the multistate design protocol produces protein sequences that have fairly low interaction energies with each of the considered binding partners. Most importantly, incorporating more constraints into the design procedure produces binding interface sequences closer to those found in the native multispecific interfaces, whose sequence is a compromise for interacting with all targets (Figure 3B) (*57*, *103*, *104*). Fromer and Shifman compared sequence profiles resulting from single-state and two-state designs and demonstrated several scenarios showing how the amino acid identity at each position trades off

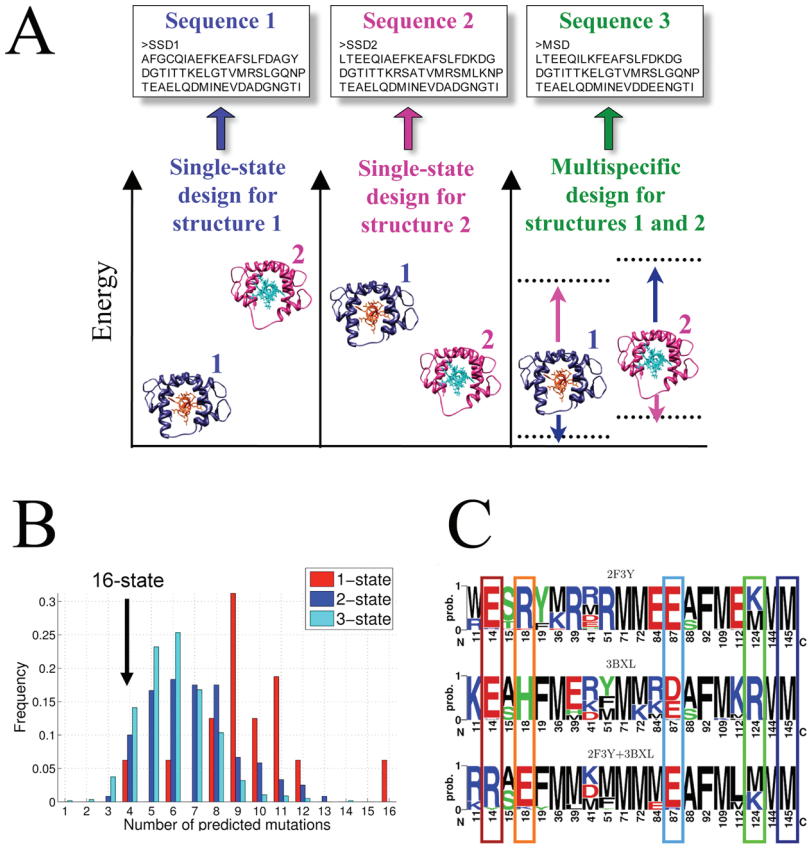

FIGURE 3: (A) Multistate design procedure. Structures 1 and 2 differ in both their backbone structures (blue vs purple) and the amino acid sequences of their binding partners (red vs cyan). Three design scenarios are shown: designing solely for structure 1, designing solely for structure 2, and simultaneously designing for both. Sequence 1 is stable only in structure 1 (binds the red ligand) and sequence 2 only in structure 2 (binds the cyan ligand), but not vice versa. On the other hand, sequence 3 is explicitly optimized to be stable in both structures (binds both ligands), yielding a multispecific protein sequence. Note that sequence 3 is not optimal for either structure, because there is some compromise of stability for structure 1 so that it can be stable in structure 2, and vice versa (downward-pointing arrows). Nevertheless, sequence 3 is significantly more stable (upward-pointing arrows) than sequence 1 for structure 2, and vice versa. Panel A reproduced from the work of Fromer et al. (*103*). Copyright 2010 Wiley-Interscience. (B) Number of mutations from the wild-type calmodulin sequence observed in the single lowest-energy sequence when redesigning 20 positions in the calmodulin binding interface in 16 calmodulin−target complexes. The average number of mutations is 9.5 for single-state designs (52% native sequence recovery), 7 for the two-state designs (65% recovery), 6 for the three-state designs (70% recovery), and 4 for the design of all 16 states (80% recovery). Panel B reproduced from ref *57*. Copyright 2009 Public Library of Science. (C) Logos of sequence profiles individually optimized in the context of the calmodulin−target complex structures with Protein Data Bank entries 2F3Y and 3BXL (one-state design), compared to the profile resulting from simultaneous optimization for interaction with both targets (two-state design). We observed five trade-off scenarios when comparing the amino acid distributions at each of the calmodulin interface positions obtained in single-state designs with those resulting from two-state design calculations. Dark blue indicates that both individual states have similar profiles and the two-state design chooses this profile. Light blue indicates that two-state design yielded a profile that is a combination of the two distributions obtained for each single-state design. Green indicates that two-state design yielded a distribution of amino acids that was similar to that of only one of the single-state designs. Orange indicates that an amino acid distribution for the two-state design was chosen that is different from that of both of the individual single-state designs. Maroon indicates that despite the individual states having similar profiles, the two-state profile is different. Interface positions are marked on the horizontal axis. Panel C reproduced from ref *57*. Copyright 2009 Public Library of Science.

upon addition of interaction constraints from an additional target (Figure 3C). The protein design studies mentioned above argue that multispecific proteins are evolved to provide reasonable but not the highest possible affinities for each of the target proteins. Affinity for each of the targets could be further enhanced, however, at the expense of the affinity for alternative targets. Experimental verification of this fact was demonstrated by our previous redesign of calmodulin (*107*, *108*). Computational enhancement of calmodulin binding specificity for a specific target was also utilized to create more efficient calmodulin-based calcium sensors for biotechnological applications (*109*).

Computational design of enzymes still presents a major challenge. Hence, design of multispecific enzymes has not been widely approached. Nevertheless, in a recent study, Suarez et al. reported a multistate computational design protocol for enhancing the promiscuous activity of an enzyme without compromising enzyme stability (*110*, *111*). They tested their approach by designing an esterase activity into a thioredoxin, introducing mutations at the periphery of the enzyme's active site (*111*). The resulting protein exhibited a 100-fold enhancement of the catalytic rate when compared to the rate of the uncatalyzed reaction and did not compromise the original function. With further development of computational methods, we expect more multispecific proteins to be rationally designed in the near future.

CONCLUDING REMARKS

Multispecificity in molecular recognition is a widely observed phenomenon, which has a different physiological role in each of the discussed groups of proteins. In PP interactions, many signaling proteins have evolved to interact with a large number of targets often with similar affinity. Their affinity for each specific target is not optimal and could be further improved. Multispecificity in

PP interactions is hence an essential part of a cell's everyday life. Multispecific antibodies and TCRs are necessary for the proper functioning of the immune system; however, the affinity of such proteins for their antigens is relatively low. This affinity could be improved through somatic mutations, however, at the expense of reducing the number of recognized antigens. Enzymes are geared to perform a single reaction, while the secondary reactions could be catalyzed with low efficiency. The promiscuous activity of enzymes is thus not a part of the organism's regular physiology but could become important under certain conditions such as a high excess of substrate or a shortage of the native enzyme. In DNA-binding proteins, promiscuous recognition is crucial to accelerating the search for the specific DNA-binding site. In addition, some DNA-binding proteins can recognize a specific set of target DNA sequences. The precise molecular mechanism for this kind of recognition has yet to be determined.

The mechanisms of multispecific recognition are similar in all the mentioned groups of proteins. Promiscuous enzymes have the ability to bind a broad range of substrates using flexible active sites and alternative binding residues. Similarly, antibodies and protein hubs frequently utilize conformational diversity to accommodate different targets. While in enzymes, TCRs, and antibodies, differences in binding modes are usually limited to small to medium-range conformational changes, in some PP interactions, conformational diversity goes to its extreme, involving folding upon binding. Additional mechanisms for multispecificity include high surface charge, posttranslational modifications, incorporation of cofactors, and rigid adaptation. Finally, several mechanisms could be utilized in tandem to achieve multispecific recognition.

Recent studies showed that multispecificity could be manipulated through directed evolution and computational protein design. In some experiments, multispecific proteins have been converted into more specific ones, yet in other experiments, the low-level promiscuous function of the protein has been amplified to yield a broad-specificity molecule. Evolution and design experiments in both directions not only unravel the molecular mechanisms of multispecificity but also assist in the creation of molecules with great potential for various biotechnological and pharmaceutical applications.

REFERENCES

1. Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. U.S.A. 98*, 4569–4574.
2. Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S., and Rothberg, J. M. (2000) A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature 403*, 623–627.
3. Giot, L., Bader, J. S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y. L., Ooi, C. E., Godwin, B., Vitols, E., Vijayadamodar, G., Pochart, P., Machineni, H., Welsh, M., Kong, Y., Zerhusen, B., Malcolm, R., Varrone, Z., Collis, A., Minto, M., Burgess, S., McDaniel, L., Stimpson, E., Spriggs, F., Williams, J., Neurath, K., Ioime, N., Agee, M., Voss, E., Furtak, K., Renzulli, R., Aanensen, N., Carrolla, S., Bickelhaupt, E., Lazovatsky, Y., DaSilva, A., Zhong, J., Stanyon, C. A., Finley, R. L., Jr., White, K. P., Braverman, M., Jarvie, T., Gold, S., Leach, M., Knight, J., Shimkets, R. A., McKenna, M. P., Chant, J., and Rothberg, J. M. (2003) A protein interaction map of *Drosophila melanogaster*. *Science 302*, 1727–1736.
4. Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S. L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., Yang, L., Wolting, C., Donaldson, I., Schandorff, S., Shewnarane, J., Vo, M., Taggart, J., Goudreault, M., Muskat, B., Alfarano, C., Dewar, D., Lin, Z., Michalickova, K., Willems, A. R., Sassi, H., Nielsen, P. A., Rasmussen, K. J., Andersen, J. R., Johansen, L. E., Hansen, L. H., Jespersen, H., Podtelejnikov, A., Nielsen, E., Crawford, J., Poulsen, V., Sorensen, B. D., Matthiesen, J., Hendrickson, R. C., Gleeson, F., Pawson, T., Moran, M. F., Durocher, D., Mann, M., Hogue, C. W., Figeys, D., and Tyers, M. (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature 415*, 180–183.
5. Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M. A., Copley, R. R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., and Superti-Furga, G. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature 415*, 141–147.
6. Gavin, A. C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L. J., Bastuck, S., Dumpelfeld, B., Edelmann, A., Heurtier, M. A., Hoffman, V., Hoefert, C., Klein, K., Hudak, M., Michon, A. M., Schelder, M., Schirle, M., Remor, M., Rudi, T., Hooper, S., Bauer, A., Bouwmeester, T., Casari, G., Drewes, G., Neubauer, G., Rick, J. M., Kuster, B., Bork, P., Russell, R. B., and Superti-Furga, G. (2006) Proteome survey reveals modularity of the yeast cell machinery. *Nature 440*, 631–636.
7. Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F. H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S., Timm, J., Mintzlaff, S., Abraham, C., Bock, N., Kietzmann, S., Goedde, A., Toksoz, E., Droege, A., Krobitsch, S., Korn, B., Birchmeier, W., Lehrach, H., and Wanker, E. E. (2005) A human protein-protein interaction network: A resource for annotating the proteome. *Cell 122*, 957–968.
8. Rual, J. F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G. F., Gibbons, F. D., Dreze, M., Ayivi-Guedehoussou, N., Klitgord, N., Simon, C., Boxem, M., Milstein, S., Rosenberg, J., Goldberg, D. S., Zhang, L. V., Wong, S. L., Franklin, G., Li, S., Albala, J. S., Lim, J., Fraughton, C., Llamosas, E., Cevik, S., Bex, C., Lamesch, P., Sikorski, R. S., Vandenhaute, J., Zoghbi, H. Y., Smolyar, A., Bosak, S., Sequerra, R., Doucette-Stamm, L., Cusick, M. E., Hill, D. E., Roth, F. P., and Vidal, M. (2005) Towards a proteome-scale map of the human protein-protein interaction network. *Nature 437*, 1173–1178.
9. Barabasi, A. L., and Albert, R. (1999) Emergence of scaling in random networks. *Science 286*, 509–512.
10. Patil, A., Kinoshita, K., and Nakamura, H. (2010) Hub promiscuity in protein-protein interaction networks. *Int. J. Mol. Sci. 11*, 1930–1943.
11. Jeong, H., Mason, S. P., Barabasi, A. L., and Oltvai, Z. N. (2001) Lethality and centrality in protein networks. *Nature 411*, 41–42.
12. Fraser, H. B., Hirsh, A. E., Steinmetz, L. M., Scharfe, C., and Feldman, M. W. (2002) Evolutionary rate in the protein interaction network. *Science 296*, 750–752.
13. Jordan, I. K., Wolf, Y. I., and Koonin, E. V. (2003) No simple dependence between protein evolution rate and the number of protein-protein interactions: Only the most prolific interactors tend to evolve slowly. *BMC Evol. Biol. 3*, 1.
14. Wuchty, S. (2004) Evolution and topology in the yeast protein interaction network. *Genome Res. 14*, 1310–1314.
15. Kim, P. M., Lu, L. J., Xia, Y., and Gerstein, M. B. (2006) Relating three-dimensional structures to protein networks provides evolutionary insights. *Science 314*, 1938–1941.
16. Valdar, W. S. J., and Thornton, J. M. (2001) Protein-protein interfaces: Analysis of amino acid conservation in homodimers. *Proteins 42*, 108–124.
17. Copley, S. D. (2003) Enzymes with extra talents: Moonlighting functions and catalytic promiscuity. *Curr. Opin. Chem. Biol. 7*, 265–272.
18. Hult, K., and Berglund, P. (2007) Enzyme promiscuity: Mechanism and applications. *Trends Biotechnol. 25*, 231–238.
19. Nobeli, I., Favia, A. D., and Thornton, J. M. (2009) Protein promiscuity and its implications for biotechnology. *Nat. Biotechnol. 27*, 157–167.
20. Babtie, A., Tokuriki, N., and Hollfelder, F. (2010) What makes an enzyme promiscuous? *Curr. Opin. Chem. Biol. 14*, 200–207.
21. Khersonsky, O., and Tawfik, D. S. (2010) Enzyme promiscuity: A mechanistic and evolutionary perspective. *Annu. Rev. Biochem. 79*, 471–505.
22. D'Ari, R., and Casadesus, J. (1998) Underground metabolism. *BioEssays 20*, 181–186.

23. Kim, J., and Copley, S. D. (2007) Why metabolic enzymes are essential or nonessential for growth of *Escherichia coli* k12 on glucose. *Biochemistry 46*, 12501–12511.
24. Patrick, W. M., Quandt, E. M., Swartzlander, D. B., and Matsumura, I. (2007) Multicopy suppression underpins metabolic evolvability. *Mol. Biol. Evol. 24*, 2716–2722.
25. de Groot, M. J. L., van Berlo, R. J. P., van Winden, W. A., Verheijen, P. J. T., Reinders, M. J. T., and de Ridder, D. (2009) Metabolite and reaction inference based on enzyme specificities. *Bioinformatics 25*, 2975–2982.
26. Macchiarulo, A., Nobeli, I., and Thornton, J. M. (2004) Ligand selectivity and competition between enzymes in silico. *Nat. Biotechnol. 22*, 1039–1045.
27. Favia, A. D., Nobeli, I., Glaser, F., and Thornton, J. M. (2008) Molecular docking for substrate identification: The short-chain dehydrogenases/reductases. *J. Mol. Biol. 375*, 855–874.
28. Hermann, J. C., Ghanem, E., Li, Y. C., Raushel, F. M., Irwin, J. J., and Shoichet, B. K. (2006) Predicting substrates by docking high-energy intermediates to enzyme structures. *J. Am. Chem. Soc. 128*, 15882–15891.
29. Hermann, J. C., Marti-Arbona, R., Fedorov, A. A., Fedorov, E., Almo, S. C., Shoichet, B. K., and Raushel, F. M. (2007) Structure-based activity prediction for an enzyme of unknown function. *Nature 448*, 775–779.
30. Cameron, D. J., and Erlanger, B. F. (1977) Evidence for multi-specificity of antibody molecules. *Nature 268*, 763–765.
31. Richards, F. F., Konigsberg, W. H., Rosenstein, R. W., and Varga, J. M. (1975) On the specificity of antibodies. *Science 187*, 130–137.
32. Mason, D. (1998) A very high level of crossreactivity is an essential feature of the T-cell receptor. *Immunol. Today 19*, 395–404.
33. Wing, M. G. (1995) The molecular basis for a polyspecific antibody. *Clin. Exp. Immunol. 99*, 313–315.
34. Wucherpfennig, K. W., Gagnon, E., Call, M. J., Huseby, E. S., and Call, M. E. (2010) Structural biology of the T-cell receptor: Insights into receptor assembly, ligand recognition, and initiation of signaling. *Cold Spring Harbor Perspect. Biol. 2*, a005140.
35. Wilson, D. B., Wilson, D. H., Schroder, K., Pinilla, C., Blondelle, S., Houghten, R. A., and Garcia, K. C. (2004) Specificity and degeneracy of T cells. *Mol. Immunol. 40*, 1047–1055.
36. Kranz, D. M. (2000) Incompatible differences: View of an allogeneic pMHC-TCR complex. *Nat. Immunol. 1*, 277–278.
37. Chow, D. A., and Bennet, R. D. (1989) Low natural antibody and low in vivo tumor resistance, in xid-bearing B-cell deficient mice. *J. Immunol. 142*, 3702–3706.
38. Sethi, D. K., Agarwal, A., Manivel, V., Rao, K. V. S., and Salunke, D. M. (2006) Differential Epitope Positioning within the Germline Antibody Paratope Enhances Promiscuity in the Primary Immune Response. *Immunity 24*, 429–438.
39. Zimmermann, J., Oakman, E. L., Thorpe, I. F., Shi, X., Abbyad, P., Brooks, C. L., III, Boxer, S. G., and Romesberg, F. E. (2006) Antibody evolution constrains conformational heterogeneity by tailoring protein dynamics. *Proc. Natl. Acad. Sci. U.S.A. 103*, 13722–13727.
40. Roggenbuck, D., Konig, H., Niemann, B., Schoenherr, G., Jahn, S., and Porstmann, T. (1994) Real-time biospecific interaction analysis of a natural human polyreactive monoclonal IgM antibody and its Fab and scFv fragments with several antigens. *Scand. J. Immunol. 40*, 64–70.
41. Poncet, P., Matthes, T., Billecocq, A., and Dighiero, G. (1988) Immunochemical studies of polyspecific natural autoantibodies: Charge, lipid reactivity, Fab′2 fragments activity and complement fixation. *Mol. Immunol. 25*, 981–989.
42. Bostrom, J., Yu, S. F., Kan, D., Appleton, B. A., Lee, C. V., Billeci, K., Man, W., Peale, F., Ross, S., Wiesmann, C., and Fuh, G. (2009) Variants of the Antibody Herceptin That Interact with HER2 and VEGF at the Antigen Binding Site. *Science 323*, 1610–1614.
43. Mariuzza, R. A. (2006) Multiple paths to multispecificity. *Immunity 24*, 359–361.
44. Badis, G., Berger, M. F., Philippakis, A. A., Talukder, S., Gehrke, A. R., Jaeger, S. A., Chan, E. T., Metzler, G., Vedenko, A., Chen, X., Kuznetsov, H., Wang, C.-F., Coburn, D., Newburger, D. E., Morris, Q., Hughes, T. R., and Bulyk, M. L. (2009) Diversity and Complexity in DNA Recognition by Transcription Factors. *Science 324*, 1720–1723.
45. Bobay, B. G., Benson, L., Naylor, S., Feeney, B., Clark, A. C., Goshe, M. B., Strauch, M. A., Thompson, R., and Cavanagh, J. (2004) Evaluation of the DNA binding tendencies of the transition state regulator AbrB. *Biochemistry 43*, 16106–16118.
46. Ishii, A., and Hihara, Y. (2008) An AbrB-like transcriptional regulator, Sll0822, is essential for the activation of nitrogen-regulated genes in *Synechocystis* sp. PCC 6803. *Plant Physiol. 148*, 660–670.
47. Sullivan, D. M., Bobay, B. G., Kojetin, D. J., Thompson, R. J., Rance, M., Strauch, M. A., and Cavanagh, J. (2008) Insights into the Nature of DNA Binding of AbrB-like Transcription Factors. *Structure 16*, 1702–1713.
48. Kalodimos, C. G., Biris, N., Bonvin, A., Levandoski, M. M., Guennuegues, M., Boelens, R., and Kaptein, R. (2004) Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science 305*, 386–389.
49. Vonhippel, P. H., and Berg, O. G. (1989) Facilitates target location in biological systems. *J. Biol. Chem. 264*, 675–678.
50. Halford, S. E., and Marko, J. F. (2004) How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res. 32*, 3040–3052.
51. Viadiu, H., and Aggarwal, A. K. (2000) Structure of BamHI bound to nonspecific DNA: A model for DNA sliding. *Mol. Cell 5*, 889–895.
52. Townson, S. A., Samuelson, J. C., Bao, Y. M., Xu, S. Y., and Aggarwal, A. K. (2007) BstYI bound to noncognate DNA reveals a "hemispecific" complex: Implications for DNA scanning. *Structure 15*, 449–459.
53. Walter, J., Trautner, T. A., and Noyer-Weidner, M. (1992) High plasticity of multispecific DNA methyltransferases in the region carrying DNA target recognizing enzyme modules. *EMBO J. 11*, 4445–4450.
54. James, L. C., and Tawfik, D. S. (2003) Conformational diversity and protein evolution: A 60-year-old hypothesis revisited. *Trends Biochem. Sci. 28*, 361–368.
55. Patil, A., and Nakamura, H. (2007) The role of charged surface residues in the binding ability of small hubs in protein-protein interaction networks. *Biophysics 3*, 27.
56. Iwahara, J., Zweckstetter, M., and Clore, G. M. (2006) NMR structural and kinetic characterization of a homeodomain diffusing and hopping on nonspecific DNA. *Proc. Natl. Acad. Sci. U.S.A. 103*, 15062–15067.
57. Fromer, M., and Shifman, J. M. (2009) Tradeoff between stability and multispecificity in the design of promiscuous proteins. *PLoS Comput. Biol. 5*, e1000627.
58. Lange, O. F., Lakomek, N. A., Fares, C., Schroder, G. F., Walter, K. F., Becker, S., Meiler, J., Grubmuller, H., Griesinger, C., and de Groot, B. L. (2008) Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science 320*, 1471–1475.
59. James, L. C., Roversi, P., and Tawfik, D. S. (2003) Antibody multispecificity mediated by conformational diversity. *Science 299*, 1362–1367.
60. Armstrong, K. M., Piepenbrink, K. H., and Baker, B. M. (2008) Conformational changes and flexibility in T-cell receptor recognition of peptide-MHC complexes. *Biochem. J. 415*, 183–196.
61. Patil, A., and Nakamura, H. (2006) Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. *FEBS Lett. 580*, 2041–2045.
62. Kim, P. M., Sboner, A., Xia, Y., and Gerstein, M. (2008) The role of disorder in interaction networks: A structural analysis. *Mol. Syst. Biol. 4*, 179.
63. 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., and Obradovic, Z. (2001) Intrinsically disordered protein. *J. Mol. Graphics Modell. 19*, 26–59.
64. Uversky, V. N., Oldfield, C. J., and Dunker, A. K. (2005) Showing your ID: Intrinsic disorder as an ID for recognition, regulation and cell signaling. *J. Mol. Recognit. 18*, 343–384.
65. Obradovic, Z., Peng, K., Vucetic, S., Radivojac, P., Brown, C. J., and Dunker, A. K. (2003) Predicting intrinsic disorder from amino acid sequence. *Proteins 53* (Suppl. 6), 566–572.
66. Dunker, A. K., Cortese, M. S., Romero, P., Iakoucheva, L. M., and Uversky, V. N. (2005) Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J. 272*, 5129–5148.
67. Wright, P. E., and Dyson, H. J. (1999) Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol. 293*, 321–331.
68. Gsponer, J., Futschik, M. E., Teichmann, S. A., and Babu, M. M. (2008) Tight regulation of unstructured proteins: From transcript synthesis to protein degradation. *Science 322*, 1365–1368.
69. Lando, D., Peet, D. J., Whelan, D. A., Gorman, J. J., and Whitelaw, M. L. (2002) Asparagine hydroxylation of the HIF transactivation domain: A hypoxic switch. *Science 295*, 858–861.
70. Elkins, J. M., Hewitson, K. S., McNeill, L. A., Seibel, J. r. F., Schlemminger, I., Pugh, C. W., Ratcliffe, P. J., and Schofield, C. J. (2003) Structure of Factor-inhibiting Hypoxia-inducible Factor

(HIF) Reveals Mechanism of Oxidative Modification of HIF-1α. *J. Biol. Chem. 278*, 1802–1806.
71. Merkley, N., and Shaw, G. S. (2004) Solution structure of the flexible class II ubiquitin-conjugating enzyme Ubc1 provides insights for polyubiquitin chain assembly. *J. Biol. Chem. 279*, 47139–47147.
72. Wilson, M. A., and Brunger, A. T. (2000) The 1.0 Å crystal structure of Ca^{2+}-bound calmodulin: An analysis of disorder and implications for functionally relevant plasticity. *J. Mol. Biol. 301*, 1237–1256.
73. Goodman, R. H., and Smolik, S. (2000) CBP/p300 in cell growth, transformation, and development. *Genes Dev. 14*, 1553–1577.
74. Parker, D., Rivera, M., Zor, T., Henrion-Caude, A., Radhakrishnan, I., Kumar, A., Shapiro, L. H., Wright, P. E., Montminy, M., and Brindle, P. K. (1999) Role of secondary structure in discrimination between constitutive and inducible activators. *Mol. Cell. Biol. 19*, 5601–5607.
75. Radhakrishnan, I., Perez-Alvarado, G. C., Parker, D., Dyson, H. J., Montminy, M. R., and Wright, P. E. (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: A model for activator:coactivator interactions. *Cell 91*, 741–752.
76. Kim, T. D., Paik, S. R., and Yang, C. H. (2002) Structural and functional implications of C-terminal regions of α-synuclein. *Biochemistry 41*, 13782–13790.
77. Park, S. M., Jung, H. Y., Kim, T. D., Park, J. H., Yang, C. H., and Kim, J. (2002) Distinct roles of the N-terminal-binding domain and the C-terminal-solubilizing domain of α-synuclein, a molecular chaperone. *J. Biol. Chem. 277*, 28512–28520.
78. Bhattacharyya, J., and Das, K. P. (1999) Molecular chaperone-like properties of an unfolded protein, α(s)-casein. *J. Biol. Chem. 274*, 15505–15509.
79. Smulders, R., Carver, J. A., Lindner, R. A., van Boekel, M. A., Bloemendal, H., and de Jong, W. W. (1996) Immobilization of the C-terminal extension of bovine αA-crystallin reduces chaperone-like activity. *J. Biol. Chem. 271*, 29060–29066.
80. Pasta, S. Y., Raman, B., Ramakrishna, T., and Rao, Ch. M. (2002) Role of the C-terminal extensions of α-crystallins. Swapping the C-terminal extension of α-Crystallin to αB-Crystallin results in enhanced chaperone activity. *J. Biol. Chem. 277*, 45821–45828.
81. Lindner, R. A., Carver, J. A., Ehrnsperger, M., Buchner, J., Esposito, G., Behlke, J., Lutsch, G., Kotlyarov, A., and Gaestel, M. (2000) Mouse Hsp25, a small shock protein. The role of its C-terminal extension in oligomerization and chaperone action. *Eur. J. Biochem. 267*, 1923–1932.
82. Tompa, P., and Csermely, P. (2004) The role of structural disorder in the function of RNA and protein chaperones. *FASEB J. 18*, 1169–1175.
83. Ikura, M., and Ames, J. B. (2006) Genetic polymorphism and protein conformational plasticity in the calmodulin superfamily: Two ways to promote multifunctionality. *Proc. Natl. Acad. Sci. U.S.A. 103*, 1159–1164.
84. Friedler, A., Veprintsev, D. B., Rutherford, T., von Glos, K. I., and Fersht, A. R. (2005) Binding of Rad51 and other peptide sequences to a promiscuous, highly electrostatic binding site in p53. *J. Biol. Chem. 280*, 8051–8059.
85. Sheinerman, F. B., Norel, R., and Honig, B. (2000) Electrostatic aspects of protein-protein interactions. *Curr. Opin. Struct. Biol. 10*, 153–159.
86. Schreiber, G., and Fersht, A. R. (1996) Rapid, electrostatically assisted association of proteins. *Nat. Struct. Biol. 3*, 427–431.
87. Rajamani, D., Thiel, S., Vajda, S., and Camacho, C. J. (2004) Anchor residues in protein-protein interactions. *Proc. Natl. Acad. Sci. U.S.A. 101*, 11287–11292.
88. Fields, B. A., Goldbaum, F. A., Ysern, X., Poljak, R. J., and Mariuzza, R. A. (1995) Molecular basis of antigen mimicry by an anti-idiotope. *Nature 374*, 739–742.
89. McFarland, B. J., Kortemme, T., Yu, S. F., Baker, D., and Strong, R. K. (2003) Symmetry recognizing asymmetry: Analysis of the interactions between the C-type lectin-like immunoreceptor NKG2D and MHC class I-like ligands. *Structure 11*, 411–422.
90. Watkins, R. E., Wisely, G. B., Moore, L. B., Collins, J. L., Lambert, M. H., Williams, S. P., Willson, T. M., Kliewer, S. A., and Redinbo, M. R. (2001) The human nuclear xenobiotic receptor PXR: Structural determinants of directed promiscuity. *Science 292*, 2329–2333.
91. Ngan, C.-H., Beglov, D., Rudnitskaya, A. N., Kozakov, D., Waxman, D. J., and Vajda, S. (2009) The Structural Basis of Pregnane X Receptor Binding Promiscuity. *Biochemistry 48*, 11572–11581.
92. Tsai, C. J., Ma, B., and Nussinov, R. (2009) Protein-protein interaction networks: How can a hub protein bind so many different partners? *Trends Biochem. Sci. 34*, 594–600.
93. Larion, M., Moore, L. B., Thompson, S. M., and Miller, B. G. (2007) Divergent evolution of function in the ROK sugar kinase superfamily: Role of enzyme loops in substrate specificity. *Biochemistry 46*, 13564–13572.
94. Matsumura, I., and Ellington, A. D. (2001) In vitro evolution of β-glucuronidase into a β-galactosidase proceeds through non-specific intermediates. *J. Mol. Biol. 305*, 331–339.
95. Jensen, R. A. (1976) Enzyme Recruitment in Evolution of New Function. *Annu. Rev. Microbiol. 30*, 409–425.
96. Levin, K. B., Dym, O., Albeck, S., Magdassi, S., Keeble, A. H., Kleanthous, C., and Tawfik, D. S. (2009) Following evolutionary paths to protein-protein interactions with high affinity and selectivity. *Nat. Struct. Mol. Biol. 16*, 1049–1067.
97. Collins, C. H., Arnold, F. H., and Leadbetter, J. R. (2005) Directed evolution of *Vibrio fischeri* LuxR for increased sensitivity to a broad spectrum of acyl-homoserine lactones. *Mol. Microbiol. 55*, 712–723.
98. Collins, C. H., Leadbetter, J. R., and Arnold, F. H. (2006) Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR. *Nat. Biotechnol. 24*, 708–712.
99. Aharoni, A., Gaidukov, L., Khersonsky, O., Gould, S. M., Roodveldt, C., and Tawfik, D. S. (2005) The 'evolvability' of promiscuous protein functions. *Nat. Genet. 37*, 73–76.
100. Tokuriki, N., and Tawfik, D. S. (2009) Protein Dynamism and Evolvability. *Science 324*, 203–207.
101. Vedha-Peters, K., Gunawardana, M., Rozzell, J. D., and Novick, S. J. (2006) Creation of a broad-range and highly stereoselective D-amino acid dehydrogenase for the one-step synthesis of D-amino acids. *J. Am. Chem. Soc. 128*, 10923–10929.
102. Allen, B. D., and Mayo, S. L. (2010) An efficient algorithm for multistate protein design based on FASTER. *J. Comput. Chem. 31*, 904–916.
103. Fromer, M., Yanover, C., and Linial, M. (2010) Design of multi-specific protein sequences using probabilistic graphical modeling. *Proteins 78*, 530–547.
104. Humphris, E. L., and Kortemme, T. (2007) Design of multi-specificity in protein interfaces. *PLoS Comput. Biol. 3*, e164.
105. Yanover, C., Fromer, M., and Shifman, J. M. (2007) Dead-end elimination for multistate protein design. *J. Comput. Chem. 28*, 2122–2129.
106. Fromer, M., Yanover, C., Harel, A., Shachar, O., Weiss, Y., and Linial, M. (2010) SPRINT: Side-chain Prediction Inference Toolbox for Multistate Protein Design. *Bioinformatics 26*, 2466–2467.
107. Shifman, J. M., and Mayo, S. L. (2003) Exploring the origins of binding specificity through the computational redesign of calmodulin. *Proc. Natl. Acad. Sci. U.S.A. 100*, 13274–13279.
108. Yosef, E., Politi, R., Choi, M. H., and Shifman, J. M. (2009) Computational design of calmodulin mutants with up to 900-fold increase in binding specificity. *J. Mol. Biol. 385*, 1470–1480.
109. Palmer, A. E., Giacomello, M., Kortemme, T., Hires, S. A., Lev-Ram, V., Baker, D., and Tsien, R. Y. (2006) Ca^{2+} indicators based on computationally redesigned calmodulin-peptide pairs. *Chem. Biol. 13*, 521–530.
110. Suarez, M., Tortosa, P., Carrera, J., and Jaramillo, A. (2008) Pareto optimization in computational protein design with multiple objectives. *J. Comput. Chem. 29*, 2704–2711.
111. Suarez, M., Tortosa, P., Garcia-Mira, M. M., RodrÌguez-Larrea, D., Godoy-Ruiz, R., Ibarra-Molero, B., Sanchez-Ruiz, J. M., and Jaramillo, A. (2010) Using multi-objective computational design to extend protein promiscuity. *Biophys. Chem. 147*, 13–19.
112. DeLano, W. L. (2008) The PyMOL Molecular Graphics System, DeLano Scientific, Palo Alto, CA.