Published in final edited form as:
Curr Opin Struct Biol. 2011 February ; 21(1): 50–61. doi:10.1016/j.sbi.2010.10.002.

Protein Binding Specificity versus Promiscuity

Gideon Schreiber[1] and **Amy E. Keating**[2]

Gideon Schreiber: gideon.schreiber@weizmann.ac.il; Amy E. Keating: keating@mit.edu

[1] Department of Biological Chemistry, Weizmann Institute of Science, Rehovot, 76100, Israel

[2] Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139

Abstract

Interactions between macromolecules in general, and between proteins in particular, are essential for any life process. Examples include transfer of information, inhibition or activation of function, molecular recognition as in the immune system, assembly of macromolecular structures and molecular machines, and more. Proteins interact with affinities ranging from millimolar to femtomolar and, because affinity determines the concentration required to obtain 50% binding, the amount of different complexes formed is very much related to local concentrations. Although the concentration of a specific binding partner is usually quite low in the cell (nanomolar to micromolar), the total concentration of other macromolecules is very high, allowing weak and non-specific interactions to play important roles. In this review we address the question of binding specificity, i.e., how do some proteins maintain monogamous relations while others are clearly polygamous. We examine recent work that addresses the molecular and structural basis for specificity versus promiscuity. We show through examples how multiple solutions exist to achieve binding via similar interfaces and how protein specificity can be tuned using both positive and negative selection (specificity by demand). Binding of a protein to numerous partners can be promoted through variation in which residues are used for binding, conformational plasticity and/or post-translational modification. Natively unstructured regions represent the extreme case in which structure is obtained only upon binding. Many natively unstructured proteins serve as hubs in protein-protein interaction networks and such promiscuity can be of functional importance in biology.

Introduction

In an organism, proteins can participate in specific interactions with just one or a few partners, in promiscuous yet functional interactions with many partners, and/or in non-specific interactions with some of the numerous functionally non-cognate partners. In the cell, 30% of the dry mass is composed of proteins [1]. It has been shown that chymotrypsin inhibitor 2 translational and rotational diffusion rates in cell extracts *in vitro* were hindered by weak, non-specific interactions [2*]. The source of interaction specificity that favors a small set of interactions over the multitude of possibilities is not well understood. Specificity involves both binding to a specific partner and *not* binding to other proteins. When defining specificity one has first to define “binding”. The simplest definition would be to use some arbitrary affinity threshold. However, this is not advisable, as functionally important binding

occurs at a range of affinities from low millimolar to femtomolar. A different definition would be to relate specificity to the concentrations and compartmentalization of the proteins in question in the cell, e.g. requiring that for biologically relevant binding, two proteins must be localized near one another at a concentration that promotes interaction. In other words, specificity is a relative trait that is context-dependent. But critical information about the relevant cellular conditions is often not available. Nevertheless, significant progress has been made towards understanding specificity by studying proteins under more controlled conditions, using biochemical and biophysical methods. Varying criteria have been used to define binding and binding specificity, e.g. designating a protein as “specific” if interaction with a desired partner is tighter than with other proteins, without considering the energy gap. This is because quantitative binding affinities and/or complete specificity profiles are often not available.

The challenge of achieving specificity is greater when candidate interaction partners are similar in sequence and/or structure. Large, paralogous gene families pose this type of problem, as does the selection of conformationally specific antibodies and the design of targeted biological therapeutics. A tradeoff between affinity and specificity of binding to similar interfaces has been suggested, but no universal relationship between these properties has been established [3]. Very tight binding to a specific partner may be one mechanism, but another involves explicit negative design elements that suppress cross-interactions, which may be an evolutionary trait. A related issue is whether it is important to consider negative design in protein engineering, and on this point the answer seems to vary.

Multi-specificity is the property of interacting with many partners, and this can be important for biological function. Protein interactome studies have identified “hub” proteins that participate in exceptionally high numbers of interactions, and multi-specificity is common for many proteins involved in signaling and regulation. A range of structural strategies to achieve multi-specificity has been observed. At one extreme, there are proteins for which numerous interactions can occur via structurally similar complexes that exhibit negligible to small, though important, variation in different instances. More commonly, promiscuous binding seems to involve degrees of structural plasticity, which may result in different subsets of residues being important for binding to different partners. More extreme examples of structural plasticity are found in natively unstructured proteins that can adopt dramatically different structures in different complexes. Finally, post-translational modifications can alter the chemistry and structure of the same sequence in its interactions with diverse partners. Structural and post-translational changes make the problems of computationally predicting protein interactions, protein complex structures, and interaction hotspots fiendishly difficult.

The evolution of protein interfaces and the engineering of new partners are related problems. Key to each is the observation that there are many solutions to the interaction problem for a given protein. Clearly, some sites on structured proteins are more suitable than others for mediating interactions, and a few of the biophysical properties that distinguish good interfaces from bad are understood [4,5]. But once a protein has evolved a good binding site, it appears that there are essentially innumerable ways to use it in different variations. For example, even within the context of a fixed backbone scaffold, computational protein design algorithms often find many different solutions predicted to complement a known interface, particularly when both sides of the interface can be changed. In computational methods that do not keep the template fixed, it is possible to swap interaction modules in and out to diversify the binding properties. *In vitro* selection is an excellent way to identify new interaction partners for a given protein, and large numbers of distinct ligands, which can use a common binding site in somewhat different ways, have been discovered this way. Over the course of evolution, interfaces can change in concert via a process of co-evolution, and in

this way diverge from homologous structures to find nearby solutions that satisfy requirements on function in new ways [6].

Another way in which a single protein can exhibit diverse solutions to the interaction problem is through participation in numerous complexes of unrelated structure. In natural systems, evolutionarily related proteins can form structurally dissimilar interactions with diverse partners using the same binding site. Studying this phenomenon, Martin recently found little evidence for convergence to chemically similar binding solutions at such interfaces: many distinct physical solutions seem to exist [7*]. From either an evolutionary or engineering perspective, intrinsically disordered proteins provide great potential for forming many interactions via structurally dissimilar complexes, as residues that are key for one interaction can often be decoupled from those important for others. Stringing a number of such semi-autonomous motifs into a disordered region provides tremendous interaction capacity, in a modular solution to the multi-specificity problem.

In this review we highlight advances in our understanding of protein-binding specificity, and how this relates to biological function. We demonstrate the principle of “specificity by demand”, which shows that the specificity of an interface is tunable. We discuss examples of positive versus negative design principles, and show how flexibility and post-translational modifications can result in a multitude of binding on the same protein domain, using different residues for interaction (and as “hotspots”) for different partners. The potential for promiscuity stems from the principle of many sequence and structure combinations that will promote binding to a given protein.

Multiple interactions via a conserved structure

Nature uses structurally conserved domains and motifs to mediate many interactions. When such domains occur multiple times in an organism, or across organisms, an opportunity for intra-family multi-specificity arises, because one protein can bind to multiple evolutionarily related partners. It is interesting to consider the commonalities and differences between largely similar complexes that determine the affinity and specificity of interactions. Elements of negative design, predicted to form bad interactions in non-cognate complexes, can be identified in many of these examples.

β-lactamase - β-lactamase inhibitor proteins

β-lactamases can be inhibited by a variety of proteins that show varying interaction specificities. The structures of five different β-lactamase - β-lactamase inhibitor protein complexes have been solved; TEM1-BLIP (1JTG), TEM1-BLIP1 (3GMW), TEM1-BLIP2 (1JTD), SHV – BLIP (2G2U) and KPC2 – BLIP (3E2L) [8–12]. Except for BLIP2, a significant degree of homology exits between the proteins (Table 1). The different BLIP and β-lactamase proteins bind one another at the same location and orientation (Fig 1) using a similar set of amino acids to form the interaction. In addition, BLIP-like protein BLP ((3GMX), 32% and 42% identity to BLIP and BLIP1, Table 1) does not bind to β-lactamase. These complexes serve as an excellent test case to investigate whether the same interactions are used in the different complexes, and whether hotspot residues are conserved. The residues conserved in all five binding sites are listed in Table 2, and these constitute a large percentage of the interface residues. However, despite conservation in location (even for TEM1 binding to BLIP2) the interface residue identities are not highly conserved. Moreover, different complexes use different residues as hotspots for binding, promoting specificity of BLIPs for four different β-lactamase proteins [13]. Even if comparing only four out of the five complexes (without TEM1-BLIP2) only two conserved sidechain-sidechain interactions (between K111 and S39 and S130 with D39) and three conserved mainchain-to-sidechain interactions are found. Thus, the analysis of these five interfaces

teaches us that although the location of the interface is highly conserved, the identity of the pairwise interactions is not. In other words, there are many solutions to obtain micromolar-to-nanomolar affinity for this interaction.

It is interesting that structure is much better conserved than sequence in this example (Fig 1). As this is not due to contacts made, and apparently not due to functional restraints (one could inhibit β-lactamase activity with less structure conservation at the interface), one may speculate that certain structural characteristics promote binding, while others do not [4]. What these characteristics are is currently not well understood, but may be important for interface design. One BLIP like protein (BLP) does not bind β-lactamase. Here, the reason is clearly a totally different sequence of the interface residues, which does not promote binding [11*].

Colicin-immunity proteins

Interactions between colicin endonucleases (DNase) and immunity proteins provide an example where cognate complexes are much more stable than non-cognate complexes, largely because of a much slower rate of dissociation resulting in six orders of magnitude difference in affinity. A recent structure of a low-affinity non-cognate complex showed only small differences between the cognate and non-cognate interfaces, suggesting that a few frustrated contacts (a form of negative design) are all that is needed to produce a large effect on binding [14*]. Iterative rounds of random mutagenesis and selection toward higher affinity of the non-cognate ColE7-Im9 pair uncovered a latent set of interactions, providing the key to the rapid divergence of Im protein variants with altered specificity [15*]. This suggests that protein-protein interactions can exploit promiscuous interactions and alternative binding configurations to facilitate the divergence of new functions, and also that specificity can be introduced following gene duplication with just a few changes in sequence.

Bcl-2 proteins

Interactions among Bcl-2 proteins are important for regulating programmed cell death. In humans, 6 globular anti-apoptotic Bcl-2 family members can bind a variety of short helical-peptide segments in “BH3-only” pro-apoptotic proteins, thereby blocking apoptosis. Sequence similarity among the anti-apoptotic receptors is low to moderate, but these proteins, as well as several viral homologs, have highly conserved structures. Pro-apoptotic BH3-only proteins appear to exhibit “specificity by demand”, i.e. the various family members have different binding specificities, presumably associated with their distinct functions. Some BH3-only proteins (e.g. Bim, Puma, Bid) bind to anti-apoptotic human paralogs promiscuously, whereas others (e.g. Bad, Noxa) selectively interact with just a subset. The anti-apoptotic proteins conversely have distinct specificity patterns, which has been exploited to profile the state of primary tumor isolates to determine which Bcl-2 proteins are involved in cell survival [16].

Both structure-based and systematic mutational studies have probed the significance of many residues for Bcl-2 family binding and interaction specificity [17–20]. Native Bcl-2 family complexes are very stable, with low nanomolar dissociation constants. Although the overall structure and interaction mode of different complexes is highly conserved, modest structural plasticity plays a role in recognition on both the anti-apoptotic and pro-apoptotic sides of the interface [21]. The plasticity of the interface accommodates small-to-large, hydrophobic-to-charged and alanine mutations at many sites with only minor changes in binding affinity, consistent with the ability of many Bcl-2 family receptors to bind a wide diversity of BH3-only proteins [22–26]. Certain interface sites are more restrictive, however, providing a mechanism for negative design against non-cognate complexes. Single-residue

or two-residue mutations of this type have been identified that give dramatic changes in BH3-only specificity [18*,19]. Many of these can be rationalized structurally [21], and a sequence-based model has been developed based on available data that shows good prediction performance in a limited sequence space [18*]. Directed studies have provided engineered peptides that display desired specificities [18*,20,23,27]. But the importance of even the relatively modest structural variation observed makes model building a challenge.

Antibody-protein Antigen

Antibodies have a structural scaffold that allows for an essentially unlimited diversity of binding partners. Naïve antibodies have tremendous capacity to bind different eptiopes, and the plasticity provided by flexible loops on a dynamic scaffold is critical to this. Antibodies undergo a complex maturation and clonal selection process against specific epitopes on the antigen, which can be associated with rigidification. The maturation process drives low affinity and low specificity recognition towards high affinity and specificity. However, affinity and specificity of antibodies do not have to be directly related. E.g., in the absence of negative selection, the same mechanism that provides higher affinity for one epitope (like a more rigid binding site with a certain composition of amino acids) may also alter promiscuity to similar molecules [3]. Thus, to achieve high specificity, a high affinity monoclonal antibody may not be preferable over lower affinity polyclonal antibodies, as promiscuous binding of the higher affinity molecule may be enhanced as well. For polyclonal antibodies, although promiscuous binding may be distributed over different molecules, specific binding can be enhanced by the repertoire, making the polyclonal antibodies effectively more specific [28].

The question of antibody selectivity was systematically evaluated by Michaud et al. by analyzing the binding of 11 antibodies raised against different antigens on a yeast protein array [29]. They found that five of the antibodies showed specific recognition towards their antigen, five others were cross reactive towards a number of other antigens, and one was not specific, binding >1000 partners. Similar results were reported more recently using an array of human proteins produced recombinantly in E. coli [30]. Although such experiments provide valuable information, one should remember that many of the proteins immobilized on microarrays are not properly folded, and may differ in posttranslational modifications from native counterparts. In the body one has also to consider the local concentrations of the antibodies and antigens, which may significantly alter the binding profile. Yet overall, the suggestion is that many antibodies exhibit promiscuous interactions, perhaps associated with antibody binding sites being “good” interfaces.

Birtalan et al. performed a comprehensive experimental analysis to relate antibody specificity to sequence by testing the functional capacity of a single Fab framework for binding VEGF, HER2, IGF-1 and protein A using phage display [31**]. This study provided insights into what makes a good binding site. For all antigens, Tyr was the preferred amino acid for contributing to both affinity and specificity. This is in line with the prevalence of Tyr in native antibody binding sites as well as with previous work on artificial antibody libraries [32]. In addition, Gly and Ser were enriched to provide conformational flexibility, whereas Ala seemed to compromise protein stability. Trp and Arg were found to contribute significantly to binding affinity, but they were detrimental to specificity (particularly Arg). Other amino acids were not effective in replacing Tyr, and were rarely found in the selected libraries. For some antigens, high affinity antibodies could be derived using only Tyr/Ser/Gly diversity, but for others, additional chemical diversity was required for high affinity. Interestingly, while the selection libraries were varied in the CDR-H3 length, the selected binders towards a specific antigen were of a specific length, showing the importance of the backbone in providing a good framework for binding. Thus, not only is an antibody binding site a physically advantageous type of protein-interaction surface, but there

are distinctly better variants of such sites (differing in residue composition and loop length) that have higher capacity than others to solve the interaction problem.

Coiled coils

Alpha helical coiled coils are good models for studying protein-protein interactions due to the simplicity of their rod-like structures, which are formed of two or more helices twisted into a superhelical bundle. From simple considerations of the shared sequence pattern and the commonalities of structure, any coiled-coil peptide is a potential interaction partner for any other, and so these proteins would seem to have high potential to form promiscuous interactions. But several large-scale studies indicate that coiled-coil interactions are not indiscriminate, even when removed from their native biological environment and their structural context in full-length proteins [33–35]. Distinct interactions profiles are observed even for different members of a common protein family. Due to the extended, rod-like structure of coiled coils, modular strategies that involve decorating the lengths of the helices with charged and polar groups can be used to encode positive and negative interaction elements. For example, asparagine residues apposed with hydrophobic residues are highly destabilizing at coiled-coil dimer interfaces, and charge repulsion between adjacent sites across interfaces is unfavorable. Using these principles alone, numerous non-cognate interactions can be prevented by appropriate placement of destabilizing interactions along the helix. Patterning of even relatively short (< 40-residue) sequences can give rise to impressively high degrees of specificity and complex interaction profiles [36**,37*].

One example of specificity in coiled-coil interactions comes from a recent study of laminin heterotrimers. Out of 45 possible heterotrimers composed of 11 different isoforms, only 16 were found to assemble in vitro, demonstrating modest multi-specificity for the component chains. The complexes observed in vitro were consistent with those known to be present and/or functional in the cell [38]. Conversely, SNARE coiled-coil peptides are more promiscuous, as assessed in large-scale yeast two-hybrid tests, but still show preferences for certain types of interactions over others [35]. An interesting example illustrating the concept of “specificity by demand” is the human and viral bZIP coiled coils, which have been studied using peptide arrays [39,40]. Structural studies show that most bZIP complexes have highly similar structures; still, some interact very selectively while others show more promiscuous multi-specificity. Two viral bZIPs (Meq and HBZ) were shown to associate with proteins from ~8 of the 20 human bZIP families. Given that the interaction partner of a bZIP is important for determining its function, the intricate specificity profiles of coiled-coil mediated interactions are probably functionally important both for human and pathogen-host transcriptional control.

Beyond the examples of selective, combinatorial coiled-coil interactions in nature, such associations can also be exploited for protein engineering. A panel of 27 synthetic coiled-coil peptides designed to be heterospecific showed a rich pattern of pair-wise interactions that can potentially be used to wire specific protein interactions into cellular circuits for applications in synthetic biology [37*].

Binding to multiple conformations of structured domains

The examples above show relatively minor changes in structure across distinct complexes, even for proteins that have numerous interaction partners. For other proteins with multiple partners, promiscuity is facilitated by greater degrees of structural plasticity. In this section we discuss conformational changes for structured proteins that have been related to multi-specificity. In some cases, thermodynamic or kinetic insights into the recognition mechanism are available.

Allosteric control

Many proteins exhibit allosteric control, shifting from an “off” to an “on” state. Gao et al. selected for antibodies that bound specifically to only one of two conformations of caspase-1 [41*]. Two selected antibodies bound to their cognate antigens with 20- to 500-fold higher affinity than to the noncognate caspase conformer. Kinetic analysis of the interaction of the two antibodies towards the “on”, “off” and “apo” forms of caspase-1 showed that in this case the specificity of binding was mostly dictated by large differences in k_{on}.

A strategy to specifically employ optimization of k_{on} towards the “off” state of Ras binding was used to enhance binding of Raf towards Ras-GDP. The 30–100-fold enhancement in affinity of the Raf mutants towards the Ras-GDP form made the interaction sufficiently tight to activate the mitogen-activated protein kinase pathway [42*,43]. Structural analysis of these Raf mutants in complex with RasGDP revealed that the loop on Ras termed switch I was found in a conformation similar to that of RasGTP and not RasGDP, however, with an increased mobility relative to that observed in RasGTP.

Internal flexibility of binding sites drives promiscuity

For some proteins, flexibility of the binding site(s) promotes multiple conformations that allow interactions with different partners [44]. Ubiquitin is a highly promiscuous protein, and a structural ensemble of ubiquitin, refined against residual dipolar couplings that captured the solution dynamics up to microseconds, accounted for the complete structural heterogeneity observed in 46 ubiquitin crystal structures, most of which are complexes with other proteins. This indicates that conformational selection, rather than induced-fit, explains the molecular recognition dynamics of this example. With a large part of the solution dynamics being concentrated in one concerted mode, the entropic cost of complex formation is kept low [45**]. A similar conclusion was obtained from computational analysis of germline antibody H3 loops. This work suggested that the H3 sequences are nearly optimal to provide conformational flexibility, while the maturation process fixes specific conformations and provides enhanced specificity in a rigidified scaffold [46]. Another example is the regulatory subunit of protein kinase A (PKA), which can interact with multiple partners at high affinity using a relatively hydrophobic interface. Relying on MD simulations, it was suggested that there is a direct linear relationship between total configurational entropy, stemming from favorable alternative contacts, and the binding energy. Sampling of these different contacts increases the number of bound configurations and thus minimizes the frustration and the entropy change upon binding. The hydrophobic interface of PKA/Ht31pep thus provides a favorable environment for a set of structures, each of which is minimally frustrated and thus energetically stable [47].

Extreme promiscuity and intrinsic disorder

Many proteins have domains or regions that are intrinsically disordered and adopt a defined conformation only upon interaction with a partner. Promiscuous proteins often have such disordered regions, and extreme structural flexibility provides one mechanism for extreme promiscuity. In this section we discuss PDZ domains and p53, two extensively studied examples of promiscuous proteins where one interaction partner is natively disordered. The p53 example illustrates how intrinsic disorder enables a wide range of complex structures, whereas many distinct PDZ complexes show significant structural homology.

PDZ domains

PDZ domains bind to short, intrinsically unstructured regions typically located at the C-termini of proteins, with dissociation constants of ~1–50 μM. Numerous PDZ domain interactions with candidate peptide partners have been experimentally tested using protein

arrays, peptide SPOT arrays and phage display [48–50]. Historically PDZ domains were divided into two classes, based on a preference for hydrophobic residues vs. Ser/Thr at the -2 position of the peptide ligand, and were viewed as relatively non-specific. But larger data sets have revealed many more intricate specificity classes, or even a continuum of specificities [48,49*]. Individual domains nevertheless interact with many partners. Some murine PDZ domains bound over 20 peptides derived from the C-termini of murine proteins on a protein array [48], and when tested against randomized phage libraries, it is common for a single PDZ domain to bind more than 50 peptides with unique sequences [49*]. Detailed studies of the Erbin PDZ domain have demonstrated that this protein can accommodate many mutations at the binding interface and still maintain peptide binding, often with interesting changes in binding specificity [51]. Clearly, there are many solutions to the PDZ-peptide interaction problem.

Available structures support a well-defined and relatively invariant complex for many PDZ domains, but for others, conformational exceptions are observed, and changes are also observed in some examples between apo and bound structures. Computational studies have proposed explanations for these differences [52]. C-terminal peptides often bind in a structurally similar position with respect to the beta strand and helix that define the PDZ binding groove [53*]. But interesting variations in binding geometry have recently been reported for Dvl2 PDZ domain, which adopts different conformations to accommodate internal, as opposed to C-terminal ligands [54]. Structural plasticity of the PDZ domain appears important in this case for promiscuity with respect to peptide ligand type.

The large data sets of PDZ interactions have inspired development of computational approaches for predicting PDZ binding specificity. Two types of approaches show promise. In one, the PDZ-peptide interface is represented by a set of residue-pair interactions. The particular pairs can be chosen based on the structure, or learned in the process of developing the predictor. Weights are assigned to a set of important residue pairs using the experimental data and a machine learning procedure [48,55*,56]. In the other type of approach, structure is considered explicitly, and energy terms are computed from modeled complexes [53*]. Predicting specificity is a very challenging computational problem, and accuracy of all methods is understandably limited. A better understanding of structural plasticity in PDZ binding could guide the further development of both approaches.

The p53 protein

The tumor suppressor p53 is a much-studied protein, famous for having many interaction partners and playing key roles in numerous critical pathways. Additional p53-interacting proteins are being identified all the time [57]. Approximate binding sites on p53 for more than 84 partners have been identified, and numerous high-resolution structures have been solved for complexes involving parts of the protein. Dunker and colleagues point out that the majority of known interaction sites on p53 are located in regions experimentally determined to be disordered in the unbound state, suggesting that extreme structural plasticity plays an important role in the interaction promiscuity of the protein [58*]. Many interactions map to the intrinsically disordered C-terminal domain. For example, 7 residues in this region were shown to adopt four distinct conformations spanning three major secondary structures (helix, extended strand, coil) in four different complexes [58*]. Different residues are important for the different interactions. Additional structures covering this region of p53 have been now been solved and show yet other binding modes [59,60]. Structures illustrating different interaction modes for residues in the 367–391 region of p53 are presented in Fig. 2.

Unstructured regions of proteins are rich in post-translational modifications. Indeed, this is true of the C-terminal domain of p53, and 5 of the 8 structures involve modification of

residues by phosphorylation, methylation or acetylation (see Fig. 2). These chemical modifications alter the conformational and interaction properties of p53. Nussinov and colleagues point out that promiscuous protein-protein interactions that have been attributed to a single protein in the literature are often not associated with a single chemical entity [61]. Rather, distinct interactions can be attributed to differently processed gene products. From a molecular recognition perspective this is entirely sensible, and illustrates a strategy used in nature to diversify gene function. p53 provides a dramatic example of interaction specificity that is tunable by protein modification, and the many available structures are starting to reveal the details of this phenomenon at high resolution.

Design of specificity and multi-specificity

In recent years we have witnessed an impressive increase in the successful use of computational tools for protein design, including the first designs of de-novo protein interfaces [62]. More common than designing new interfaces is re-designing known ones with demonstrated interaction capacity. Computer-based design was used to replace a group of connected amino acids within the interface of TEM1-BLIP (termed a module) with another module selected from an unrelated protein structure. This resulted in an alternative interface, where the mutant proteins bound one another specifically and with high affinity, but binding to the wild-type proteins was severely compromised, resulting in a new interaction with a specificity factor of ~300-fold [63*]. An alternative strategy for introducing specificity is to insert deleterious mutations into an interface, and then to search for second-site repressor mutations that re-establish the binding affinity. Kuhlman and colleagues implemented such a protocol computationally by scanning an interface for knobs and holes or changing hydrophobic residues to polar or vice versa. In 4 out of 8 cases, a specificity factor of 10–50 was achieved [64].

Calmodulin (CaM) can bind and regulate different protein targets mediating processes such as inflammation, metabolism, apoptosis and more. CaM is regulated through the binding of calcium, which induces a structural shift that enables it to bind specific partners. Shifman and coworkers aimed to design a CaM variant with greater binding specificity. This was achieved solely through computational optimization of interaction with a specific peptide, without using negative design. Experimental measurements showed a small increase in affinity for the specific peptide but a large drop in binding affinity to other peptides [65*]. This is an example where the targets are apparently sufficiently different such that optimization towards one is not coupled to increasing the affinity towards another. Further computational analysis has suggested that the CaM interface can be divided into positions that provide affinity while others provide specificity of binding [66].

In contrast to CaM, negative design appears to be critical for engineering specific coiled-coil complexes. Our current understanding of positive and negative elements in coiled-coil recognition is sufficient to design molecules with desired interaction properties, and engineering of interaction specificity has been demonstrated for both simple homo vs. heterodimer examples, as well as for complex multi-state problems with many undesired states [36**,67,68]. Negative design was strongly predicted to be necessary in many of these examples to disfavor competing complexes, especially undesired homodimerization of the designed protein. This is consistent with selection experiments that found strong off-target interactions when there was no counter selection against these [69].

Most prokaryotes harbor dozens of histidine kinases that interact with their respective response regulators to form two-component signaling pathways. While the protein pairs are similar, there is no promiscuous crosstalk between them [70]. Laub and colleagues have used computational co-variation analysis to localize the interaction specificity to a point at

the tip of the kinase 4-helix bundle DHp domain, and demonstrated that modest changes in sequence in this local region can re-wire the interaction specificity [71*]. The fact that changes of just a few amino acids are sufficient to completely switch specificity is another example, similar to colicin-immunity protein complexes discussed above, of how multiple solutions to the interaction problem can be close in sequence space, facilitating the divergence of duplicated genes for distinct functions.

Summary

In this review, through the many examples given, we highlight how protein complexes can accommodate extensive variation in sequence, even at structurally similar interfaces. Any protein can interact with many proteins, however at different affinities. Differences in affinity determine specificity, but in a biological context the energy gap that is relevant depends on functional requirements as well as on local protein concentrations. The lack of absolute specificity is a result of the many solutions that exist to obtain binding, which can be achieved at a single interface or at different interfaces. Therefore, the need for negative design, whether evolutionary or man-made, has to be considered with respect to *in vivo* conditions that are difficult to define. Despite these challenges, advances in systems biology, structural methods, combinatorial high-throughput approaches and computer modeling and design have improved our understanding of both promiscuous and specific interactions. We have also seen the first steps towards successful prediction and rational design of interaction specificity.

Acknowledgments

We thank members of the Schreiber and Keating laboratories for their thoughtful contributions to our studies of specificity. G.S acknowledges support from the MINERVA foundation (grant 5670). A.E.K. acknowledges support of work on interaction specificity from NIH awards GM084181 and GM067681.

References and recommended reading

1. Elcock AH. Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. Curr Opin Struct Biol 2010;20:196–206. [PubMed: 20167475]

2*. Wang Y, Li C, Pielak GJ. Effects of proteins on protein diffusion. J Am Chem Soc 2010;132:9392–9397. In this NMR study it is shown that weak interactions at high concentrations in an E.coli cell lysate significantly affect the rotational diffusion of a protein, an effect that is absent when using synthetic polymers, which apparently are inert to the probed protein. [PubMed: 20560582]

3. Greenspan NS. Cohen's Conjecture, Howard's Hypothesis, and Ptashne's Ptruth: an exploration of the relationship between affinity and specificity. Trends Immunol 2010;31:138–143. [PubMed: 20149744]

4. Neuvirth H, Raz R, Schreiber G. ProMate: A Structure Based Prediction Program to Identify the Location of Protein-Protein Binding Sites. J Mol Biol 2004;338:181–199. [PubMed: 15050833]

5. Nooren IM, Thornton JM. Diversity of protein-protein interactions. EMBO J 2003;22:3486–3492. [PubMed: 12853464]

6. Khersonsky O, Tawfik DS. Enzyme promiscuity: a mechanistic and evolutionary perspective. Annu Rev Biochem 2010;79:471–505. [PubMed: 20235827]

7*. Martin J. Beauty is in the eye of the beholder: proteins can recognize binding sites of homologous proteins in more than one way. PLoS Comput Biol 2010;6:e1000821. In this bioinformatic study of protein-protein interfaces it is suggested that different proteins interact with similar partners using alternate strategies. In other words, there are many solutions to bind the same protein. The different solutions may use the same surfaces and amino acids on the shared binding sites, or alternative ones. [PubMed: 20585553]

8. Strynadka NC, Jensen SE, Alzari PM, James MN. A potent new mode of beta-lactamase inhibition revealed by the 1.7 A X-ray crystallographic structure of the TEM-1-BLIP complex. Nat Struct Biol 1996;3:290–297. [PubMed: 8605632]

9. Lim D, Park HU, De Castro L, Kang SG, Lee HS, Jensen S, Lee KJ, Strynadka NC. Crystal structure and kinetic analysis of beta-lactamase inhibitor protein-II in complex with TEM-1 beta-lactamase. Nat Struct Biol 2001;8:848–852. [PubMed: 11573088]

10. Reynolds KA, Thomson JM, Corbett KD, Bethel CR, Berger JM, Kirsch JF, Bonomo RA, Handel TM. Structural and computational characterization of the SHV-1 beta -lactamase/beta -lactamase inhibitor protein(BLIP) interface. J Biol Chem 2006;281:26745–26753. [PubMed: 16809340]

11*. Gretes M, Lim DC, de Castro L, Jensen SE, Kang SG, Lee KJ, Strynadka NC. Insights into positive and negative requirements for protein-protein interactions by crystallographic analysis of the beta-lactamase inhibitory proteins BLIP, BLIP-I, and BLP. J Mol Biol 2009;389:289–305. This study provides insights into positive and negative requirements for protein-protein interactions by crystallographic analysis of the beta-lactamase inhibitory proteins BLIP, BLIP-I, and BLP, showing how a similar set of amino acids can be used differently for binding, and how relatively few changes can hinder binding. [PubMed: 19332077]

12. Hanes MS, Jude KM, Berger JM, Bonomo RA, Handel TM. Structural and biochemical characterization of the interaction between KPC-2 beta-lactamase and beta-lactamase inhibitor protein. Biochemistry 2009;48:9185–9193. [PubMed: 19731932]

13. Zhang Z, Palzkill T. Dissecting the protein-protein interface between beta-lactamase inhibitory protein and class A beta-lactamases. J Biol Chem 2004;279:42860–42866. [PubMed: 15284234]

14*. Meenan NA, Sharma A, Fleishman SJ, Macdonald CJ, Morel B, Boetzel R, Moore GR, Baker D, Kleanthous C. The structural and energetic basis for high selectivity in a high-affinity protein-protein interaction. Proc Natl Acad Sci U S A 2010;107:10080–10085. Colicins bind only their specific immunity proteins with very high affinity. This manuscript describes how a low affinity non-cognate complex was trapped for crystallographic study and illuminates the structural basis for weak binding, showing that frustration of a few bonds is sufficient to destroy good binding. [PubMed: 20479265]

15*. Levin KB, Dym O, Albeck S, Magdassi S, Keeble AH, Kleanthous C, Tawfik DS. Following evolutionary paths to protein-protein interactions with high affinity and selectivity. Nat Struct Mol Biol 2009;16:1049–1055. Colicins bind only their specific immunity proteins at a very high affinity. Here, selection methods were used to change the binding selectivity of Im9 from ColE9 to ColE7. Mutations outside the binding site altered the binding orientation of the two proteins such that “latent” interactions were introduced between native Im9 residues and ColE7; these are interactions that resemble those of corresponding Im7 residues with ColE7. It was shown that gaining binding to a new partner did not diminish binding to the original partner without the use of negative selection as well. [PubMed: 19749752]

16. Certo M, Del Gaizo Moore V, Nishino M, Wei G, Korsmeyer S, Armstrong SA, Letai A. Mitochondria primed by death signals determine cellular addiction to antiapoptotic BCL-2 family members. Cancer Cell 2006;9:351–365. [PubMed: 16697956]

17. Boersma MD, Sadowsky JD, Tomita YA, Gellman SH. Hydrophile scanning as a complement to alanine scanning for exploring and manipulating protein-protein recognition: application to the Bim BH3 domain. Protein Sci 2008;17:1232–1240. [PubMed: 18467496]

18*. Dutta S, Gulla S, Chen TS, Fire E, Grant RA, Keating AE. Determinants of BH3 binding specificity for Mcl-1 versus Bcl-xL. J Mol Biol 2010;398:747–762. Yeast surface display screening that involved both positive and negative selection was used to identify high affinity peptides that bound either Mcl-1 or Bcl-xL selectively. Peptide SPOT arrays were used to classify the effects of many point substitutions on binding specificity, elucidating both positive and negative design elements. Within the sampled sequence space, a position weight matrix derived from the SPOT data could largely explain the behavior of 90 unique peptides classified as binding Mcl-1 and/or Bcl-xL in the yeast screening. [PubMed: 20363230]

19. Lee EF, Czabotar PE, Smith BJ, Deshayes K, Zobel K, Colman PM, Fairlie WD. Crystal structure of ABT-737 complexed with Bcl-xL: implications for selectivity of antagonists of the Bcl-2 family. Cell Death Differ 2007;14(9):1711–1713. [PubMed: 17572662]

20. Stewart ML, Fire E, Keating AE, Walensky LD. The MCL-1 BH3 helix is an exclusive MCL-1 inhibitor and apoptosis sensitizer. Nat Chem Biol 2010;6:595–601. [PubMed: 20562877]
21. Rautureau GJ, Day CL, Hinds MG. Intrinsically disordered proteins in bcl-2 regulated apoptosis. Int J Mol Sci 2010;11:1808–1824. [PubMed: 20480043]
22. Fire E, Gulla SV, Grant RA, Keating AE. Mcl-1-Bim complexes accommodate surprising point mutations via minor structural changes. Protein Sci 2010;19:507–519. [PubMed: 20066663]
23. Lee EF, Czabotar PE, van Delft MF, Michalak EM, Boyle MJ, Willis SN, Puthalakath H, Bouillet P, Colman PM, Huang DC, Fairlie WD. A novel BH3 ligand that selectively targets Mcl-1 reveals that apoptosis can proceed without Mcl-1 degradation. J Cell Biol 2008;180:341–355. [PubMed: 18209102]
24. Day CL, Smits C, Fan FC, Lee EF, Fairlie WD, Hinds MG. Structure of the BH3 domains from the p53-inducible BH3-only proteins Noxa and Puma in complex with Mcl-1. J Mol Biol 2008;380:958–971. [PubMed: 18589438]
25. Lee EF, Czabotar PE, Yang H, Sleebs BE, Lessene G, Colman PM, Smith BJ, Fairlie WD. Conformational changes in Bcl-2 pro-survival proteins determine their capacity to bind ligands. J Biol Chem 2009;284:30508–30517. [PubMed: 19726685]
26. Smits C, Czabotar PE, Hinds MG, Day CL. Structural plasticity underpins promiscuous binding of the prosurvival protein A1. Structure 2008;16:818–829. [PubMed: 18462686]
27. Lee EF, Fedorova A, Zobel K, Boyle MJ, Yang H, Perugini MA, Colman PM, Huang DC, Deshayes K, Fairlie WD. Novel Bcl-2 homology-3 domain-like sequences identified from screening randomized peptide libraries for inhibitors of the pro-survival Bcl-2 proteins. J Biol Chem 2009;284:31315–31326. [PubMed: 19748896]
28. Hershberg, U.; Solomon, S.; Cohen, IR. What is the basis of the immune system's specificity. In: Capasso, V., editor. Mathematical Modelling and Cumputing in Biology and Medicine. 2003. p. 377-384.The MIRIAM Project Series, Progetto Leonardo, ESCULAPIO Pub.Co., Bologna, Italy.
29. Michaud GA, Salcius M, Zhou F, Bangham R, Bonin J, Guo H, Snyder M, Predki PF, Schweitzer BI. Analyzing antibody specificity with whole proteome microarrays. Nat Biotechnol 2003;21:1509–1512. [PubMed: 14608365]
30. Kijanka G, Ipcho S, Baars S, Chen H, Hadley K, Beveridge A, Gould E, Murphy D. Rapid characterization of binding specificity and cross-reactivity of antibodies using recombinant human protein arrays. J Immunol Methods 2009;340:132–137. [PubMed: 18996391]
31**. Birtalan S, Fisher RD, Sidhu SS. The functional capacity of the natural amino acids for molecular recognition. Mol Biosyst 2010;6:1186–1194. This manuscript provides a comprehensive experimental analysis of the relation between specificity and sequence by testing the functional capacity of a single fragmented antigen binding region of an Ab (Fab) to bind vascular endothelial growth factor (VEGF), human epidermal growth factor 2 (HER2), insulin-like growth factor 1 (IGF-1) and protein A. For all cases, Tyr was the preferred amino acid contributing to both affinity and specificity. In addition, Gly and Ser were found to be preferable to provide conformational flexibility, while Ala seems to compromise protein stability. Trp and Arg were found to contribute significantly to binding affinity; they were detrimental to specificity (particularly Arg). For some antigens, high affinity Abs could be derived using only Tyr/Ser/Gly diversity, but for others, additional chemical diversity was required to achieve high affinity. [PubMed: 20383388]
32. Gilbreth RN, Esaki K, Koide A, Sidhu SS, Koide S. A dominant conformational role for amino acid diversity in minimalist protein-protein interfaces. J Mol Biol 2008;381:407–418. [PubMed: 18602117]
33. Newman JR, Wolf E, Kim PS. A computationally directed screen identifying interacting coiled coils from Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 2000;97:13203–13208. [PubMed: 11087867]
34. Zizlsperger N, Malashkevich VN, Pillay S, Keating AE. Analysis of coiled-coil interactions between core proteins of the spindle pole body. Biochemistry 2008;47:11858–11868. [PubMed: 18850724]
35. Zhang H, Chen J, Wang Y, Peng L, Dong X, Lu Y, Keating AE, Jiang T. A computationally guided protein-interaction screen uncovers coiled-coil interactions involved in vesicular trafficking. J Mol Biol 2009;392:228–241. [PubMed: 19591838]

36**. Grigoryan G, Reinke AW, Keating AE. Design of protein-interaction specificity gives selective bZIP-binding peptides. Nature 2009;458:859–864. A general formalism for the computational design of protein interaction specificity is presented and applied to the human bZIP transcription factors. The human bZIPs dimerize via a coiled coil, and this work reports successful design of synthetic coiled-coil interaction partners for 19 out of 20 bZIP families. For 8 of the 20 families, coiled-coil microarray experiments used to profile binding specificity indicated that interaction with the intended target was greater than with any other bZIP on the array. Explicit negative design was predicted to be very important in achieving specific binders. The designed molecules showed a wide range of interaction profiles distinct from native bZIPs, illustrating that nature has sampled only a small portion of the bZIP solution space in the human proteome. [PubMed: 19370028]

37*. Reinke AW, Grant RA, Keating AE. A synthetic coiled-coil interactome provides heterospecific modules for molecular engineering. J Am Chem Soc 2010;132:6025–6031. The 48 synthetic proteins designed in ref. 36 were tested for pair-wise interactions with one another using coiled-coil microarrays. These molecules were designed not to homo-associate and they are largely heterospecific. Putative negative design elements important for the specificity were identified. The pair-wise interaction network made up of the 48 “SYNZIPs” is complex and contains numerous motifs that may prove useful for molecular engineering. The synthetic pairs illustrate yet more solutions to the coiled-coil interface problem that are not found in nature. [PubMed: 20387835]

38. Macdonald PR, Lustig A, Steinmetz MO, Kammerer RA. Laminin chain assembly is regulated by specific coiled-coil interactions. J Struct Biol 2010;170:398–405. [PubMed: 20156561]

39. Newman JR, Keating AE. Comprehensive identification of human bZIP interactions with coiled-coil arrays. Science 2003;300:2097–2101. [PubMed: 12805554]

40. Reinke AW, Grigoryan G, Keating AE. Identification of bZIP interaction partners of viral proteins HBZ, MEQ, BZLF1, and K-bZIP using coiled-coil arrays. Biochemistry 2010;49:1985–1997. [PubMed: 20102225]

41*. Gao J, Sidhu SS, Wells JA. Two-state selection of conformation-specific antibodies. Proc Natl Acad Sci U S A 2009;106:3071–3076. A general strategy for identification of conformation-specific antibodies using phage display is presented. It was applied to the selection of two Fabs, each of which bound to its cognate (active versus non-active) caspase-1 conformer 20- to 500-fold more tightly than its noncognate conformer. The Fabs were used to localize the different states in cells and revealed the activated caspase-1 is concentrated in a central structure in the cytosol, similar to what has been described as the pyroptosome. [PubMed: 19208804]

42*. Kiel C, Filchtinski D, Spoerner M, Schreiber G, Kalbitzer HR, Herrmann C. Improved binding of raf to Ras.GDP is correlated with biological activity. J Biol Chem 2009;284:31893–31902. Only Ras bound to GTP is able to interact strongly with its effector Raf, whereas in the GDP-bound state, the stability of the complex is strongly decreased. In this manuscript, computer-aided protein design was used to improve the interaction between RasGDP and Raf. This improved binding resulted in the loss in specificity for RasGTP versus RasGDP, making RasGDP biologically active in signaling. The study suggests that binding affinity is the main factor dictating specificity of signaling in this system. [PubMed: 19776012]

43. Filchtinski D, Sharabi O, Ruppel A, Vetter IR, Herrmann C, Shifman JM. What Makes Ras an Efficient Molecular Switch: A Computational, Biophysical, and Structural Study of Ras-GDP Interactions with Mutants of Raf. J Mol Biol 2010;399:422–435. [PubMed: 20361980]

44. Matthews JM, Bhati M, Lehtomaki E, Mansfield RE, Cubeddu L, Mackay JP. It takes two to tango: the structure and function of LIM, RING, PHD and MYND domains. Curr Pharm Des 2009;15:3681–3696. [PubMed: 19925420]

45**. Lange OF, Lakomek NA, Fares C, Schroder GF, Walter KF, Becker S, Meiler J, Grubmuller H, Griesinger C, de Groot BL. Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. Science 2008;320:1471–1475. In this manuscript the solution dynamics of a structural ensemble of ubiquitin was obtained from residual dipolar coupling data. The ensemble covers the complete structural heterogeneity observed in 46 ubiquitin crystal structures, most of which are complexes with other proteins. This shows that conformational selection, rather than induced-fit motion, explains the molecular recognition

dynamics of ubiquitin. A large part of the solution dynamics relates to one concerted mode, ensuring a low entropic cost for complex formation. [PubMed: 18556554]

46. Babor M, Kortemme T. Multi-constraint computational design suggests that native sequences of germline antibody H3 loops are nearly optimal for conformational flexibility. Proteins 2009;75:846–858. [PubMed: 19194863]
47. Chang CE, McLaughlin WA, Baron R, Wang W, McCammon JA. Entropic contributions and the influence of the hydrophobic environment in promiscuous protein-protein association. Proc Natl Acad Sci U S A 2008;105:7456–7461. [PubMed: 18495919]
48. Stiffler MA, Chen JR, Grantcharova VP, Lei Y, Fuchs D, Allen JE, Zaslavskaia LA, MacBeath G. PDZ domain binding selectivity is optimized across the mouse proteome. Science 2007;317:364–369. [PubMed: 17641200]
49*. Tonikian R, Zhang Y, Sazinsky SL, Currell B, Yeh JH, Reva B, Held HA, Appleton BA, Evangelista M, Wu Y, et al. A specificity map for the PDZ domain family. PLoS Biol 2008;6:e239. Phage display was used to derive peptide binding profiles for 330 human and worm PDZ domains and 91 Erbin PDZ domain point mutants. The results support many specificity classes with relatively subtle differences between them. Comparison of several human and worm orthologs suggests that binding specificity is conserved. Binding site residue identity is a good predictor of binding profile similarity. [PubMed: 18828675]
50. Wiedemann U, Boisguerin P, Leben R, Leitner D, Krause G, Moelling K, Volkmer-Engert R, Oschkinat H. Quantification of PDZ domain specificity, prediction of ligand affinity and rational design of super-binding peptides. J Mol Biol 2004;343:703–718. [PubMed: 15465056]
51. Ernst A, Sazinsky SL, Hui S, Currell B, Dharsee M, Seshagiri S, Bader GD, Sidhu SS. Rapid evolution of functional complexity in a domain family. Sci Signal 2009;2:ra50. [PubMed: 19738200]
52. Ho BK, Agard DA. Conserved tertiary couplings stabilize elements in the PDZ fold, leading to characteristic patterns of domain conformational flexibility. Protein Sci 2010;19:398–411. [PubMed: 20052683]
53*. Smith CA, Kortemme T. Structure-Based Prediction of the Peptide Sequence Space Recognized by Natural and Synthetic PDZ Domains. J Mol Biol. 2010 Structure based methods incorporating modest backbone flexibility were used to predict binding profiles for 17 PDZ domains of known structure. The results were compared to profiles generated by phage display. In some cases, the modeling was able to recognize that a residue other than the native residue in the PDB structure was stabilizing. Significant changes in peptide binding specificity upon PDZ mutation were also recapitulated for some variants of the Erbin PDZ domain.
54. Zhang Y, Appleton BA, Wiesmann C, Lau T, Costa M, Hannoush RN, Sidhu SS. Inhibition of Wnt signaling by Dishevelled PDZ peptides. Nat Chem Biol 2009;5:217–219. [PubMed: 19252499]
55*. Chen JR, Chang BH, Allen JE, Stiffler MA, MacBeath G. Predicting PDZ domain-peptide interactions from primary sequences. Nat Biotechnol 2008;26:1041–1045. A statistical learning approach that addressed the problem of sparse training data was used to derive a PDZ binding model from protein array data (ref. 48). The model assigns amino-acid-dependent weights to a set of 38 domain-peptide residue pairs and it significantly outperformed a method for predicting interaction specificity that used general residue pair potentials. It also showed good performance on fly and worm interactions, after training on murine PDZ domain data. [PubMed: 18711339]
56. Thomas J, Ramakrishnan N, Bailey-Kellogg C. Graphical models of protein-protein interaction specificity from correlated mutations and interaction data. Proteins 2009;76:911–929. [PubMed: 19306342]
57. Lunardi A, Di Minin G, Provero P, Dal Ferro M, Carotti M, Del Sal G, Collavin L. A genome-scale protein interaction profile of Drosophila p53 uncovers additional nodes of the human p53 network. Proc Natl Acad Sci U S A 2010;107:6322–6327. [PubMed: 20308539]
58*. Oldfield CJ, Meng J, Yang JY, Yang MQ, Uversky VN, Dunker AK. Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. BMC Genomics 2008;9(Suppl 1):S1. Using p53 and 14-3-3 as case studies, the authors illustrate ways in which extreme structural plasticity can be associated with promiscuity. The two examples are complementary because p53 participates in many interactions via its intrinsically disordered regions, whereas14-3-3 is a structured protein that binds disordered partners into a structurally

conserved groove. Interestingly, as for TEM-BLIP, 14-3-3 can bind different partners at a common site, but using different amino-acid interactions. The large amount of structural data available for p53 and, to a lesser extent, 14-3-3, makes these examples compelling models for studying promiscuity.

59. Roy S, Musselman CA, Kachirskaia I, Hayashi R, Glass KC, Nix JC, Gozani O, Appella E, Kutateladze TG. Structural insight into p53 recognition by the 53BP1 tandem Tudor domain. J Mol Biol 2010;398:489–496. [PubMed: 20307547]
60. Schumacher B, Mondry J, Thiel P, Weyand M, Ottmann C. Structure of the p53 C-terminus bound to 14-3-3: implications for stabilization of the p53 tetramer. FEBS Lett 2010;584:1443–1448. [PubMed: 20206173]
61. Tsai CJ, Ma B, Nussinov R. Protein-protein interaction networks: how can a hub protein bind so many different partners? Trends Biochem Sci 2009;34:594–600. [PubMed: 19837592]
62. Mandell DJ, Kortemme T. Computer-aided design of functional protein interactions. Nat Chem Biol 2009;5:797–807. [PubMed: 19841629]
63*. Potapov V, Reichmann D, Abramovich R, Filchtinski D, Zohar N, Ben Halevy D, Edelman M, Sobolev V, Schreiber G. Computational redesign of a protein-protein interface for high affinity and binding specificity using modular architecture and naturally occurring template fragments. J Mol Biol 2008;384:109–119. Protein-protein interfaces are built from distinct modules. In this manuscript, one such module was replaced by another taken from a non-related structure, resulting in highly specific binding of the mutant proteins to one another, but not to the wild-type partner. Thus, specificity was achieved without the explicit use of negative design. The success of the design was attributed to utilizing the modular construction of the interface, capitalizing on native rather than artificial templates, and ranking with an accurate atom-atom contact surface scoring function. [PubMed: 18804117]
64. Sammond DW, Eletr ZM, Purbeck C, Kuhlman B. Computational design of second-site suppressor mutations at protein-protein interfaces. Proteins 2010;78:1055–1065. [PubMed: 19899154]
65*. Yosef E, Politi R, Choi MH, Shifman JM. Computational design of calmodulin mutants with up to 900-fold increase in binding specificity. J Mol Biol 2009;385:1470–1480. The authors attempted redesign of calmodulin to improve both affinity and specificity of binding to two helical peptide targets. Affinity was not much improved in the designs, but significant binding specificity gaps could be introduced using only positive design for the desired partner. The authors suggest that calmodulin interactions are already largely optimized, so changes in sequence will tend to be destabilizing in the absence of positive design. This is an example of a case where specificity design does not require explicit negative design (contrast to ref. 36). [PubMed: 18845160]
66. Fromer M, Shifman JM. Tradeoff between stability and multispecificity in the design of promiscuous proteins. PLoS Comput Biol 2009;5:e1000627. [PubMed: 20041208]
67. Havranek JJ, Harbury PB. Automated design of specificity in molecular recognition. Nat Struct Biol 2003;10:45–52. [PubMed: 12459719]
68. Barth P, Schoeffler A, Alber T. Targeting metastable coiled-coil domains by computational design. J Am Chem Soc 2008;130:12038–12044. [PubMed: 18698842]
69. Mason JM, Muller KM, Arndt KM. Positive aspects of negative design: simultaneous selection of specificity and interaction stability. Biochemistry 2007;46:4804–4814. [PubMed: 17402748]
70. Skerker JM, Prasol MS, Perchuk BS, Biondi EG, Laub MT. Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. PLoS Biol 2005;3:e334. [PubMed: 16176121]
71*. Skerker JM, Perchuk BS, Siryaporn A, Lubin EA, Ashenberg O, Goulian M, Laub MT. Rewiring the specificity of two-component signal transduction systems. Cell 2008;133:1043–1054. Prior work established that two-component histidine kinases show a kinetic preference in vitro for phosphorylating their functionally cognate response regulator over the many alternatives in E. coli (ref. 70). In this work, computational co-variation analysis of large multiple-sequence alignments was used to identify residues responsible for encoding phosphotransfer specificity. Out of 12 residues with high inter-protein mutual information scores, three were identified that were sufficient to switch the response-regulator specificity of EnvZ to that of RstB, another E. coli histidine kinase. Other successful specificity swaps were engineering as well, promoting co-

variation analysis as an important tool for deciphering and re-wiring specificity. [PubMed: 18555780]

72. Rustandi RR, Baldisseri DM, Weber DJ. Structure of the negative regulatory domain of p53 bound to S100B(betabeta). Nat Struct Biol 2000;7:570–574. [PubMed: 10876243]
73. Lowe ED, Tews I, Cheng KY, Brown NR, Gul S, Noble ME, Gamblin SJ, Johnson LN. Specificity determinants of recruitment peptides bound to phospho-CDK2/cyclin A. Biochemistry 2002;41:15625–15634. [PubMed: 12501191]
74. Cosgrove MS, Bever K, Avalos JL, Muhammad S, Zhang X, Wolberger C. The structural basis of sirtuin substrate affinity. Biochemistry 2006;45:7511–7521. [PubMed: 16768447]
75. Mujtaba S, He Y, Zeng L, Yan S, Plotnikova O, Sachchidanand, Sanchez R, Zeleznik-Le NJ, Ronai Z, Zhou MM. Structural mechanism of the bromodomain of the coactivator CBP in p53 transcriptional activation. Mol Cell 2004;13:251–263. [PubMed: 14759370]
76. Chuikov S, Kurash JK, Wilson JR, Xiao B, Justin N, Ivanov GS, McKinney K, Tempst P, Prives C, Gamblin SJ, et al. Regulation of p53 activity through lysine methylation. Nature 2004;432:353–360. [PubMed: 15525938]

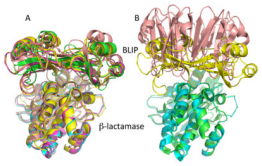

Figure 1.
Overlay of β-lactamase – BLIP structures. A. cyan – TEM1-BLIP (1JTG), magenta – SHV-BLIP (2G2U), yellow – KPC2-BLIP (3E2L), pink – TEM1-BLIP1 (3GMW) and green – is BLP (3GMX) overlaid on TEM1-BLIP. B. TEM1-BLIP2 (1JTD) structure (pink over cyan) overlaid on TEM1-BLIP.

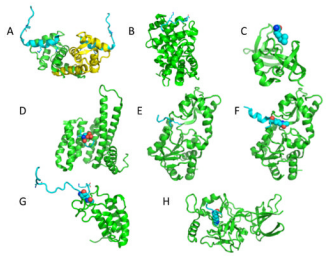

Figure 2.

Interactions formed by p53 residues 367–391. In each panel, the p53 fragment is in cyan. The p53 sequence is 367-SHLKS**K**KGQSTSRHK**K**LMF**K**TEGPD-391, and in each structure the bold lysine residues are shown in stick representation, if unmodified. Post-translationally modified residues are shown using spheres. (A) S100 calcium-binding protein with p53 residues 367–388 (1DT7) [72], (B) Cyclin A2 with p53 residues 378–386 (1H26) [73], (C) Tumor suppressor p53-binding protein 1 with p53 dimethylated lysine residue 382 (3LGL); residues 377–381 and 383–387 did not show clear density [59], (D) 14-3-3 protein sigma with p53 residues 385–391, phosphothreonine 387 (3LW1) [60], (E) NAD-dependent deacetylase Sir2 with p53 residues 378–384 (2H2F) [74], (F) NAD-dependent deacetylase Sir2 with p53 residues 373–385; acetyllysine 382 (2H2D) [74], (G) CREB-binding protein with p53 residues 367–386; acetyllysine 382 (1JSP) [75], (H) Histone-lysine N-methyltransferase with p53 residues 369–374; N-methyl-lysine 372 (1XQH) [76].

Table 1

Amino-acid identity between pairs of BLIP and β–lactamase proteins

protein 1	aa	protein 2	aa	% identity
BLP	154	BLIP	165	32
BLP	154	BLIP1	156	42
BLP	154	BLIPII	273	3
BLIP	165	BLIP1	156	39
BLIP	165	BLIPII	273	12
BLIP1	156	BLIPII	273	3
TEM1	262	KPC	259	40
TEM1	262	SHV1	265	67
KPC	259	SHV1	265	41

Table 2

Amino-acid identity at conserved locations in the BLIP- β–lactamase interface

aa number β-lactamase	Residue identity	aa number BLIP	Residue identity
99	N/K	31	E/D (S)
104	E/P/D	35	S/V (-)
105	Y/W	41	H/L (L)
106	S	48	G (D)
107	P	49	D (H)
108	E/-	50	Y (-)
110	E	53	Y (Q)
111	K	71	S/R (R)
112	H/Y/-	73	E (M)
129	M/Y	74	K/Y (E)
130	S/-	142	F/L (R)
167	P/L/T/-	143	Y/F (F)
168	E/-	162	W/R (S)
215	K/T/R		
216	V/T		
236	G/-		
237	A/T		
238	G/C/-		

-is designated for a location not conserved in TEM1 in complex with BLIP2 (pdb 1JTD). In brackets are the residues on BLP at the designated location, with – indicating no residue.