Published in final edited form as:

Methods Enzymol. 2013 ; 523: 285–302. doi:10.1016/B978-0-12-394292-0.00013-8.

Generation of high-performance binding proteins for peptide motifs by affinity clamping

Shohei Koide^{1,*} and Jin Huang²

¹Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, IL 60637, USA

²Beijing Prosperous Biopharm Co., Ltd., Beijing, 100085, China

Abstract

We describe concepts and methodologies for generating “Affinity Clamps”, a new class of recombinant binding proteins that achieve high affinity and high specificity toward short peptide motifs of biological importance, which is a major challenge in protein engineering. The Affinity Clamping concept exploits the potential of nonhomologous recombination of protein domains in generating large changes in protein function and the inherent binding affinity and specificity of the so-called modular interaction domains toward short peptide motifs. Affinity Clamping creates a clamshell architecture that clamps onto a target peptide. The design processes involve (i) choosing a starting modular interaction domain appropriate for the target and applying structure-guided modifications, (ii) attaching a second domain, termed “enhancer domain” and (iii) optimizing the peptide-binding site located between the domains by directed evolution. The two connected domains work synergistically to achieve high levels of affinity and specificity that are unattainable with either domain alone. Because of the simple and modular architecture, affinity clamps are particularly well suited as building blocks for designing more complex functionalities. Affinity Clamping represents a major advance in protein design that is broadly applicable to the recognition of peptide motifs.

Keywords

molecular recognition; nonhomologous recombination; directed evolution; domain interface; affinity reagents; modular domain; molecular scaffold

1. Introduction

Generating proteins that perform novel molecular-recognition functions has been a major goal of protein design. The central approach in the field has been to choose a starting “molecular scaffold” and to introduce modifications that create surfaces suitable for the new function (Koide, 2010). This concept is most evidently illustrated in the mechanism by which natural and synthetic antibodies are generated, where the amino acid sequences of a set of surface loops, or the complementarity determining regions (CDRs), are diversified to generate a large ensemble from which functional molecules are selected (Sidhu and Fellouse, 2006; Winter et al., 1994). Most structure-guided and computational design of binding sites also follows this concept (Kortemme and Baker, 2004).

Although the current protein design strategies are reasonably successful in generating binding interfaces that recognize highly structured protein molecules, or the “domainome”

*To whom correspondence should be addressed. skoide@uchicago.edu.

portion of the proteome (Colwill and Graslund, 2011), these strategies have not been particularly effective in generating binding interfaces for the “peptidome” portion, comprised of short, unstructured peptide segments (Cobaugh et al., 2008). Such short peptide motifs that are present in the context of larger proteins, in particular those harboring a post-translational modification, are important “hubs” in signal transduction and epigenetics. However, most affinity reagents currently in use for this class of targets are polyclonal antibodies that are inherently undefined, under-characterized and not renewable. The difficulty of designing binding interfaces for short peptides lies in the fundamental thermodynamics where a large entropic cost of immobilizing a flexible peptide must be offset by forming a small number of interactions.

Another challenge in protein design is to generate binding interfaces that recognize a predefined epitope within a targeted molecule. For example, when one generates an antibody to a protein target, the epitope of the antibody is typically poorly defined, and its precise definition requires laborious methods (Parmley and Smith, 1989; Rockberg et al., 2008). This situation stands in a stark contrast to the recognition of short nucleic acid motifs, where one can readily design an oligonucleotide that recognizes a predefined “epitope” within a much larger genome simply using sequence information. A major challenge in protein design has been to develop a platform technology that rapidly creates binding interfaces for a predefined peptide motif. Developing such a technology would enable systematic generation of affinity reagents on a proteomic scale with epitopes known a priori. Clearly, such a system would have a strong impact on biomedical research, biotechnology and medicine.

2. The Affinity Clamping concept

We have developed a new protein-engineering concept, termed “Affinity Clamping”, to systematically generate capture reagents directed to predefined peptide motifs (Huang et al., 2008; Koide, 2009). This concept was inspired by a series of observations regarding structure-function relationships and the evolution of proteins. In the evolution of natural proteins, nonhomologous recombination of domains is considered the main driving force for a large leap in function (Bashton and Chothia, 2007). Comparative structural analyses suggest that more specific and complex protein functions are achieved by two or multi-domain proteins with the active site frequently occurring at the interface between domains (Rossmann et al., 1974). Many proteins that act on peptides, e.g. enzymes, fall in this category. There exist numerous modular interaction domains that bind to particular classes of peptide motifs, but their binding sites are small and shallow grooves located on the protein surface, which explains the low levels of affinity and specificity of the interactions between modular interaction domains and their respective target peptide motifs (Pawson and Nash, 2003).

The observations discussed above suggest that one could design a high-performance binding protein for peptide motifs by first connecting two nonhomologous domains and then optimizing the surfaces at the newly formed interface between them. Specifically, in affinity clamping, a natural peptide-binding domain (referred to as a “primary domain”) is combined with another unrelated domain (an “enhancer domain”), followed by directed evolution of the surfaces of the “enhancer domain” that are expected to contact the peptide bound to the primary domain (Figure 1) (Huang et al., 2008). The enhancer domain is essentially an antibody mimic, in that its role is to recognize the target peptide presented on the surface of the primary domain. The resulting binding proteins, termed “Affinity Clamps”, with their clamshell architecture would clamp on the target peptides.

We have successfully applied Affinity Clamping to a PDZ domain as the primary domain and dramatically enhanced its target recognition function (Huang et al., 2008; Huang et al., 2009a). The binding affinity and specificity of the resulting Affinity Clamps were respectively >500-fold (to single nanomolar K_d values) and >2,000-fold higher than those of the primary domain alone. The PDZ clamps were produced at high levels in *E. coli* and were highly stable. They outperformed commercial monoclonal antibodies to the same target in Western blotting. These results validated the Affinity Clamping strategy in producing high-performance binding proteins.

Affinity Clamping addresses two complementary challenges in generating a high-affinity binding site for a short, flexible peptide. First, the affinity clamp architecture creates peptide-interacting surfaces that are much larger than those of the primary domain alone, thereby enabling higher affinity and higher specificity. Second, the pre-existing peptide-binding site of the primary domain provides weak but significant affinity and specificity for the targeted peptide motif, and thus Affinity Clamping converts the challenge of *de novo* generation of peptide-recognition surfaces into maturation of affinity and specificity.

3. Design of affinity clamps

The major breakthrough associated with Affinity Clamping is more conceptual than technical. Thus, we will focus this article primarily on descriptions of key concepts and precautions pertinent to setting up an Affinity Clamping project.

3.1. Choice of the primary domain

Affinity Clamping takes advantage of the fact that many small domains exist in nature that bind to a particular class of short peptide motif. Many complex signal transduction behaviors are mediated by regulatory proteins with highly modular architectures (Pawson and Nash, 2003). These proteins are built from linear combinations of individually folded domains (or “modules”), each of which performs a specific function. Close to 100 commonly occurring modular domains have been identified in eukaryotic proteins. Among these, the interaction domains are small proteins, usually less than 100 amino acid residues in length, and are independently folded and stable. Examples include PDZ, WW and SH2 domains. These interaction domains usually bind to small peptides containing a chemical signature, e.g. free C-terminus, phosphorylation or other post-translational modification. Importantly, the target peptide binds to a shallow cleft on the surface of these interaction domains (Pawson and Nash, 2003), and consequently approximately half of the peptide surface is still exposed for making additional interactions with the enhancer domain.

Although it is straightforward to choose a particular domain family as a potential primary domain for Affinity Clamping (e.g. PDZ domain for the recognition of protein C-termini), it is not trivial to select a specific domain within a family. Because affinity clamps will ultimately be used in practical applications, they should be robust and easy to produce. Therefore, biophysical and chemical properties such as conformational stability, solubility and the absence of reactive Cys residues are important aspects to consider, in addition to binding specificity. Such information is often unavailable in the literature and one may need to use guesswork from available, circumstantial data. For example, the availability of crystallographic and/or NMR structures is a good indication that the domain of interest has good biophysical properties. Several recent studies have systematically and comprehensively characterized functional and biophysical properties of domain family members (Machida et al., 2007; Tonikian et al., 2008). Clearly, such studies are highly valuable for choosing a primary domain.

3.2. Molecular display of the primary domain

Affinity Clamping requires the use of a directed evolution technology, such as phage display or yeast surface display (Koide et al., 2012), to produce large sequence diversity in the enhancer domain and to identify clones from such a repertoire. Therefore, a primary domain of choice must be compatible with the directed evolution technology of choice. Here, we will limit our discussion to phage display, although the concept is applicable to other directed evolution technologies.

The first test is to confirm that the primary domain of choice is robustly displayed on the surface of phages. Briefly, the gene for the domain is cloned in a phage display vector and an epitope tag, such as the V5 or FLAG tag, is added adjacent to the domain (Wojcik et al., 2010). The gene for the primary domain can be obtained from a natural source or can be synthesized. As the cost of gene synthesis continues to decline, it is often beneficial to use a synthetic gene that can be easily optimized for codon usage and for the inclusion and exclusion of convenient restriction enzyme sites. After confirming the DNA sequence of the phage display vector, phage particles are produced, and the level of surface display of the cloned domain is tested by performing phage ELISA against an anti-epitope tag antibody (Sidhu et al., 2000).

After surface display of the cloned primary domain is confirmed, phage ELISA is repeated using a peptide known to bind to the domain as the target. A peptide is synthesized with a biotin attached at a terminus (see 3.6 for target preparation), and the biotinylated peptide is immobilized to microtiter wells that have been coated with streptavidin and blocked with BSA. Because the interaction between the domain and peptide is generally weak, the phage ELISA signals may be much weaker than those for phage ELISA with an anti-epitope tag antibody. Still, it is important to confirm functional display of a domain of interest on the phage surface.

Although molecular display technologies for antibody fragments (single-chain Fv and Fab) have been refined over the years to be quite robust, it is important to recognize that one must fine-tune a molecular display system for individual proteins. Thus, extensive experimentation may be required to achieve a good level of surface display for the primary domain of your choice. Important parameters include: plasmid copy number; the promoter that drives the expression of the domain fused with a display partner (e.g. the phage p3 protein); signal sequence; growth temperature; aeration; induction method and the length of expression. See Supplementary Materials of Wojcik et al. (2010) for an example of systematic optimization of a phage display system.

If functional display of the primary domain cannot be confirmed after adjusting experimental parameters, one needs to make a strategic decision. One may wish to try directed evolution of the chosen primary domain, or test other family members with similar function. Also it is useful to experimentally characterize the conformational stability of the chosen domain in the exact format used for molecular display, because the stability of protein domains is often sensitive to the choice of boundary positions (Hamill et al., 1998). The definition of a domain discerned based on homology can be too strict, and constructs designed based on such a definition may eliminate terminal residues critical for maintaining stability. Together, one must realize that establishing a molecular display system for a new protein can be a lengthy and laborious process and it often requires intimate knowledge of the function, structure and stability of the protein, before embarking an Affinity Clamping project.

3.3. Choice of the enhancer domain

The primary role of the enhancer domain is to present a binding surface of sufficient size and good complementarity for forming the interaction with a peptide target bound on the primary domain (Figure 1). Because most Affinity Clamping projects would aim to generate high-affinity and high-specificity Affinity Clamps to the multitude of targets that the primary domain can recognize, the enhancer domain should be able to present diverse shapes and chemistry. This requirement is essentially identical to that for establishing an antibody library or an antibody-mimic library. Because Affinity Clamps contain a primary domain with preexisting affinity to a class of target peptides, the level of affinity enhancement required from an enhancer domain is much lower than that required to generate *de novo* target binding using the enhancer domain alone. This reduced requirement may give protein designers greater freedom in choosing an enhancer domain and/or in designing combinatorial libraries. Affinity Clamps to be developed should ideally have several attractive attributes so as to make them versatile and convenient tools. Such attributes include small size, high stability, the absence of disulfide bonds, ease of production and compatibility with molecular display methods (Binz et al., 2005; Koide, 2010). Consideration of these attributes can help a protein designer select an enhancer domain suitable for her/his methodologies.

The so-called non-antibody scaffolds serve as particularly attractive candidates for enhancer domains. A number of such non-antibody scaffolds have been developed for the purpose of generating antibody-like, target-binding proteins that do not inherit undesirable properties associated with conventional antibodies (Binz et al., 2005; Koide, 2010). A diverse array of protein scaffolds, including fibronectin type III (FN3) domains (“monobodies”), designed ankyrin repeat proteins (“DARPs”), lipocalins (“anticalins”) and helix-bundle proteins (“affibodies”), have been established as viable platforms for generating high-performance binding proteins. In principle, any of these non-antibody scaffolds could be used as enhancer domains.

We have successfully used a human FN3 domain as the enhancer domain in constructing Affinity Clamps (Huang et al., 2008; Huang et al., 2009a). FN3 is a small (~10 kDa) immunoglobulin-like β -sandwich protein with three loops located at one end of the molecule that can be extensively diversified to create a library of binding surfaces. Since the Koide group introduced it as a protein scaffold in 1998 (Koide et al., 1998), the FN3 domain has become the most widely used non-antibody scaffold and has been used to generate binders against a wide range of targets (Koide et al., 2012). FN3 has a highly stable core and it lacks disulfide bonds. FN3 domains are often found in multi-domain proteins, and thus FN3 is inherently capable of functioning in the presence of another domain fused to either of its termini. A large number of crystal structures of FN3-based binding proteins, termed monobodies, have revealed inherent plasticity of the three recognition loops (Gilbreth and Koide, 2012). At a more practical level, our group has invested significant effort in establishing effective molecular display systems for FN3 and in constructing high-performance combinatorial libraries that utilize the concept of highly biased amino acid compositions (Gilbreth et al., 2008; Koide et al., 2007; Wojcik et al., 2010). Naturally such know-how and the availability of reagents are invaluable in establishing an Affinity Clamp project.

One could use a protein domain that has not been validated as a molecular scaffold as an enhancer domain. However, one may need to expend considerable effort in establishing and customizing a molecular display format.

3.4. Linking the primary and enhancer domains

Once the primary and enhancer domains are chosen, their genes need to be connected to form a fusion protein. Most importantly, the two domains should be linked in such a way that the diversified surfaces of the enhancer domain can make extensive contacts with the target peptide presented on the primary domain (Figure 1). This condition can be easily met if either terminus of the primary domain is located near its peptide-binding site. In such a case, a short linker comprising of a stretch of Gly and Ser residues can be added between the two domains to provide linkage and flexibility between them.

The study of the role of the linker in PDZ-based affinity clamps showed that the inter-domain linker impacts the range of accessible inter-domain geometries and thus it is an important parameter to consider in designing affinity clamps (Huang et al., 2009a). If the linker is too short, the diversified surfaces of the enhancer domain may be able to make contacts only with a small portion of the target peptide. If the linker is too long, there may be too much conformational freedom, making it difficult to achieve high affinity. However, it is difficult to determine the single best length of the linker. Modeling utilizing molecular graphics is useful for estimating a reasonable range of linker lengths. One can introduce diversity in the linker length as a part of combinatorial library design.

Unfortunately, most of the modular domains found in nature have their termini located on the opposite side of the peptide-binding site (Pawson and Nash, 2003). This architecture is advantageous for ensuring the modularity of these domains, because changes in their adjacent domains would minimally affect their target-binding function. However, it is clearly a major problem in constructing the clamp architecture, because simple concatenation would place the enhancer domain far away from the peptide-binding site of the primary domain (Figure 2). In Affinity Clamping, the goal is to design new architecture that substantially affects the binding function of the primary domain, a goal that is fundamentally in conflict with the natural design of modular interaction domains. Therefore, although the functions of the modular interaction domains make them highly attractive primary domains, their architecture presents a design challenge.

Fortunately, another common attribute of the modular interaction domains provides a general solution to the topological challenge. The N- and C-termini of the modular interaction domains are usually juxtaposed (Pawson and Nash, 2003). Thus, one can relocate the termini by circular permutation, i.e., by connecting the original termini and disjoining another location to create a new set of termini. For the PDZ domains, there is a subset of family members, the HtrA, that are naturally circularly permuted with respect to the rest of the PDZ family members (Runyon et al., 2007), indicating that circular permutation is evolutionarily accessible and so should also be by protein design.

To perform circular permutation, one needs to identify a location at which the primary domain polypeptide may be disjoined so as to create the new termini. The location of the new termini needs to be carefully determined so that it minimally impacts the structure, function and stability of the primary domain. Because disruption of a helix or strand most likely leads to a substantial loss of stability and also structural distortions, it makes sense to choose a position within a surface loop or turn for disjoining. Because a loop or a turn may also be important for stability, not all such positions are suitable. We have developed a simple but effective strategy for identifying such a location (Huang et al., 2008). We first analyze the three-dimensional structure of the chosen primary domain (or a close homolog) and the sequence variability among the domain family members. We choose several locations as candidates that are (i) close to the peptide binding site, (ii) located within a loop/turn, and (iii) not highly conserved. We then insert a segment of four Gly residues to each of the candidate positions and examine the effects of the insertions on target binding using

standard techniques such as ELISA and SPR. We select the location where the insertion mutation exhibits the minimal effect as the new termini. Figure 2 outlines these processes.

The next step is to design a sequence with which the original termini are linked. Even when a three-dimensional structure is available, it often is not obvious where exactly a domain “starts” and “ends”. Although one could blindly link the termini with a flexible linker, e.g. several Gly residues, a long, flexible linker can reduce the stability of the redesigned domain because a long linker may have large conformational entropy that is lost upon folding. If the linker is too short, or if one removes a residue near the termini that is important for stability, the redesigned protein may be distorted or severely destabilized. We examine the crystallographic B-factors of residues near the termini to assess the level of flexibility. NMR data are also valuable. Based on such examination, we remove residues that are likely to be highly mobile and connect the termini with a linker that has a high propensity to form a turn, such as Asn-Gly (Ramirez-Alvarado et al., 1996). The function of the circularly permuted protein is then tested. In the case of erbin PDZ, this operation resulted in an ~10-fold loss in binding affinity (Huang et al., 2008). Finally, the gene for the enhancer domain is attached to the gene for the redesigned primary domain. In most cases, the attachment of an inert enhancer domain causes little perturbation of the function of the primary domain.

3.5. Combinatorial library construction

Once molecular display of the fusion construct of the primary and enhancer domains is established, the system is ready for constructing a combinatorial library. Whereas amino acid diversity can be introduced into both primary and enhancer domains, our initial efforts have focused on diversification of the enhancer domain (Huang et al., 2008; Huang et al., 2009a). Methods for constructing a combinatorial library depend on the format of molecular display and the enhancer domain. Those for the FN3 scaffold have recently been detailed in another volume of this series (Koide et al., 2012). Regardless, we recommend the use of highly biased amino acid compositions toward Tyr and varying the loop length, both of which are highly effective in generating high-affinity binding interfaces (Koide and Sidhu, 2009).

3.6. Choices and preparation of targets

Because Affinity Clamping is essentially affinity maturation of the primary domain, the protein designer already has a good idea of what peptide motifs to target by the time she/he has chosen the primary domain. In many cases, it is most convenient to produce a target peptide via chemical synthesis. Affinity Clamps can extend the recognition motif beyond that recognized by the primary domain, so it is important to prepare a peptide target that is extended from the core recognition motif of the primary domain in both N-terminal and C-terminal directions. For immobilization required for library sorting, we typically conjugate biotin to peptide targets, which can be readily achieved during chemical synthesis. Alternatively, simple peptide targets that do not contain post-translational modification can be produced as a recombinant protein. For this purpose, we have found yeast small ubiquitin-like modifier (SUMO) fusion proteins with a His-tag to be particularly convenient, because they are highly expressed, easily purified and easily quantified (Huang et al., 2008; Malakhov et al., 2004). A SUMO fusion protein can be further biotinylated by chemical modification or by biosynthetic biotinylation using the Avi-tag and biotin ligase (www.avidity.com).

It is useful and important to prepare additional peptides that are highly homologous to the actual target peptide. They can be used for assessing the level of specificity of Affinity Clamps and also as competitors during library sorting. Such peptides can be designed by searching a proteome(s) of interest with the recognition pattern of the primary domain.

Internet tools, such as Scansite (scansite.mit.edu) (Obenauer et al., 2003), are particularly useful for this purpose.

3.7. Library sorting

Phage display sorting is performed using a biotinylated peptide target and streptavidin-coated magnetic beads (Huang et al., 2008; Wojcik et al., 2010). Although direct immobilization of a target molecule to a polystyrene plate is commonly employed in antibody generation using phage display, it is not recommended for peptide targets for Affinity Clamps, because most of the peptide surface needs to be available for recognition by Affinity Clamps. One must carefully adjust library sorting conditions so as to be able to discriminate improved variants from the majority of clones that are not improved from the starting primary domain but nevertheless have detectable binding to the target. On the one hand, if the sorting conditions are too generous, one would simply be diluting the library without enriching the best clones. On the other hand, if the conditions are too stringent, even improved variants would be lost. This requirement is distinctly different from those for *de novo* generation of binding proteins in which a small number of functional variants are recovered in the presence of the vast majority of nonfunctional variants.

In order to find appropriate conditions, one needs to first perform a set of control experiments in which the target concentration is systematically varied and the recovery of phages displaying the primary domain and a functionally inert enhancer domain is determined by titering. The goal here is to find a target concentration where the phage recovery starts to increase above the background level (Figure 3). The target concentration for actual library sorting can then be set at a value slightly lower than the identified threshold. Clearly, library sorting needs to be highly reproducible for this type of optimization to be effective. We have found that automated handling of magnetic beads using a robot (Thermo Kingfisher) makes library sorting highly reproducible (Fellouse et al., 2007).

Typically a total of four rounds of library sorting are performed using decreasing target concentrations, e.g., 100, 20, 4 and 1 nM, respectively, although conditions depend on each system, as described above. Successful selection is confirmed with two parameters, enrichment ratio and hit rate. The enrichment ratio is the ratio of the number of phages recovered from the selection with a target to the number of phages from a parallel selection without the target. Usually an enrichment ratio of > 10 indicates a successful selection. The hit rate is simply the number of the binding-positive clones in a set of the randomly picked clones as tested in phage ELISA, and it is the most relevant to successful generation of Affinity Clamps. It is important to perform these tests at a target concentration where the phage displaying the starting primary domain does not exhibit measurable binding and to include the phage displaying the primary domain only (or the primary domain with an inert enhancer domain) as a negative control. The amino acid sequences of the hits are deduced by DNA sequencing.

When extremely high affinity is desired, a cycle of affinity maturation can be performed (Huang et al., 2008). A secondary library is constructed from lead molecules obtained from the initial library sorting. A variety of methods are available for the design of such affinity maturation libraries, including extensive randomization of a single loop, DNA shuffling and error-prone PCR (Koide et al., 2012). The resulting library is sorted with further reduced concentrations of the target.

4. Production and characterization of Affinity Clamps

The Affinity Clamp genes are transferred from phage clones to an expression vector, and individual clones are expressed in *Escherichia coli* and purified as soluble proteins. Because Affinity Clamps are simple, single-polypeptide proteins, standard protocols for protein expression work well for them. Our default is to express Affinity Clamps as a fusion with a His₁₀ tag. Protein samples are purified with Ni affinity chromatography followed by a size-exclusion chromatography, such as Superdex 75 (GE Life Sciences). It is critically important to ensure that Affinity Clamp proteins are monomeric, because the enhancer domain could induce oligomerization of an affinity clamp construct, which can lead to an apparent affinity increase due to multivalency.

To determine the binding affinity of selected Affinity Clamps using SPR, we express them as a fusion with a His₁₀ tag for efficient immobilization on the Ni-NTA chip (BIAcore). Although short peptides can be used in SPR measurements, their small masses would result in low sensitivity. Thus, we utilized target peptides fused to the SUMO protein (Koide et al., 2012). Typically, an Affinity Clamp is immobilized on the chip at a level of 100–300 RU. High-affinity interactions are measured in a kinetic method, whereas for measuring weak binding affinity, such as for the epitope clamp prototype, an equilibrium method may be required.

6. Applications of Affinity Clamps

Because of their high affinity, high specificity and high stability, Affinity Clamps can be excellent affinity reagents in biological and biomedical research. Immediate and obvious utilities of Affinity Clamps are as antibody alternatives in applications such as immunoprecipitation and Western blotting. Because the target peptides for Affinity Clamps usually are a short peptide motif recognized by a modular interaction domain, they are likely to be accessible in the context of the full-length protein. Thus, unlike antibodies that are raised using a peptide derived from within a folded domain, Affinity Clamps are likely to recognize their targets in both native and denatured states, which was indeed the case for the PDZ clamps (Huang et al., 2008).

Affinity Clamps are single polypeptides and also modular, making them particularly suitable as building blocks for making fusion proteins. A matched pair of an Affinity Clamp and its target peptide can be used as a building block for forming tight and stable protein-protein interactions. We have developed such a pair specifically for this purpose by optimizing the peptide sequence for binding to one of the PDZ clamps. This pair has been demonstrated to be highly effective in protein immobilization and labeling in single molecule measurements (Huang et al., 2009b) and in SPR analysis of antibody-antigen interactions (Dyson et al., 2011). The target peptide was also incorporated within the LOV2 domain in such a way to confer light-dependent control of protein-protein interaction for synthetic biology experiments (Strickland et al., 2012).

By further expanding the idea of using Affinity Clamps as building blocks for higher functionality, we have developed ratiometric fluorescence sensors (Huang and Koide, 2010). Inter-domain movements upon ligand binding are a common mechanism underlying allostery. The architecture of Affinity Clamps immediately suggests such ligand-induced conformational changes. By attaching a pair of FRET-optimized fluorescent proteins to a PDZ clamp, we were able to transduce conformational changes of the Affinity Clamp into changes in fluorescence emission.

7. Conclusion

The Affinity Clamp technology represents a major breakthrough in protein design, as it provides a rational pathway for producing high-performance affinity reagents for a predefined peptide motif. The presence of diverse families of modular interaction domains suggests that Affinity Clamps can be developed for a diverse panel of peptide motifs of high biological significance. Although the procedures for generating Affinity Clamps are still laborious and require high levels of skills in protein design, iterative refinement may lead to highly facile pipelines. The small and simple architecture of Affinity Clamps makes them ideal building blocks for designing functionalities beyond simple binding. We are confident that numerous Affinity Clamps will be developed for diverse purposes in the near future.

Acknowledgments

We thank Drs. Ryan Gilbreth, Akiko Koide, Robert Wells and Norihisa Yasui for critical reading of the manuscript. This work was supported in part by the National Institutes of Health grant R01-GM090324 to S.K. and by the University of Chicago Comprehensive Cancer Center.

References

- Bashton M, Chothia C. The generation of new protein functions by the combination of domains. *Structure*. 2007; 15:85–99. [PubMed: 17223535]
- Binz HK, Amstutz P, Pluckthun A. Engineering novel binding proteins from nonimmunoglobulin domains. *Nat Biotechnol*. 2005; 23:1257–1268. [PubMed: 16211069]
- Cobaugh CW, Almagro JC, Pogson M, Iverson B, Georgiou G. Synthetic antibody libraries focused towards Peptide ligands. *J Mol Biol*. 2008; 378:622–633. [PubMed: 18384812]
- Colwill K, Graslund S. A roadmap to generate renewable protein binders to the human proteome. *Nat Methods*. 2011; 8:551–558. [PubMed: 21572409]
- Dyson MR, Zheng Y, Zhang C, Colwill K, Pershad K, Kay BK, Pawson T, McCafferty J. Mapping protein interactions by combining antibody affinity maturation and mass spectrometry. *Anal Biochem*. 2011; 417:25–35. [PubMed: 21704603]
- Fellouse FA, Esaki K, Birtalan S, Raptis D, Cancasci VJ, Koide A, Jhurani P, Vasser M, Wiesmann C, Kossiakoff AA, et al. High-throughput Generation of Synthetic Antibodies from Highly Functional Minimalist Phage-displayed Libraries. *J Mol Biol*. 2007; 373:924–940. [PubMed: 17825836]
- 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]
- Gilbreth RN, Koide S. Structural Insights for Engineering Binding Proteins Based on Non-Antibody Scaffolds. *Curr Opin Struct Biol*. 2012 in press.
- Hamill SJ, Meekhof AE, Clarke J. The effect of boundary selection on the stability and folding of the third fibronectin type III domain from human tenascin. *Biochemistry*. 1998; 37:8071–8079. [PubMed: 9609701]
- Huang J, Koide A, Makabe K, Koide S. Design of protein function leaps by directed domain interface evolution. *Proc Natl Acad Sci U S A*. 2008; 105:6578–6583. [PubMed: 18445649]
- Huang J, Koide S. Rational conversion of affinity reagents into label-free sensors for Peptide motifs by designed allostery. *ACS Chem Biol*. 2010; 5:273–277. [PubMed: 20143785]
- Huang J, Makabe K, Biancalana M, Koide A, Koide S. Structural basis for exquisite specificity of affinity clamps, synthetic binding proteins generated through directed domain-interface evolution. *J Mol Biol*. 2009a; 392:1221–1231. [PubMed: 19646997]
- Huang J, Nagy SS, Koide A, Rock RS, Koide S. A Peptide tag system for facile purification and single-molecule immobilization. *Biochemistry*. 2009b; 48:11834–11836. [PubMed: 19928925]
- Koide A, Bailey CW, Huang X, Koide S. The fibronectin type III domain as a scaffold for novel binding proteins. *J Mol Biol*. 1998; 284:1141–1151. [PubMed: 9837732]

- Koide A, Gilbreth RN, Esaki K, Tereshko V, Koide S. High-affinity single-domain binding proteins with a binary-code interface. *Proc Natl Acad Sci U S A*. 2007; 104:6632–6637. [PubMed: 17420456]
- Koide S. Generation of new protein functions by nonhomologous combinations and rearrangements of domains and modules. *Curr Opin Biotechnol*. 2009; 20:398–404. [PubMed: 19700302]
- Koide, S. Design and engineering of synthetic binding proteins using nonantibody scaffolds. In: Park, SJ.; Cochran, JR., editors. *Protein Engineering and Design*. Boca Raton, FL: CRC Press; 2010. p. 109-130.
- Koide S, Koide A, Lipovsek D. Target-Binding Proteins Based on the 10th Human Fibronectin Type III Domain ((10)Fn3). *Methods Enzymol*. 2012; 503:135–156. [PubMed: 22230568]
- Koide S, Sidhu SS. The importance of being tyrosine: lessons in molecular recognition from minimalist synthetic binding proteins. *ACS Chem Biol*. 2009; 4:325–334. [PubMed: 19298050]
- Kortemme T, Baker D. Computational design of protein-protein interactions. *Curr Opin Chem Biol*. 2004; 8:91–97. [PubMed: 15036162]
- Machida K, Thompson CM, Dierck K, Jablonowski K, Karkkainen S, Liu B, Zhang H, Nash PD, Newman DK, Nollau P, et al. High-throughput phosphotyrosine profiling using SH2 domains. *Molecular cell*. 2007; 26:899–915. [PubMed: 17588523]
- Malakhov MP, Mattern MR, Malakhova OA, Drinker M, Weeks SD, Butt TR. SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J Struct Funct Genomics*. 2004; 5:75–86. [PubMed: 15263846]
- Obenauer JC, Cantley LC, Yaffe MB. Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res*. 2003; 31:3635–3641. [PubMed: 12824383]
- Parmley SF, Smith GP. Filamentous fusion phage cloning vectors for the study of epitopes and design of vaccines. *Advances in experimental medicine and biology*. 1989; 251:215–218. [PubMed: 2481962]
- Pawson T, Nash P. Assembly of cell regulatory systems through protein interaction domains. *Science*. 2003; 300:445–452. [PubMed: 12702867]
- Ramirez-Alvarado M, Blanco FJ, Serrano L. De novo design and structural analysis of a model beta-hairpin peptide system. *Nat Struct Biol*. 1996; 3:604–612. [PubMed: 8673604]
- Rockberg J, Lofblom J, Hjelm B, Uhlen M, Stahl S. Epitope mapping of antibodies using bacterial surface display. *Nat Methods*. 2008; 5:1039–1045. [PubMed: 19029907]
- Rossmann MG, Moras D, Olsen KW. Chemical and biological evolution of nucleotide-binding protein. *Nature*. 1974; 250:194–199. [PubMed: 4368490]
- Runyon ST, Zhang Y, Appleton BA, Sazinsky SL, Wu P, Pan B, Wiesmann C, Skelton NJ, Sidhu SS. Structural and functional analysis of the PDZ domains of human HtrA1 and HtrA3. *Protein Sci*. 2007; 16:2454–2471. [PubMed: 17962403]
- Sidhu SS, Fellouse FA. Synthetic therapeutic antibodies. *Nat Chem Biol*. 2006; 2:682–688. [PubMed: 17108986]
- Sidhu SS, Lowman HB, Cunningham BC, Wells JA. Phage display for selection of novel binding peptides. *Methods Enzymol*. 2000; 328:333–363. [PubMed: 11075354]
- Strickland D, Lin Y, Wagner E, Hope CM, Zayner J, Antoniou C, Sosnick TR, Weiss EL, Glotzer M. TULIPs: tunable, light-controlled interacting protein tags for cell biology. *Nat Methods*. 2012; 9:379–384. [PubMed: 22388287]
- 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. [PubMed: 18828675]
- Winter G, Griffiths AD, Hawkins RE, Hoogenboom HR. Making antibodies by phage display technology. *Annu Rev Immunol*. 1994; 12:433–455. [PubMed: 8011287]
- Wojcik J, Hantschel O, Grebien F, Kaupe I, Bennett KL, Barkinge J, Jones RB, Koide A, Superti-Furga G, Koide S. A potent and highly specific FN3 monobody inhibitor of the Abl SH2 domain. *Nat Struct Mol Biol*. 2010; 17:519–527. [PubMed: 20357770]

A. Conventional interface design **B. Affinity Clamping** **Figure 1.**

Schematics comparing conventional design of protein interaction interfaces (A) with Affinity Clamping (B). In conventional protein interface design, surfaces of an inert scaffold are diversified to produce an ensemble of variants (“library”) from which functional variants are identified. In Affinity Clamping, a framework for constructing new protein architecture is designed by attaching an inert enhancer domain to a primary domain that has weak affinity to a desired class of targets. Next, surfaces of the enhancer domain are improved by directed evolution, resulting in the formation of a clamshell architecture that provides a much larger interface for the peptide target than the interface between the primary domain and the target.

Figure 2.

Solution of a common topological problem in constructing a clamshell framework with modular interaction domains by circular permutation. The erbin PDZ domain is used as an example (Huang et al., 2008). (A and B) Crystal structure of the erbin PDZ domain in two orthogonal views. Note that the termini are located on the opposite face of the molecule to the peptide-binding site. (B) Identification of new termini by insertion mutagenesis and functional characterization. The positions for Gly4 insertions are indicated with arrows and labeled with residue numbers. The effects of insertions at each location are also described. (C) The circularly permuted erbin PDZ domain. The labels show the original termini that are now connected and the newly created N- and C-termini.

Figure 3. Schematic illustration of the importance of fine-tuning library sorting conditions for identifying Affinity Clamps. Recovery efficiency of phages displaying a low-affinity primary domain (solid line) and that of a high-affinity clone, corresponding to an evolved Affinity Clamp, are shown as a function of target concentration used in library sorting. The concentration range in which the Affinity Clamp can be effectively identified is shown as the shaded box.