Design of protein function leaps by directed domain interface evolution

Jin Huang, Akiko Koide, Koki Makabe, and Shohei Koide*

Department of Biochemistry and Molecular Biology, University of Chicago, 929 East 57th Street, Chicago, IL 60637

Edited by David Baker, University of Washington, Seattle, WA, and approved March 4, 2008 (received for review February 3, 2008)

Most natural proteins performing sophisticated tasks contain multiple domains where an active site is located at the domain interface. Comparative structural analyses suggest that major leaps in protein function occur through gene recombination events that connect two or more protein domains to generate a new active site, frequently occurring at the newly created domain interface. However, such functional leaps by combination of unrelated domains have not been directly demonstrated. Here we show that highly specific and complex protein functions can be generated by joining a low-affinity peptide-binding domain with a functionally inert second domain and subsequently optimizing the domain interface. These directed evolution processes dramatically enhanced both affinity and specificity to a level unattainable with a single domain, corresponding to >500-fold and >2,000-fold increases of affinity and specificity, respectively. An x-ray crystal structure revealed that the resulting “affinity clamp” had clathrin architecture as designed, with large additional binding surface contributed by the second domain. The affinity clamps having a single-nanomolar dissociation constant outperformed a monoclonal antibody in immunochemical applications. This work establishes evolutionary paths from isolated domains with primitive natural interaction domain (referred to as “primary domain”) (14) with another unrelated domain (“enhancer domain”), followed by combinatorial optimization of the enhancer domain in a process that mimics natural sequence divergence under selective pressure (Fig. 1B). The resulting two-domain protein possesses an active site that is distinctly different from those of either starting domain. This method is distinct from concatenation of multiple interaction domains in the rewiring studies of signaling networks where individual domains retain their respective functions (15). It also differs from homologous recombination and DNA shuffling applied to members within a protein family (16, 17). Here we demonstrate the impact of directed domain interface evolution by using it to address two major challenges in the design of protein-based affinity reagents (18): development of high-affinity reagents for short peptide motifs and targeting affinity reagents to a predefined epitope. We collectively term this class of engineered molecules “affinity clamps.”

Results and Discussion

Affinity Clamp Design. In our design, the chosen primary domain is the Erbin PDZ domain (96 residues), one of the best characterized protein domains known (19). PDZ domains are small globular domains that primarily bind to C-terminal sequences in globular domains that primarily bind to C-terminal sequences in target proteins with low affinity. Erbin-PDZ binds with a low-micromolar dissociation constant (Kd) to the C termini of p120-related catenins [β-catenin and Armadillo repeat gene deleted in Velo-cardio-facial syndrome (ARVC)] (20) using a shallow cleft for peptide binding (Fig. 1C), a commonly observed mode of binding among interaction domains. For our enhancer domain, we chose the 10th fibronectin type III domain of human fibronectin (FN3; 91 residues). FN3 is a robust scaffold for producing antibody-like binding proteins with three loops located at one end of the molecule that can be extensively diversified to create a repertoire of binding surfaces (Fig. 1D) (21, 22). No known natural proteins contain both PDZ and FN3 combined, evolutionally unrelated domains. However, such a path has not been directly demonstrated or exploited in the design of new protein functions.

Here we present directed evolution of protein function using a new protein-engineering concept, termed “directed domain interface evolution,” which is inspired by natural protein evolution through domain combination. This method combines a natural interaction domain (referred to as “primary domain”) (14) with another unrelated domain (“enhancer domain”), followed by combinatorial optimization of the enhancer domain surface in a process that mimics natural sequence divergence under selective pressure (Fig. 1B). The resulting two-domain protein possesses an active site that is distinctly different from those of either starting domain. This method is distinct from concatenation of multiple interaction domains in the rewiring studies of signaling networks where individual domains retain their respective functions (15). It also differs from homologous recombination and DNA shuffling applied to members within a protein family (16, 17). Here we demonstrate the impact of directed domain interface evolution by using it to address two major challenges in the design of protein-based affinity reagents (18): development of high-affinity reagents for short peptide motifs and targeting affinity reagents to a predefined epitope. We collectively term this class of engineered molecules “affinity clamps.”


Conflict of interest statement: J.H., A.K., and S.K. are named as the inventors of a patent application on the technology described in this article.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2QBW).

*To whom correspondence should be addressed. E-mail: skoide@uchicago.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0801097105/DCSupplemental.

© 2008 by The National Academy of Sciences of the USA
combinatorial libraries are labeled. The termini are also labeled. Note that the structure of FN3 (PDB entry 1FNF). The loops that are diversified to construct the peptide as a stick model, illustrating the shallow binding pocket. The affected the PDZ function with an 1N7T; Fig. 1 new C terminus (residue numbering is according to PDB entry is evolutionarily accessible and relevant (23). Indeed, the Htr permutation can occur as a result of gene duplication, and so it affinity toward the ARVCF peptide. In nature, such circular permutation, with Pro21 as the new N terminus and Asp20 as the new N and C termini were created in the PDZ domain by circular proximal to the peptide-binding cleft of the PDZ domain. Thus, not orient FN3 in such a way that its recognition loops are located on the opposite side of the peptide-binding site (Fig. 1). Consequently, simply connecting FN3 at a PDZ terminus would not orient FN3 in such a way that its recognition loops are proximal to the peptide-binding cleft of the PDZ domain. This, new N and C termini were created in the PDZ domain by circular permutation, with Pro21 as the new N terminus and Asp20 as the new C terminus (residue numbering is according to PDB entry 1N7T; Fig. 1C and Fig. S1). This structural alteration mildly affected the PDZ function with an ∼10-fold reduction in the affinity toward the ARVCF peptide. In nature, such circular permutation can occur as a result of gene duplication, and so it is evolutionarily accessible and relevant (23). Indeed, the Htr family of PDZ domains has a topology similar to the circularly permuted erbin PDZ domain (24).

The C terminus of the circularly permuted PDZ (hereafter termed cpPDZ) and the N terminus of FN3 were connected with a five-residue linker (GGSGG). The resulting two-domain protein is termed cpPDZFN. As expected, this domain combination did not significantly affect the peptide-binding function of the PDZ domain (Table 1; note that the affinity decrease seen in cpPDZFN relative to PDZ in Table 1 is due to circular permutation). We then constructed a combinatorial phage-display library of ∼10⁶ independent sequences in which three surface loops of FN3 were diversified (Table 1). After three rounds of library sorting using an eight-residue peptide corresponding to the C-terminal sequence of ARVCF, two clones exhibiting high affinity to the ARVCF peptide were identified (termed ePDZ-a and ePDZ-b, respectively; “e” stands for “enhanced”; Table 1).

Affinity and Specificity of Affinity Clamps. The two ePDZ clones were then expressed as free proteins in Escherichia coli. They were produced at high levels and highly soluble. Their binding kinetics measured by surface plasmon resonance (SPR) showed that they bound to the ARVCF peptide with Kd values of ∼60 nM, a 500 times higher affinity than that of the naive cpPDZ-FN3 fusion protein (Fig. 2B and Table 1). A cycle of affinity maturation of ePDZ-b produced second-generation affinity clamps with Kd values in the single-nanomolar range and dissociation half-lives of nearly 1 h (termed ePDZ-b1 and ePDZ-b2, respectively; Fig. 2B and Table 1). These values are comparable to those found for antibody–antigen interactions. Importantly, the affinity enhancement of >6,000-fold relative to cpPDZ (>500-fold relative to wild-type PDZ) (Table 1) by the affinity clamp strategy is far superior to the enhancement achieved by simple optimization of the peptide-binding interface of another PDZ domain alone (25), demonstrating the capacity of directed domain interface evolution to acquire function that is otherwise unattainable by manipulating only the primary domain.

Directed domain interface evolution was also highly effective in enhancing binding specificity. The homologous peptide targets ARVC and δ-catenin (Fig. 2A) provide a stringent test for the binding specificity of affinity clamps. Whereas the parent PDZ domain and ePDZ-a did not greatly discriminate the two peptides, the ePDZ-b family bound very weakly to δ-catenin (Kd > 25 nM), discriminating the two targets by as much as 6,000-fold (Fig. 2B and Table 1). Interestingly, the affinity of ePDZ-b1 and -b2 toward the δ-catenin peptide was weaker than that of the parent PDZ domain, suggesting that the enhancer domain can not only enhance the affinity toward a cognate target but also reduce the binding affinity of the primary domain, probably by competing against a noncognate target.

In the absence of the attached PDZ domain, the FN3 variants of these affinity clamps showed no detectable binding to the ARVCF peptide (data not shown). These results indicate that the FN3 domain of the affinity clamps recognize their target only when it is connected to the primary domain.

The X-Ray Crystal Structure of an Affinity Clamp. To determine whether the observed enhancements in affinity and specificity were due to successful construction of the designed architecture, we then characterized the structure of an affinity clamp. The NMR spectrum of ePDZ-a showed excellent dispersion, indicative of a well structured protein. An addition of the ARVCF peptide caused significant changes in the spectrum, consistent with the presence of a large binding interface (Fig. S2).

The x-ray crystal structure of the ePDZ-a/ARVC peptide complex at 1.8-Å resolution revealed clamshell architecture as designed (Fig. S4; statistics are given in Table S1), and the peptide was deeply buried at the domain interface (Fig. 3B). The circular permutation did not significantly alter the overall fold of domains, and the naive FN3 scaffold used here has no significant affinity to the PDZ targets.

We first sought to address the topological challenge in constructing the clamshell architecture necessary for the formation of the domain interface surrounding the active site(Fig. 1B and supporting information (SI) Fig. S1). As is common among interaction domains (14), the N and C termini of Erbin-PDZ are located on the opposite side of the peptide-binding site (Fig. 1C). Consequently, simply connecting FN3 at a PDZ terminus would not orient FN3 in such a way that its recognition loops are proximal to the peptide-binding cleft of the PDZ domain. Thus, new N and C termini were created in the PDZ domain by circular permutation, with Pro21 as the new N terminus and Asp20 as the new C terminus (residue numbering is according to PDB entry 1N7T; Fig. 1C and Fig. S1). This structural alteration mildly affected the PDZ function with an ∼10-fold reduction in the affinity toward the ARVCF peptide. In nature, such circular permutation can occur as a result of gene duplication, and so it is evolutionarily accessible and relevant (23). Indeed, the Htr

Fig. 1. The concept of directed domain interface evolution and building blocks used in this work. (A and B) Comparison of domain interface engineering with conventional protein engineering. In the conventional engineering that mimics gene duplication and sequence divergence (A), the interface predefined in the starting scaffold is altered/defined, which tends to produce incremental changes in function. In contrast, domain interface engineering that mimics gene combination and sequence divergence (B) produces a new functional site at the interface between two domains, which can result in a major leap in protein function. (C) The structure of the Erbin PDZ bound to a peptide (PDB entry 1MFG). The N and C termini are indicated. The positions for the new termini of the circularly permutated PDZ (cpPDZ) are shown with a triangle and residue numbers. Right shows the surface of the PDZ domain with the peptide as a stick model, illustrating the shallow binding pocket. (D) The structure of FN3 (PDB entry 1FNF). The loops that are diversified to construct combinatorial libraries are labeled. The termini are also labeled. Note that the N terminus and the recognition loops are located on the same side of the FN3 protein.
Table 1. Library design and binding parameters of affinity clamps

<table>
<thead>
<tr>
<th>Protein</th>
<th>BC loop (25–30)*</th>
<th>DE loop (52–55)*</th>
<th>FG loop (75–83)*</th>
<th>$k_{on}$, M$^{-1}$S$^{-1}$ (ARVCF)</th>
<th>$k_{off}$, S$^{-1}$ (ARVCF)</th>
<th>$K_d$, nM</th>
<th>SPR ARVCF</th>
<th>SPR δ-catenin</th>
<th>Affinity enhancement</th>
<th>Specificity index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Library*</td>
<td>X4–8 VX</td>
<td>(SY)(GSY) (SY)$^2$</td>
<td>X6–14</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>PDZ</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>ND</td>
<td>ND</td>
<td>2,200</td>
<td>6,300</td>
<td>3,000</td>
<td>—</td>
<td>3</td>
</tr>
<tr>
<td>cpPDZFN</td>
<td>SSSSVS</td>
<td>GSKS</td>
<td>SSSSSSSSS</td>
<td>ND</td>
<td>ND</td>
<td>24,800</td>
<td>&gt;25,000</td>
<td>ND</td>
<td>1</td>
<td>ND</td>
</tr>
<tr>
<td>ePDZ-a</td>
<td>SYYGVS</td>
<td>YSSS</td>
<td>YSDYGGSHHY</td>
<td>2.9 × 10$^5$</td>
<td>1.5 × 10$^{-2}$</td>
<td>56 ± 5</td>
<td>429</td>
<td>40</td>
<td>520</td>
<td>8</td>
</tr>
<tr>
<td>ePDZ-b</td>
<td>YYDSH VS</td>
<td>GSKS</td>
<td>HYNYHYYS</td>
<td>1.9 × 10$^5$</td>
<td>1.1 × 10$^{-2}$</td>
<td>56 ± 6</td>
<td>&gt;25,000</td>
<td>59</td>
<td>520</td>
<td>446</td>
</tr>
<tr>
<td>ePDZ-b1</td>
<td>YRELPS</td>
<td>GSKS</td>
<td>HYHYHYYS</td>
<td>7.3 × 10$^4$</td>
<td>&lt;3.7 × 10$^{-4}$</td>
<td>&lt;5</td>
<td>&gt;25,000</td>
<td>ND</td>
<td>4,960</td>
<td>&gt;5,000</td>
</tr>
<tr>
<td>ePDZ-b2</td>
<td>FTDLPLVS</td>
<td>GSKS</td>
<td>HYHYHYYS</td>
<td>7.0 × 10$^4$</td>
<td>&lt;2.9 × 10$^{-4}$</td>
<td>&lt;4</td>
<td>&gt;25,000</td>
<td>ND</td>
<td>&gt;6,200</td>
<td>&gt;6,250</td>
</tr>
</tbody>
</table>

*The residue numbering is according to that in Koide et al. (21).
†The $k_{on}$ and $k_{off}$ are shown for the ARVCF peptide. $I_{50}$ for the ARVCF peptide was determined by using competitive phage ELISA.
‡The specificity index is defined as the ratio of the $K_d$ for δ-catenin in the $K_d$ for ARVCF.
§The combinatorial library was constructed by diversifying the BC, DE, and FG loops of FN3. X denotes an amino acid mixture consisting of 40% Tyr, 20% Ser, 10% Gly, and 5% each of A, D, H, L, N, and R.
¶Taken from Zhang et al. (42).

the PDZ domain (Fig. 3D), and consequently the peptide–PDZ interactions were well conserved. Likewise, the FN3 structure was well maintained except for the mutated loops.

The interface structure rationalized the dramatic affinity enhancement. The affinity clamp buried 69% (844 Å$^2$) of the solvent-accessible surface area of the peptide, a 3.5-fold increase with respect to that by the PDZ domain only. The three mutated loops of the FN3 domain all interacted with the peptide and/or the PDZ domain, with a binding interface diagonally covering the surfaces of both cpPDZ and the peptide (Fig. 3C), reminiscent of the mode of recognition of the MHC/peptide complex by the T cell receptor (26). The newly introduced contacts mediated by the FN3 loops are predominantly hydrophobic, with only three hydrogen bonds between FN3 and the peptide. Details of the interactions are shown in Fig. S3. The shape complementarity between FN3 and the cpPDZ/peptide complex was unusually high with a shape correlation ($S_c$) value (27) of 0.79. Thus, the significantly enlarged interface with tight packing provides the enhanced affinity of ePDZ-a to the target peptide.

**Affinity Clamps as Antibody Alternatives.** The affinity clamps are excellent tools in common immunochemical analyses. The affinity clamps are robust and stable proteins. They remained monomeric and retained full activity after incubation at 50°C for 2 h (Fig. 4 A and B). Both ePDZ-b1 and ePDZ-b2 in the form of an alkaline phosphatase fusion protein outperformed a monoclonal antibody that was raised against a C-terminal region of ARVCF (28) in Western blotting for ARVCF expressed in mammalian cells (Fig. 4C). They specifically precipitated a protein tagged with the ARVCF peptide from E. coli lysate with nearly stoichiometric efficiency, whereas naive cpPDZ did not precipitate an appreciable amount of the target (Fig. 4D). These results demonstrate the effectiveness of affinity clamps as an antibody alternative. Because affinity clamps are recombinant proteins, they offer additional favorable attributes including the “monoclonal” nature, the ease of production, storage and distribution, the immediate access to amino acid and DNA sequences, and the ease of reformatting into various fusion proteins.

**Conclusions**

Our results demonstrate evolutionarily accessible paths where major functional leaps are produced by domain combination followed by interface optimization. They imply that distinct and sophisticated functions can emerge from combinations of a limited number of primordial domains. Directed domain interface evolution considerably expands the scope of protein engineering. It provides a rational guideline for producing high-performance affinity reagents targeted to a
predefined target motif. These unique features offer a clear advantage over conventional antibodies and antibody mimics whose epitopes cannot be defined a priori. Clearly, binding specificity of affinity clamps can be significantly altered by replacing the primary domain or introducing mutations in the primary domain. There exist diverse families of interaction domains that possess characteristics similar to the PDZ domain used in this work (14). Therefore, we anticipate that affinity clamps with novel functions can be readily produced by using directed domain interface evolution. Such affinity clamps will be effective tools in many fields including proteomics, cell biology, and drug discovery.

**Experimental Procedures**

**Preparation of Peptide Targets.** Target peptides corresponding to the C-terminal eight residues of ARVCF and β-catenin (NH₂-PASPDSWV-COOH and NH₂-PASPDSWV-COOH, respectively) were prepared as C-terminal fusion to the yeast SUMO (29) and cloned in the pHFT2 vector. The pHFT2 vector is a derivative of pHHT1 that contains a His₁₀ tag instead of a His₆ tag (30). The SUMO-peptide fusion proteins were cleaved with SUMO-hydrolase (29), and the cleaved peptide was purified by reversed-phase HPLC.

The ARVCF peptide was synthesized by using standard Fmoc protocols and purified by reversed-phase HPLC. The peptide was biotinylated by using sulfo-NHS-SS-Biotin (Pierce) following the protocols provided by the manufacturer and purified by using reversed-phase HPLC. The purity and identity of the peptide were verified by liquid chromatography/mass spectrometry (LC/MS).

**Construction of Phage Display Vectors and Combinatorial Libraries.** The locations of the new PDZ termini were chosen based on effects of Gly₄ insertion mutagenesis. The phage display vectors contained the DsbA signal sequence (31), the circularly permuted Erbin PDZ domain, and the FN3 domain (22), in which the PDZ and FN3 domains were connected with a GGSGG linker. A stabilization mutation, DTN (32), was also introduced to FN3. Combinatorial libraries in which the BC, DE, and FG loops of FN3 were diversified were constructed by using Kunkel mutagenesis as previously described (22, 33, 34). Phage particles were prepared in the absence of isopropyl β-D-1-thiogalactopyranoside.

**Library Sorting.** A combinatorial library of ~10⁸ independent clones was sorted with the biotinylated ARVCF peptide as the target. Phage display sorting was performed by using streptavidin-coated magnetic beads (Streptavidin MagneSphere Paramagnetic Particles; Promega) and a Thermo Kingfisher instrument as previously described (33). A total of four rounds of sorting of the initial library were performed by using decreasing target concentrations of 100, 20, 4, and 1 nM. The amino acid sequences of affinity clamp clones were deduced by DNA sequencing.

The second-generation library of εPDZ-b was constructed by diversifying either only the BC or FG loop of εPDZ-b. The resulting library was sorted twice as described above, followed by two rounds of off-rate selection performed as follows. Phages were first mixed with the biotinylated target, and, after 15 min of incubation, 1,000-fold excess of the nonbiotinylated form of the SUMO-ARVCF peptide fusion protein was added as a competitor. Phages were captured with streptavidin-coated magnetic beads, and the magnetic beads were incubated in the washing buffer [TBS containing 0.05% Tween (TBST)] with vigorous mixing for 3 minutes. After a total of three washing periods, phages were recovered.

**Protein Purification.** The affinity clamp genes in the enriched pool of phage clones were transferred into the pHFT-2 expression vector, and individual clones were expressed as soluble proteins with a His₁₀ tag. Proteins were purified with Ni affinity chromatography followed by a Sephacryl S-100 column in PBS [50 mM sodium phosphate containing 150 mM NaCl (pH 7.4)] to ensure that proteins were monomeric. For protein crystallization, the N-terminal His₁₀ tag was cleaved with TEV protease and the cleaved protein was purified by passing through a Ni-Sepharose column. For NMR characterization, proteins were expressed in the M9 minimal media with ¹⁵NH₄Cl as the sole nitrogen source.

**Affinity Measurements.** SUMO fusion peptides were used to facilitate SPR detection. SPR measurements were performed in 20 mM Hepes (pH 7.4), 150 mM NaCl, and 0.005% Tween 20 on a Biacore 2000 instrument. An affinity clamp was immobilized on a Ni-NTA chip to the level of approximate 300 RU, and the interaction with a SUMO-peptide fusion protein was monitored. The lack of interaction between SUMO and the affinity clamps was confirmed through competition experiments using the synthetic ARVCF peptide in the phage display format.

**Competition Phage ELISA.** Biotinylated ARVCF peptide was immobilized in the wells of a Maxisorp plate (NUNC) that had been coated with neutravidin and blocked with BSA. Phages displaying an affinity clamp were preincubated for 45 min with serial dilutions of nonbiotinylated ARVCF peptide in 0.05% TBST...
Fig. 4. Stability and applications of affinity clamps. (A) Gel-filtration chromatograms of ePDZ-a before (dashed line) and after (solid line) heat treatment (2 hr at 50°C). Gel filtration was performed by using a Superdex75 column (Amersham Biosciences) in PBS (pH 7.4). (B) The SPR sensograms of ePDZ-a before (dashed line) and after (solid line) the same heat treatment as in A. (C) Western blotting of wild-type ARVCF in mammalian cell lysates. Lysates of an MDCK cell line stably expressing human ARVCF (denoted as -) (41) and the parent cell line (-) were detected with ePDZ-b2 fused with alkaline phosphatase. mAb indicates a positive control with anti-ARVCF monoclonal antibody 4B1 (41). Left shows Coomassie brilliant blue staining, and Right shows Western blotting. (D) Pull-down (immunoprecipitation) of SUMO tagged with the ARVCF peptide from E. coli lysate by affinity clamps. The lysate was mixed with an affinity clamp immobilized to streptavidin magnetic beads. SDS/PAGE of the input (I), unbound (U), wash (W), and bound (B) fractions visualized with Coomassie brilliant blue staining are shown. “Beads” indicates a control experiment without an immobilized affinity clamp. “cpPDZFN” indicates a control experiment using cpPDZ fused to the unmodified F3N scaffold. The position of the captured target is marked with the triangle for ePDZ-b2 (lane B), and the equivalent position is also marked for cpPDZFN. “Binder” indicates the position of ePDZ-b2 and cpPDZFN, and “SA” indicates the position of streptavidin.

and then captured in the wells containing the immobilized peptide. After incubating for 30 min, the plate was washed with TBST and incubated with horseradish peroxidase-conjugated anti-M13 antibody (Amersham) and detected with one-step Turbo-EELISA (Pierce).


Western Blotting. Alkaline phosphate fusion proteins of affinity clamps were made by cloning the alkaline phosphatase gene (a gift of Brian Kay, University of Illinois, Chicago, IL) (40) to the 3’ of the affinity clamp gene in the phage display vector described above in such a way that the alkaline phosphatase replaces the phage pill protein. The fusion proteins were expressed in XL1-Blue cells, and the periplasmic fraction containing AP-affinity clamp fusion proteins was used for detection.

The Madin–Darby canine kidney II (MDCK II) cell line stably transfected with a human ARVCF expression vector were kindly provided by Albert B. Reynolds (Vanderbilt University, Nashville, TN) and cultured as described (41). Cells were washed once with PBS at room temperature, then lysed in ice-cold lysis buffer (0.5% Nonidet P-40/50 mM Tris, pH 7.4/150 mM NaCl) containing protease inhibitor mixture. Lysates were cleared with centrifugation, and protein concentration was estimated with SDS/PAGE. Cell lysates were separated by SDS/PAGE using 4–20% Precise Gel (Pierce) and blotted to the Immobilon-FL membrane (Millipore). The membrane was probed with the alkaline phosphatase fusion protein described above using BCIP/NBT (5-bromo-4-chloro-3-indolyl-phosphate/nitro blue tetrazolium) Color Development Substrate (Promega). A monoclonal antibody for ARVCF (mAb 481; Santa Cruz Biotechnology) was used as a positive control, and its binding was detected with alkaline phosphatase-conjugated secondary antibody (Promega).

ACKNOWLEDGMENTS. We thank V. Tereshko, B. Pentelute, R. Jones, and W. Jun for assistance; A. Reynolds and B. Kay for reagents; and A. Kossiakoff and R. Gilbreth for discussion. This work was supported in part by National Institutes of Health Grants R01-GM072688 and R21-CA132700 and by the University of Chicago Cancer Research Center. The General Medicine and Cancer Institutes Collaborative Access Team (GM/CA CAT) has been funded in whole or in part with federal funds from the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Science (Y1-GM-1104). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract DE-AC02-06CH11357.

X-Ray Crystallography. The ePDZ-a ARVCF peptide complex was crystallized in 1.0 M (β)-2-propanol containing 1.1 M Tris/HCl (pH 9) by using the hanging drop vapor diffusion method. Crystals were cryoprotected in the mother solution containing 20% glycerol and flash-frozen in liquid nitrogen. The x-ray data were collected at APS beamline 24-ID (Advanced Photon Source at Argonne National Laboratory). X-ray diffraction data were processed with HKL2000 (36). The structure was determined by molecular replacement with the program MOLREP in CCP4 (37). The structures of an FN3 mutant (PDB ID code 2OBG) and Erbin PDZ domain (PDB ID code 1MFG) were used as the search models. Refmac5 (38) was used for the structural refinement. Model building was carried out by using the program Coot (39). The structure of engineered loops, linker, and ARVCF peptide were built at this stage. Molecular graphics were generated by using Pymol (www.pymol.org). Data collection and refinement statistics are listed in Table S1.

Pull-Down Assay. Affinity clamps were biotinylated as described above. A biotinylated affinity clamp was immobilized onto streptavidin-coated magnetic beads. After washing the beads with PBS (pH 7.4), E. coli BL21(DE3) cell lysate containing a trace amount of SUMO-ARVCF was added to the beads and incubated for 45 min. After further washing steps, the beads were resuspended in PBS and SDS/PAGE sample buffer. The captured proteins were analyzed by SDS/PAGE stained with Coomassie brilliant blue. A control experiment was carried out with the magnetic beads to which no affinity clamp had been added.
Fig. S1. The domain-interface engineering strategy for generating affinity clamps. (A) Step-by-step description of the procedures for affinity clamp engineering. The target (yellow) binds to the shallow groove part of the primary peptide-binding domain (gray). Because the N and C termini of the primary domain are usually juxtaposed and located on the opposite side of the peptide-binding site, new termini are created by circular permutation so as to allow the enhancer domain (white) to be located near the active site of the primary domain. The interface (asterisks) is diversified and optimized by combinatorial library selection (in our case by phage display). Finally, a selected affinity clamp is reformatted as an isolated protein. The primary and enhancer domains of an affinity clamp form clamshell architecture that clamps the target.
Fig. S2. $^1$H, $^{15}$N-HSQC spectra of the free form (black) and the peptide-bound form (red) of $^{15}$N-labeled ePDZ-a. NMR spectra were taken at 30°C on a Varian INOVA 600 spectrometer.
Key

- **Blue**: Bonds in the peptide
- **Orange**: Bonds in PDZ
- **Green**: Bonds in FN3
- **Red**: Hydrogen bond and its length
- **Black**: Corresponding atoms involved in hydrophobic contact(s)
- **Reddish Black**: PDZ residues involved in hydrophobic contact(s)
- **Greenish Red**: FN3 residues involved in hydrophobic contact(s)

Fig. S3. Schematic drawings of the interactions between the peptide and PDZ (A) and those between the peptide and the affinity clamp, ePDZ-a (B). The analysis was performed using LIGPLOT.
### Table S1. X-ray crystallography data collection and refinement statistics

<table>
<thead>
<tr>
<th>Data collection</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Space group</strong></td>
<td>I2_1,2_1</td>
</tr>
<tr>
<td><strong>Cell parameters</strong></td>
<td>(a = 54.82,\text{Å},\ b = 79.35,\text{Å},\ c = 119.84,\text{Å},\ \alpha = \beta = \gamma = 90^\circ)</td>
</tr>
<tr>
<td><strong>Beamline</strong></td>
<td>APS 23-ID-D</td>
</tr>
<tr>
<td><strong>Wavelength, Å</strong></td>
<td>0.9793</td>
</tr>
<tr>
<td><em><em>Resolution,</em> Å</em>*</td>
<td>50–1.8 (1.86–1.80)</td>
</tr>
<tr>
<td><strong>(R_{\text{merge}})</strong></td>
<td>0.112 (0.423)</td>
</tr>
<tr>
<td><em><em>Completeness,</em> %</em>*</td>
<td>94.3 (97.0)</td>
</tr>
<tr>
<td><strong>Redundancy</strong>*</td>
<td>4.3 (4.1)</td>
</tr>
<tr>
<td><strong>(I/s(I))</strong></td>
<td>18.76 (5.78)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Refinement statistics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unique reflections:</strong></td>
<td></td>
</tr>
<tr>
<td>Working set*</td>
<td>22,059 (1,652)</td>
</tr>
<tr>
<td>Free set*</td>
<td>1,164 (90)</td>
</tr>
<tr>
<td>(R) factor*</td>
<td>0.188 (0.204)</td>
</tr>
<tr>
<td>(R_{\text{free}})*</td>
<td>0.227 (0.234)</td>
</tr>
<tr>
<td><strong>No. of amino acid residues</strong></td>
<td>195</td>
</tr>
<tr>
<td><strong>No. of water molecules</strong></td>
<td>229</td>
</tr>
<tr>
<td><strong>Matthews coefficient</strong></td>
<td>3.0 Å³/Da (water content: 58.4%)</td>
</tr>
</tbody>
</table>

| rmsd from ideal values |     |
| Bonds, Å/angles, ° | 0.007/1.104 |

| Ramachandran plot statistics |     |
| Residues in most favored regions | 92.4% (146) |
| Residues in additional allowed regions | 7.0% (11) |
| Residues in generously allowed regions | 0.0% (0) |
| Residues in disallowed regions | 0.6% (1) |

*Highest-resolution shell is shown in parentheses.

\(R_{\text{merge}} = \frac{\sum_{hkl} |(hf_{hkl}) - \langle |hf_{hkl}| \rangle / \langle |hf_{hkl}| \rangle|^2}{\sum_{hkl} |hf_{hkl}|^2 / \langle |hf_{hkl}| \rangle^2}\) over \(i\) observations of a reflection \(hkl\).

\(R\) factor = \(\sum |F_{\text{obs}}| - |F_{\text{calc}}| / \sum |F_{\text{obs}}|\).