De novo selection of oncogenes

Kelly M. Chacón[a], Lisa M. Petti[a], Elizabeth H. Scheideman[a], Valentina Pirazzoli[b], Katerina Politi[b,c], and Daniel DiMaio[a,c,d,e,1]

[a]Department of Genetics, Yale School of Medicine, New Haven, CT 06520-8005; [b]Department of Pathology, Yale School of Medicine, New Haven, CT 06520-8023; [c]Yale Cancer Center, New Haven, CT 06520-8028; [d]Department of Therapeutic Radiology, Yale School of Medicine, New Haven, CT 06520-8040; and [e]Department of Molecular Biophysics and Biochemistry, Yale School of Medicine, New Haven, CT 06520-8024

Edited by Patrick S. Moore, University of Pittsburgh, Pittsburgh, PA, and approved November 19, 2013 (received for review August 12, 2013)

All cellular proteins are derived from preexisting ones by natural selection. Because of the random nature of this process, many potentially useful protein structures never arose or were discarded during evolution. Here, we used a single round of genetic selection in mouse cells to isolate chemically simple, biologically active transmembrane proteins that do not contain any amino acid sequences from preexisting proteins. We screened a retroviral library expressing hundreds of thousands of proteins consisting of hydrophobic amino acids in random order to isolate four 29-aa proteins that induced focus formation in mouse and human fibroblasts and tumors in mice. These proteins share no amino acid sequences with known cellular or viral proteins, and the simplest of them contains only seven different amino acids. They transformed cells by forming a stable complex with the platelet-derived growth factor β receptor transmembrane domain and causing ligand-independent receptor activation. We term this approach de novo selection and suggest that it can be used to generate structures and activities not observed in nature, create prototypes for novel research reagents and therapeutics, and provide insight into cell biology, transmembrane protein–protein interactions, and possibly virus evolution and the origin of life.

synthetic biology | protein engineering | receptor tyrosine kinase | E5 protein | traptamer

All extant naturally occurring proteins are derived from preexisting ones by the incremental process of natural selection. Cellular proteins are thus the products of a long chain of evolutionary decisions driven by chance mutations and selective pressures, which allowed some profitable lineages to flourish but drove many others to extinction. Indeed, because of the immense diversity that can result from the arrangement of 20 different amino acids into long polymers, the roster of distinct proteins on earth is a miniscule fraction of possible chemical structures (1). Because they are trapped by their evolutionary history, naturally occurring proteins are not necessarily the optimal proteins for accomplishing a given task. Therefore, one of the major goals of synthetic biology is to generate artificial proteins for a variety of practical applications, such as optimizing catalytic activity and even developing entirely new catalytic activities, generating molecules suitable for industrial processes, and fabricating novel biomaterials.

Computational methods have been developed to construct artificial proteins based on detailed understanding of the chemistry, energetics, and structure–function relationships of existing proteins (2–4). This process, known as de novo or computational protein design, typically starts from a known protein structure or a conserved motif and iteratively samples numerous parameters such as amino acid substitutions and side-chain rotamers to identify predicted low-energy states. Although this approach requires enormous computational power, it has been used successfully for both globular and transmembrane proteins. For example, Yin et al. (5) used computational design to build transmembrane peptides that bound the transmembrane domains of integrins $\alpha_{IIb}\beta_3$ and $\alpha_v\beta_3$.

Libraries expressing large collections of proteins with long randomized segments have also been successfully screened for active proteins (reviewed in refs. 6 and 7). In some cases, randomization schemes are constrained by an underlying design motif modeled on existing proteins. For example, Hecht and coworkers constructed a library that formed a stable 102-aa globular structure, a four-helix bundle, by incorporating alternating random stretches of polar and nonpolar amino acids at 56 positions (which formed the soluble exterior surface and hydrophobic core of the proteins, respectively), linked by loops of other randomized amino acids (1). This library was screened in bacteria for proteins that rescued defects caused by mutations in metabolic pathways. In a pioneering study using a library without any prior design constraints, Keefe and Szostak isolated 108-aa ATP-binding proteins from an mRNA display library in which 80 consecutive amino acids were totally randomized (8). In this approach, DNA with a randomized ORF was transcribed in vitro, and the resulting RNAs were ligated to a puromycin linker and translated in vitro, so that each protein product became covalently coupled to its cognate mRNA. In vitro enrichment of proteins with ATP-binding activity resulted in the recovery of mRNAs, which were amplified and subjected to serial rounds of synthesis and enrichment. Randomized mRNA display libraries are very large, ranging from $\sim 10^{11}$ to 1.5×10^{13} different sequences, necessitating up to 18 rounds of selection to obtain the desired activity (8–10). These approaches have been used to generate novel enzymes, protein–protein interactions, and proteins that bind metals and other small molecules (e.g., refs. 11–15), but the products of these efforts are often not tested in cells or have limited activity in vivo.

To isolate protein structures that have not been observed in nature, we developed a simple genetic method, termed de novo selection, in which we screened a relatively small retroviral library expressing randomized proteins. We reasoned that short proteins

Significance

Artificial proteins may have improved properties compared with proteins that arose during evolution, but approaches to construct active artificial proteins are cumbersome and often constrained by existing protein structures. Here, we used mouse cells to select proteins that formed tumors from a library of small transmembrane proteins with randomized hydrophobic amino acid sequences. The resulting oncoproteins lack amino acid sequences from any known protein and function by activating a cellular growth factor receptor. This approach can be used to generate structures not observed in nature, create prototypes for research and possibly clinical uses, and provide insight into cell biology, protein–protein interactions, and evolution.

Author contributions: K.M.C., L.M.P., E.H.S., V.P., K.P., and D.D. designed research; K.M.C., L.M.P., E.H.S., and V.P. performed research; K.M.C., L.M.P., E.H.S., V.P., K.P., and D.D. analyzed data; and K.M.C., L.M.P., E.H.S., V.P., K.P., and D.D. wrote the paper.

Conflict of interest: Yale University has filed a provisional patent covering the construction and use of novel traptamers for various purposes.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: daniel.dimaio@yale.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1315298111/-/DCSupplemental.

would be desirable for this purpose, to minimize the size of the library and the scale of the screen. The size of the library could be further reduced by restricting it to proteins lacking some of the naturally occurring amino acids. It is also preferable that a large fraction of the expressed proteins fold into stable structures to increase the frequency of active proteins in the library. By these criteria, short transmembrane proteins are ideally suited for de novo selection. Less than 25 aa are sufficient to span membranes, so the selected proteins will be quite small, and the very hydrophobic nature of these segments reduces the chemical diversity of the amino acid side chains in any given transmembrane domain. Furthermore, irrespective of amino acid sequence, most hydrophobic protein segments adopt a stable α-helical structure in the membrane environment to reduce the energetic cost of membrane insertion. Finally, the amino acid side chains of transmembrane segments can mediate highly specific protein–protein interactions, so biological selections can be designed to recover rare active sequences from a vast excess of inactive ones.

These features are exemplified by the 44-aa bovine papillomavirus (BPV) E5 protein, the smallest known naturally occurring oncoprotein (Fig. 1*A*). The E5 protein is essentially an isolated transmembrane domain, which interacts directly and specifically with the transmembrane domain of the platelet-derived growth factor (PDGF) β receptor, resulting in ligand-independent receptor activation and cell transformation (16–18). We previously constructed libraries expressing small proteins in which various portions of the E5 protein were replaced with randomized hydrophobic segments. These proteins are designated traptamers, for transmembrane aptamers. We isolated traptamers that specifically activate the PDGF β receptor or the erythropoietin receptor or down-regulate the HIV coreceptor, C-C chemokine receptor type 5 (CCR5) (19–24). However, because the E5 protein was selected during viral evolution for its ability to insert into cell membranes, fold properly, and interact with the PDGF β receptor, the E5 sequences retained in these traptamers may be required for biological activity.

Here, we used genetic selection to isolate artificial transmembrane proteins lacking any preexisting sequences. To engineer such proteins, minimal sequence constraints were imposed, namely a methionine codon in a favorable context for translation initiation, a randomized 25-residue primarily hydrophobic stretch to span the membrane and serve as a specific protein–protein interaction domain, and two C-terminal aromatic amino acids, which are frequently found at the end of transmembrane domains near the polar head groups of the lipid bilayer (25). By carrying out a single round of selection for transformation of mouse cells, we isolated four proteins with amino acid sequences that do not occur in nature but bind and activate the PDGF β receptor in a ligand-independent fashion, resulting in cell transformation and tumorigenicity. These results demonstrated that simple genetic selections can be used to isolate biologically active transmembrane proteins in the absence of a template.

A MPNLWFLLFLGLVAAMQLLLLLFLLLFFLVYWDHFECSCTGLPF

B MA**XXXXXXXXXXXXXXXXXXXXXXXXX**YW

C
6A-1: MA**ALLLVLLTVICIMFMIMIVILIIII**YW
9C-3: MA**IIVAILLLIVMLAIEILLLITLIIA**YW
12A-5: MA**LIVIAVSITVMILLLLFFLIGFVIV**YW
3A-2: MA**LLLTVIVLVLLIFIIVMILLVLLVL**YW

Fig. 1. Design of the UDv3 library and sequences of active small transmembrane proteins. (*A*) Amino acid sequence of the wild-type BPV E5 protein. (*B*) Design of the UDv3 library. Randomized positions are indicated by "X," flanked by the amino acids encoded by the Kozak start site (MA) and the C-terminal aromatic anchor (YW). (*C*) Amino acid sequences of the four transforming clones recovered from the library. All randomized segments are shown in bold.

Results

Library Design. To isolate artificial transmembrane proteins that transform cells but lack any naturally occurring amino acid sequences, we constructed a retroviral library expressing many 29-aa traptamers, each consisting of methionine–alanine (encoded by a Kozak translation start site), a randomized transmembrane domain, and tyrosine–tryptophan to anchor the protein in the membrane (Fig. 1*B*). To randomize the transmembrane domain of the traptamers, we synthesized a degenerate oligonucleotide that encoded a randomized 25-residue segment, in which ~82% of amino acids are hydrophobic, an amino acid composition matching that of naturally occurring transmembrane domains (26). The degenerate oligonucleotide was converted into double-stranded DNA, amplified, and cloned into the pMSCV-puro retroviral expression vector. The resulting mixture of plasmids was packaged into retroviral particles to generate the UDv3 traptamer expression library. Based on deep sequencing of a similar library, we estimate this library encodes several hundred thousand different traptamers, a tiny fraction of the $\sim 10^{32}$ unique sequences possible given the randomization scheme. To verify the structure of this library, we sequenced 20 randomly sampled clones. As designed, the majority of these clones encoded 29-aa proteins that retained the invariant amino acids flanking a 25-residue randomized hydrophobic segment (Fig. S1). None of these unselected clones transformed C127 cells.

Isolation of Traptamers That Transform Mouse and Human Fibroblasts. C127 mouse cells were infected with the library at a low multiplicity of infection (<0.5 infectious units per cell) and maintained at confluence for 14 d. Transformed foci appeared at a low frequency in cells infected with the UDv3 library but not in cells infected with the empty vector. The infrequent appearance of foci implied that the vast majority of clones in the library did not have transforming activity. Individual transformed foci were expanded to establish transformed cell lines. Retroviral inserts were recovered by PCR from genomic DNA harvested from these cell lines and reintroduced individually into naïve C127 cells to test their transforming activity. Four sequences recovered from different foci (6A-1, 9C-3, 12A-5, and 3A-2; Fig. 1*C*) induced focus formation, some at levels exceeding that of E5 (Fig. 2*A*; and Fig. 2*B*, dark gray bars). In addition, cell lines stably expressing each of these proteins appeared transformed, as evidenced by the appearance of piled-up, disorganized, and refractile cells, in marked contrast to cells expressing empty vector, which formed a monolayer and remained flat and organized at confluence (Fig. S2).

The sequences of the transmembrane domains of the four active traptamers were not homologous to each other or the E5 protein (Fig. 1). The active traptamers were strongly hydrophobic, but each of them contained one or two hydrophilic residues. BLASTp searches identified a number of naturally occurring proteins that shared short, interrupted stretches of amino acids with the traptamers, as expected for these simple sequences, but did not reveal any convincing homologs in the >26 million member nonredundant translated protein database or in the >6 million entries in the metagenomics sequence database (Table S1).

We noted that the traptamer 9C-3 contains a glutamic acid at position 17. Interestingly, at this position, BPV E5 contains a glutamine that modules its activity. To test whether Glu17 in 9C-3 played a role in transformation, we mutated it to other hydrophilic amino acids and tested the transforming activity of the mutants. As shown in Fig. 2*C*, the focus forming activity of 9C-3 in C127 cells was eliminated by mutation to aspartic acid but stimulated by mutation to serine or threonine. A glutamic acid to glutamine substitution had little effect on transforming activity.

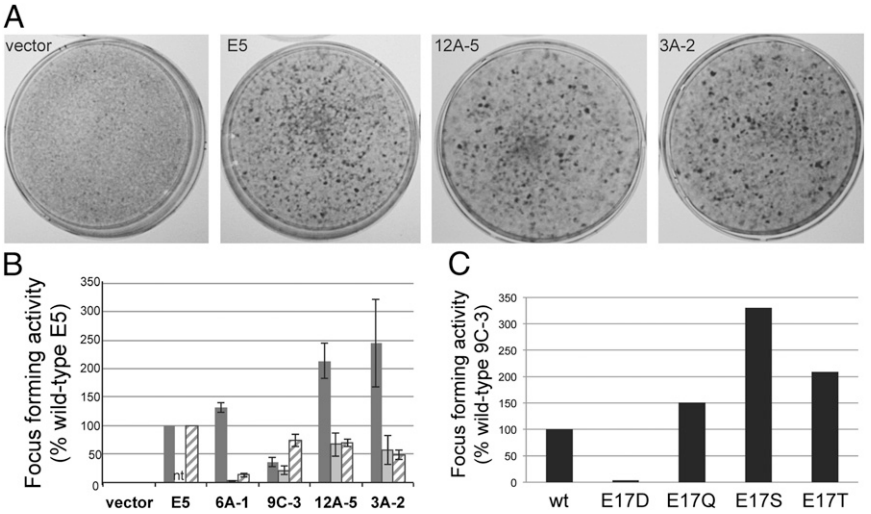

Fig. 2. Traptamers transform C127 mouse cells and HFFs. (*A*) C127 cells were infected with empty retroviral vector or a virus expressing E5 or the indicated traptamer and incubated for 14 d. Stained plates for the controls and two representative traptamers are shown. (*B*) Percentage of focus-forming activity relative to E5 in C127 cells (dark and light gray bars) and HFFs (cross-hatched bars), normalized for viral titer. Wild-type traptamers are shown with dark gray and cross-hatched bars and FF mutants with light gray bars. Results shown are the average of multiple independent experiments, with SE. The E5 FF mutant was not tested (nt), so the activity of the FF mutants was normalized to wild-type E5. (*C*) Focus formation by 9C-3 mutants. C127 cells were infected with retroviruses expressing wild-type 9C-3 or 9C-3 containing substitutions at position 17 (all containing an N-terminal FLAG tag). The graph is representative of two independent experiments and shows the percentage of focus-formation activity normalized to virus titer, relative to wild-type tagged 9C-3.

To determine whether the traptamers also transformed human cells, they were expressed individually in primary mortal, diploid, human foreskin fibroblasts (HFFs), which were then incubated at confluence for 21 d. As seen in Fig. 2*B*, cross-hatched bars, and Fig. S3*A*, three of the four traptamers induced numerous foci in these cells, although they were less active than E5 and the relative transforming activity of the various traptamers differed in HFFs compared C127 cells. Notably, 6A-1 induced few foci in HFFs. HFFs stably expressing 12A-5, 3A-2, and 9C-3 exhibited a transformed morphology, whereas cells expressing 6A-1 displayed an intermediate morphology (Fig. S3*B*).

The UDv3 library was designed to encode proteins, like E5, which contained a C-terminal tyrosine–tryptophan sequence. To determine whether traptamers lacking this sequence retained transforming activity, we replaced tyrosine–tryptophan with two phenylalanines and assessed the ability of these mutants to induce transformed foci in C127 cells. As seen in Fig. 2*B*, light gray bars, the 9C-3, 12A-5, and 3A-2 "FF" mutants induced focus formation, albeit at reduced levels compared with the original traptamers, whereas 6A-1.FF was inactive. Thus, three of the four active traptamers do not require any sequences derived from E5 for transforming activity.

Traptamers Activate the PDGF β Receptor. Tyrosine phosphorylation of the PDGF β receptor was assessed to determine whether the traptamers activated the PDGF β receptor. The PDGF β receptor was immunoprecipitated from detergent lysates of C127 cells and HFFs expressing empty vector, E5, or an active traptamer. The immunoprecipitates were electrophoresed and immunoblotted with an antibody specific for the PDGF β receptor or for phosphotyrosine. As shown in Fig. 3*A*, the levels of the mature and the more rapidly migrating, intracellular precursor forms of the PDGF β receptor were similar in all cell lines tested. As expected, there was little tyrosine phosphorylation of PDGF β receptor in cells expressing empty vector, but the mature (and, to a lesser extent, the precursor) form of the PDGF β receptor was phosphorylated in response to E5. All four of the traptamers also induced abundant tyrosine phosphorylation of the PDGF β receptor in C127 cells (Fig. 3*A*, left blots); 6A-1, 12A-5, and 3A-2 caused phosphorylation of the mature form, and 9C-3 induced substantial phosphorylation of both receptor forms. In HFFs, 12A-5 caused robust phosphorylation of both receptor forms, and 9C-3 and 3A-2 caused phosphorylation of primarily the mature form (Fig. 3*A*, right blots); 6A-1 induced only a low level of tyrosine phosphorylation of the PDGF β receptor in HFFs, consistent with its low transforming activity in these cells. Tyrosine phosphorylation of the PDGF β receptor was similar in C127 cells expressing 12A-5, 9C-3, and 3A-1 and their cognate FF mutants (Fig. 3*B*). Phosphorylation induced by 6A-1.FF was markedly reduced compared with that induced by wild-type 6A-1, again consistent with its low transforming activity.

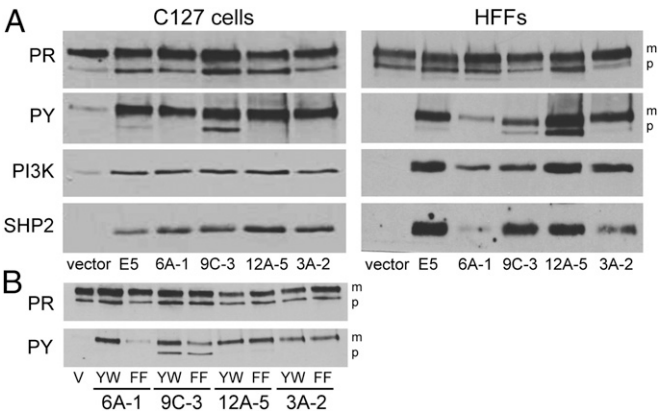

Fig. 3. Traptamers activate the PDGF β receptor. (*A*) Detergent extracts were prepared from C127 cells (left blots) and HFFs (right blots) expressing empty vector, E5, or the indicated traptamer. Extracts were immunoprecipitated with anti-PDGF β receptor antibody and immunoblotted with the indicated antibodies (PR, PDGF β receptor; PY, phosphotyrosine) to detect expression and tyrosine phosphorylation of the PDGF β receptor, respectively, or with antibody recognizing PI3K and SHP2 to assess association of these substrates with the receptor. (*B*) C127 cells expressing vector (V) or the indicated wild-type (YW) or mutant (FF) traptamers were analyzed as above for PDGF β receptor expression and tyrosine phosphorylation. The mature (m) and precursor (p) forms of the PDGF β receptor are indicated.

To determine whether the active traptamers recruited signaling effectors to the PDGF β receptor in transformed cells, PDGF receptor immunoprecipitates were immunoblotted with antibodies that recognized phosphoinositide 3′-kinase (PI3K) or SHP2. As expected, little or no association was detected between these effectors and the inactive receptor in cells expressing empty vector (Fig. 3*A*). In contrast, the PDGF β receptor formed a stable complex with both PI3K and SHP2 in transformed C127 cells and HFFs expressing E5 or the active traptamers. In accordance with the weak transforming activity of 6A-1 in HFFs, this traptamer recruited minimal SHP2 to the PDGF receptor in these cells. These results provided biochemical evidence that transformation-competent traptamers activated the PDGF β receptor in C127 cells and HFFs.

Traptamers Form a Stable Complex with the PDGF β Receptor. The E5 protein forms a stable complex with the PDGF β receptor, resulting in receptor activation. To determine whether the traptamers interacted with the PDGF β receptor, we added a FLAG epitope tag to the N terminus of the traptamers. The resulting FLAG-tagged constructs transformed C127 cells and induced tyrosine phosphorylation of the PDGF β receptor (Fig. 4, top blots). Detergent lysates of C127 cells expressing a FLAG-tagged traptamer (or the E5 protein or untagged 9C-3 as controls) were immunoprecipitated with an anti-FLAG antibody, electrophoresed, and immunoblotted with an anti-FLAG or an anti-PDGF β receptor antibody. As shown in Fig. 4, bottom blots, the FLAG-tagged proteins were expressed in C127 cells. We note that a small amount of 9C-3 migrates more slowly than the major species, which may represent a dimeric form stable in SDS. Notably, both the mature and precursor forms of the PDGF β receptor were coimmunoprecipitated from cells expressing the FLAG-tagged traptamers, but the receptor was not present in the anti-FLAG immunoprecipitates from cells expressing empty vector, E5, or untagged 9C-3. Reproducibly more PDGF β receptor was coimmunoprecipitated with 9C-3 than with the other traptamers and less with 6A-1. These results indicated that the traptamers stably interact with the PDGF β receptor in transformed cells.

Transformation Results from Ligand-Independent Activation of the PDGF β Receptor. To assess whether the biological activity of the traptamers required the PDGF β receptor, they were coexpressed with exogenous murine PDGF β receptor in BaF3 cells, a mouse hematopoietic cell line that lacks endogenous PDGF β receptor and requires IL-3 for proliferation. IL-3 dependence can be abrogated in these cells by coexpression of PDGF β receptor and proteins that activate the receptor, thus providing an alternative functional assay for activation of the receptor (27). Empty MSCV-puro vector, E5, v-*sis* (the viral homolog of PDGF-BB), or a traptamer was expressed in BaF3 cells stably expressing either empty LXSN vector or an exogenous mouse PDGF β receptor (BaF3/mPR cells). After selection for puromycin resistance, the ability of the cells to proliferate in the absence of IL-3 was assessed. As expected, E5 and v-*sis* induced IL-3 independent proliferation in cells expressing PDGF β receptor (Fig. 5*B*) but not in cells lacking receptor expression (Fig. 5*A*). Similarly, 9C-3, 12A-5, and 3A-2 induced IL-3 independent proliferation only in BaF3 cells expressing the PDGF β receptor; 6A-1 did not induce growth factor independence in either cell line. These results demonstrated that 9C-3, 12A-5, and 3A-2 induce IL-3–independent proliferation in BaF3 cells and that the PDGF β receptor is required for this activity.

To assess whether the tyrosine kinase activity of the PDGF β receptor was required for the focus forming activity of 6A-1, which was inactive in BaF3/mPR cells, we added a selective PDGF receptor kinase inhibitor (AG1296) to the medium of subconfluent C127 cells expressing empty vector, E5, 6A-1, or 12A-5. We previously showed that AG1296 and the similar inhibitor, AG1295, caused morphologic reversion of cells expressing the E5 protein or v-*sis* but not of cells expressing activated Neu, an oncogenic receptor tyrosine kinase unrelated to the PDGF β receptor (28, 29). In the absence of AG1296, cells expressing E5 or a traptamer displayed a transformed phenotype. AG1296 caused the cell lines transformed by E5 or a traptamer to revert to a nontransformed, flat morphology (Fig. S4), indicating that kinase activity of the PDGF β receptor is required for the transforming activity of 12A-5 and 6A-1 in C127 cells.

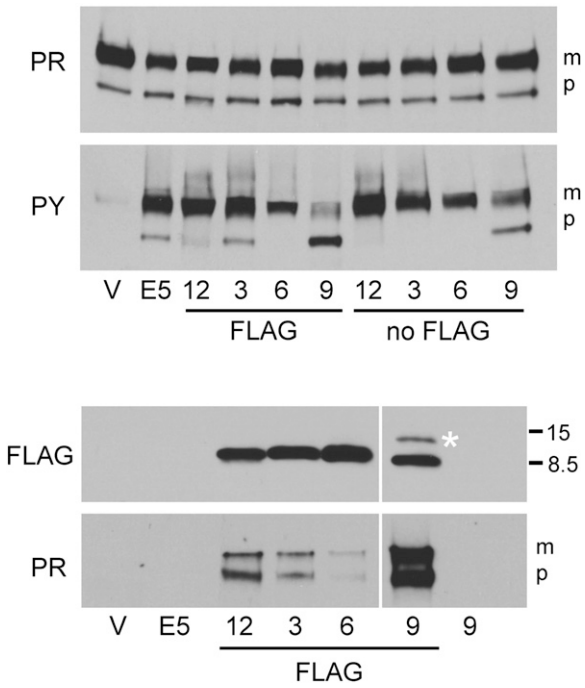

Fig. 4. Traptamers are expressed in C127 cells and form a stable complex with the PDGF β receptor. Detergent extracts were prepared from C127 cells expressing empty vector (V), E5, or an untagged or a FLAG-tagged traptamer, as indicated (12, 12A-5; 3, 3A-2; 6, 6A-1; 9, 9C-3). In the upper two blots, extracts were immunoprecipitated with anti-PDGF β receptor antibody and immunoblotted with the same antibody (PR) or with a phosphotyrosine (PY) antibody. In the lower four blots, extracts were immunoprecipitated with an anti-FLAG antibody and immunoblotted with either the FLAG antibody to detect expression of the tagged traptamers or the PDGF β receptor antibody to detect PDGF β receptor in complex with the traptamer. The lanes in the bottom blots were all from a single gel and exposed the same amount of time; an empty lane was removed as indicated by the space between the blots. A slower migrating band marked by an asterisk in the 9C-3 lane blotted with the FLAG antibody may represent a dimeric form of this traptamer. The mature (m) and precursor (p) forms of the mPDGF β receptor are indicated.

E5 induces cell transformation by interacting with the transmembrane domain of the PDGF β receptor and activating the receptor in a ligand-independent manner. To determine whether the traptamers acted similarly, we used two PDGF β receptor mutants, designated βαβ and TPR, in BaF3 cells. The βαβ chimeric receptor retains the extracellular ligand-binding domain and the intracellular signaling domain of the PDGF β receptor, but the PDGF β receptor transmembrane domain is precisely replaced with the transmembrane domain of the closely related PDGF α receptor (24). Therefore, v-*sis*, which binds to the extracellular domain of the PDGF receptor, induced growth factor independence in BaF3 cells expressing βαβ, whereas the E5 protein was inactive in these cells because it does not recognize the PDGF α receptor transmembrane domain contained in this chimera (Fig. 5*C*). Similarly, none of the traptamers induced

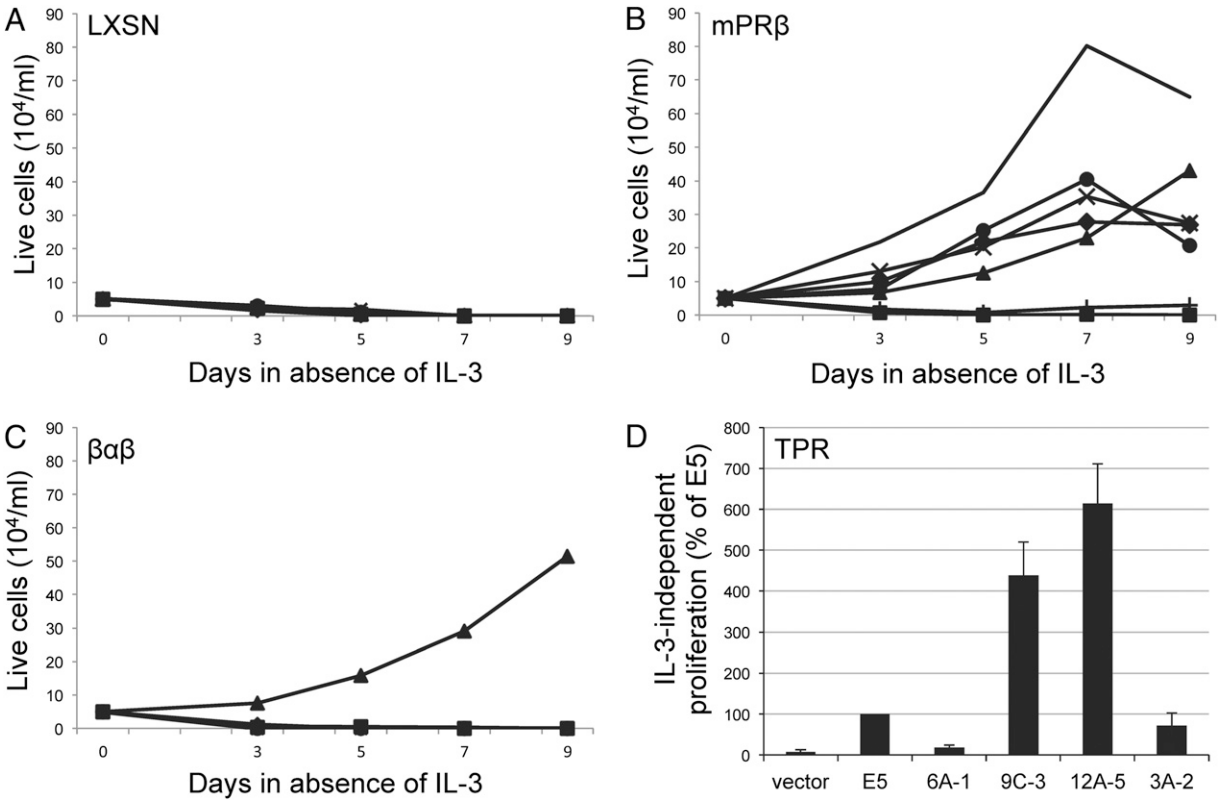

Fig. 5. Genetic requirements for traptamer activity. BaF3 cells expressing empty vector (LXSN) (*A*), the mouse PDGF β receptor (mPRβ) (*B*), or the chimeric βαβ receptor (*C*) were infected with empty retroviral vector (square) or retroviruses expressing E5 (no symbol), the *v-sis* oncogene (triangle), 6A-1 (+), 9C-3 (X), 12A-5 (circle), or 3A-2 (diamond). After selection for puromycin resistance, cells were incubated in medium lacking IL-3, and live cells were counted on the indicated days. Results shown are representative of multiple independent experiments. (*D*) BaF3 cells expressing the truncated PDGF β receptor (TPR) were infected with empty retroviral vector or retrovirus expressing E5 or the indicated traptamer. After selection for puromycin resistance, cells were incubated in the absence of IL-3 for between 5 and 7 d in different experiments, and live cells were counted. The graph shows results from three independent experiments, expressed as the average percentage of live cells relative to the number of live cells in E5-expressing cultures, with SEM shown.

growth factor independence in cells expressing the chimeric receptor, demonstrating that 12A-5, 3A-2, and 9C-3 required the transmembrane domain of the PDGF β receptor for biological activity.

The TPR receptor mutant lacks the extracellular ligand-binding domain of the PDGF β receptor but retains the transmembrane and intracellular domains of the PDGF β receptor. Therefore, it does not support growth factor independence in response to v-*sis* but does cooperate with the E5 protein (Fig. 5*D*) (27). Strikingly, 9C-3 and 12A-5 induced substantially greater proliferation when expressed with TPR than did the E5 protein; 3A-2 also conferred grow factor independence, but it was slightly less active than the E5 protein; and 6A-1 was inactive. Thus, the transforming activity of the traptamers in BaF3 cells does not require the extracellular domain of the PDGF β receptor and is therefore ligand-independent.

C127 Cells Expressing Traptamers Are Tumorigenic. To determine whether the traptamers caused tumor formation in animals, C127 cells expressing empty vector, E5, 9C-3, 12A-5, or 3A-2 were injected s.c. into immunodeficient nude mice (10^7 cells per injection site). As shown in Fig. 6*A*, C127 cells expressing vector alone did not form tumors, whereas cells transformed by the E5 protein were strongly tumorigenic. Cells expressing 3A-2 and 12A-5 also formed progressively growing tumors at all injection sites, whereas cells expressing 9C-3 formed tumors that did not grow in size (Fig. 6*B*). The tumors induced by 12A-5 grew more rapidly and to a larger size than the tumors caused by E5 itself, an effect that was statistically significant (12A-5 vs. E5-induced tumors, $P = 0.004$).

Tumors were excised 5 wk after inoculation and subjected to histologic examination. As shown in Fig. 6*C*, the tumors contained small, densely packed cells with a fibroblastic appearance. The tumors caused by the E5 protein or any of the traptamers were morphologically indistinguishable, including invasion of tumor cells into the adjacent muscle (an example of muscle invasion is shown in the right half of the 12A-5 tumor).

Because the traptamers used in the tumorigenicity studies lacked an epitope tag, we were not able to directly assess traptamer expression in the tumors. To test whether the PDGF β receptor was constitutively activated in the tumor cells, extracts were prepared from the tumors induced by 12A-5, 9C-3, and 3A-2 and analyzed by immunoprecipitation with anti-PDGF receptor antibody, followed by Western blotting for phosphotyrosine. All tumors displayed PDGF receptor tyrosine phosphorylation, with a higher signal from the 12A-5 tumors, correlating with their more rapid growth (Fig. 6*D*). Thus, the traptamers also activated the PDGF β receptor in vivo.

Discussion

To isolate biologically active proteins lacking naturally occurring amino acid sequences, we took advantage of several properties of transmembrane domains, in particular their small size, limited chemical diversity, stable conformation in membranes regardless of their amino acid sequence, and ability to engage in specific interactions with cellular transmembrane proteins to confer a selectable phenotype. With these factors in mind, we constructed

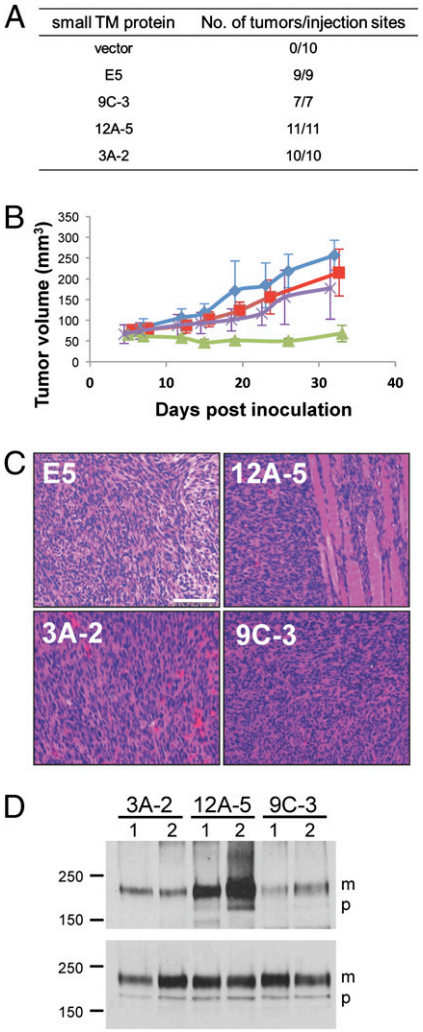

Fig. 6. C127 cells expressing traptamers are tumorigenic. (*A*) C127 cells expressing the empty vector, E5, or the indicated traptamer were injected into nude mice, and tumor formation was scored after 4 wk. (*B*) C127 cells expressing E5 (purple X), 9C-3 (green triangle), 12A-5 (blue diamond), or 3A-2 (red square) were injected s.c. into nude mice (two sites per mouse). At various times, palpable masses were measured in two dimensions, and tumor volume was calculated. The average tumor volume per mouse was used to calculate the group mean and SD, which are shown. (*C*) Tumors were excised, imbedded in paraffin, sectioned, and stained with hematoxylin/eosin. (Scale bar: 50 μm.) (*D*) RIPA extracts prepared from tumors induced by 12A-5, 9C-3, or 3A-2 were immunoprecipitated with antibody recognizing the PDGF β receptor and immunoblotted for phosphotyrosine (*Upper*) or PDGF β receptor (*Lower*). The same amount of extracted protein was immunoprecipitated for each lane. Molecular mass markers (in kilodaltons) and the mature (m) and precursor (p) forms of the mPDGF β receptor are indicated.

a retroviral library that expressed hundreds of thousands of short, randomized transmembrane proteins and selected proteins that induced oncogenic transformation of mouse cells. These proteins are not homologous to each other, the E5 protein, or any other known protein, and they act by binding and activating the PDGF β receptor in a ligand-independent fashion. Thus, de novo selection of small transmembrane proteins can be used to isolate totally novel proteins that induce complex phenotypes in mammalian cells.

Our results provide insight into various aspects of cell biology and transmembrane domain recognition. All of the traptamers selected for focus formation in C127 cells activated the PDGF β receptor. Furthermore, their ability to induce morphologic transformation in C127 cells required PDGF receptor activation (Fig. S4), as does their ability to induce growth factor independence in BaF3 cells (Fig. 5). In addition, PDGF receptor activation persists in the tumor cells, implying that it is important in vivo as well. These results imply that the PDGF β receptor is the primary cellular transmembrane protein able to induce focus formation in C127 cells, although it is possible that other activities ascribed to the E5 protein, such as its association with the vacuolar ATPase (30), may play a role. Importantly, the diverse amino acid sequences of the active traptamers demonstrate that transmembrane domains with no obvious sequence similarity can have similar biological activity.

Despite their shared mechanism of action, the traptamers display cell-type specificity. For example, 9C-3 is more active in HFFs than in C127 cells, whereas 6A-1 is inactive in HFFs and BaF3 cells; 6A-1 is also markedly inhibited by the FF substitution, suggesting that this traptamer is particularly sensitive to sequence perturbations and cellular context. These differences may reflect differential signaling outputs from the activated PDGF β receptor, which may arise from differences in the signaling pathways available in various cell types, the intracellular localization of the traptamers or the activated receptor, the sites of receptor tyrosine phosphorylation, or the structures of the activated PDGF β receptor dimer. Any of these factors may be influenced by the cell type and the particular traptamer under study. Thus, further analysis of these proteins is likely to reveal new aspects of PDGF receptor signaling. Furthermore, the traptamers may engage the PDGF β receptor in subtly different ways. For example, although the E5 protein acts as a disulfide-linked dimer, three of the four active traptamers lack cysteine residues and therefore are monomers or noncovalent oligomers. Structure–function analysis of traptamers and their effect on the PDGF β receptor will be facilitated by their small size and simple amino acid composition.

Transmembrane domains display less amino acid side-chain diversity than most proteins because they are short and hydrophobic. None of the traptamers isolated here contains more than 10 different amino acids; 3A-2.FF contains only 7 and lacks charged amino acids, strongly polar groups, or the ability to form disulfide bonds. Despite their limited chemical diversity, these traptamers specifically and productively engage the PDGF β receptor transmembrane domain. Experiments with libraries with even more restricted coding potential should define the minimal chemical diversity required to generate biological activity and specificity.

Although the traptamers are not homologous to BPV E5, there are some common sequence elements. For example, 12A-5 and E5 both have the sequence LLLLF. This presumably merely reflects the hydrophobic composition of the traptamers and the BPV E5 protein itself, because this sequence is not required for E5-mediated transformation, and it varies among the E5 proteins from different fibropapillomaviruses (31, 32). A more significant similarity is the glutamic acid at position 17 of 9C-3. The E5 protein has a glutamine at this position, which stabilizes complex formation with the PDGF β receptor and facilitates E5 homodimerization (28). Similarly, Glu17 in 9C-3 may account for the ability of this traptamer to coimmunoprecipitate the PDGF β receptor more efficiently than the other traptamers and for its apparent propensity to homodimerize noncovalently (Fig. 4 lower blots). As is the case for the BPV E5 protein (28, 33), hydrophilic substitutions at position 17 in 9C-3 stimulate or inhibit transforming activity. Other small oncoproteins with randomized transmembrane domains flanked by various amounts of E5 sequence also contain an essential hydrophilic amino acid at or near this position (21, 34). These observations imply that there are relatively few structural motifs that can support PDGF β receptor activation by small transmembrane proteins. However, we estimate that 1 in 20,000 clones in the UDv3 traptamer

library has transforming activity. Because the randomization scheme we used could theoretically generate an immense number of different sequences, this suggests that a very large number of transmembrane sequences can productively engage the PDGF β receptor.

Unlike most other protein-engineering approaches, de novo selection does not require complicated library synthesis or enrichment schemes, numerous rounds of selection, sophisticated computational expertise, high-resolution structural information, or detailed structure–function or chemical understanding of model proteins. Importantly, the sequence of traptamers is not constrained by the repertoire of sequences that survived natural selection or by the complex folds adopted by existing globular proteins. Although the utility of traptamers is presumably restricted to transmembrane protein targets, most biological processes are likely to be susceptible to modulation by traptamers because up to 30% of all eukaryotic cell proteins span membranes, and many biological pathways include transmembrane protein components (26). Although we previously isolated traptamers that modulated the PDGF β receptor, the erythropoietin receptor or CCR5 (19, 23), our earlier libraries encoded traptamers containing various segments of the E5 protein. These E5 sequences may impose restrictions on the structure of the traptamers encoded by these libraries, thereby preventing them from recognizing some potential targets. By eliminating all preexisting sequences from the UDv3 library, we diversified the structures encoded by this library, thereby expanding the pool of possible targets to many, if not all, transmembrane proteins.

The work reported here is similar in some regards to the study by Keefe and Szostak, who carried out 18 rounds of serial in vitro selection to isolate ATP-binding proteins from a library in which 80 aa were totally randomized (8). In contrast, we used a single round of biological selection, which not only identified active proteins but also discarded toxic ones. Strikingly, Keefe and Szostak estimated that only ∼1 in 10^{11} random-sequence proteins had ATP-binding activity, compared with the ∼1 in 20,000 clones in the traptamer library that induced focus formation. Thus, active proteins were more than one million-fold more abundant in our library. We believe that this difference is primarily attributable to the stable folded conformation most of the traptamers adopt in the membrane. In addition, it may be simpler to generate a protein able to bind productively to a large target protein than to form a 3D fold able to bind to a small molecule. Fisher et al. also reported a relatively high frequency of active proteins in a library encoding proteins engineered to adopt a stable globular fold (1). However, these proteins were much less active in bacteria than the proteins they replaced, possibly because of the sequence constraints imposed by this library design.

Proteins isolated by de novo selection may have structures and properties that do not occur in existing cellular proteins. These proteins may use novel mechanisms to modulate the activity of cell surface proteins, which constitute a large fraction of drug targets, and traptamers themselves may generate new activities, such as channels and pumps with novel specificity. Studies of such artificial proteins may suggest new therapeutic strategies or provide new classes of hydrophobic lead compounds for further development.

Our experiments may also model certain aspects of virus and cell evolution. Random nucleotide sequences can be incorporated on rare occasion into viral genomes, attributable, for example, to replication errors. These sequences will become fixed if their protein products provide selective advantage to the virus, in a process similar to de novo selection. This mechanism may explain why many viruses express short, transmembrane proteins that are not homologous to cellular proteins. Such proteins include the short hydrophobic (SH) protein of respiratory syncytial virus, the M2 ion channel of influenza virus, I5L and O3 of vaccinia virus, the p7 viroporin of swine fever virus, and the E5 protein of the fibropapillomaviruses (17, 35–39). Finally, we note that the properties of transmembrane domains that make them ideal for de novo selection—namely small size, simple chemistry, and a stable fold in the proper environment—may have also allowed them to arise early during biological evolution.

Materials and Methods

Retroviral Library Construction. To construct the UDv3 library expressing small, randomized, predominantly hydrophobic proteins, a long 5′ degenerate oligonucleotide that encoded the small hydrophobic proteins was first synthesized. This oligonucleotide consisted of a fixed 5′ end with an XhoI site and a Kozak consensus translation start sequence (ACCATGGG, encoding methionine–alanine), followed by 25 randomized codons, a 3′ fixed end encoding tyrosine and tryptophan, and two stop codons. A randomization scheme was used that encoded amino acids at frequencies similar to those found in naturally occurring transmembrane domains (26). In this scheme, the ratio of A:C:G:T was 1:1:1:0.5 at the first position of each randomized codon; 0.1:0.25:0.1:1 at the second position; and 0:1:0.1:0 at the third position. A 3′ nondegenerate oligonucleotide containing an EcoRI site was annealed to the 3′ fixed end of the degenerate oligonucleotide and extended to form double-stranded DNA. After PCR amplification with short primers, the product was digested with XhoI and EcoRI and ligated into a pMSCV-puro retroviral vector. Electroporation was used to transform *Escherichia coli* strain DH10β with the purified ligation mixture. Approximately 1.6 million ampicillin-resistant bacterial colonies were pooled, and plasmid DNA was harvested to generate the UDv3 library. To confirm the amino acid composition and structure of clones in this library, DNA from randomly picked ampicillin-resistant colonies was sequenced. Details of library construction are presented in *SI Appendix, Supplemental Methods*.

Mutagenesis and Cloning. The QuikChange site-directed mutagenesis protocol (Agilent Technologies) was used to construct FF mutants by substituting the C-terminal tyrosine and tryptophan residues of the active traptamers with two phenylalanines. A PCR-based protocol was used to insert an in-frame N-terminal FLAG tag followed by a di-glycine linker (DYKDDDDKGG) between the initiating methionine and the invariant alanine at the second position of 12A-5, 6A-1, 3A-2, and 9C-3. Briefly, PCR was performed using a long forward primer encoding the tag sequence at the appropriate position, and the resulting PCR product was digested and inserted into pMSCV-puro. QuikChange mutagenesis was also used to construct mutations at position 17 of FLAG-tagged 9C-3.

Cells and Viruses. C127 cells were maintained in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% (vol/vol) FBS, 2 mM L-glutamine, 20 mM Hepes, and antibiotics (penicillin–streptomycin and amphotericin B) (DMEM-10). Early-passage HFFs were obtained from the Yale Skin Diseases Research Center and maintained in minimal essential alpha medium supplemented with 10% (vol/vol) FBS and antibiotics (MEM alpha-10). BaF3 cells were maintained in RPMI medium 1640 supplemented with 10% (vol/vol) heat-inactivated FBS (HI-FBS), 6.7% (vol/vol) WEHI-3B cell-conditioned medium (as an IL-3 source), 0.05 mM β-mercaptoethanol (β-ME), and antibiotics (RPMI/IL-3); 293T cells were maintained in DMEM-10 or in DMEM supplemented with 5% (vol/vol) FBS, 5% (vol/vol) bovine calf serum, 1 mM L-glutamine, 20 mM Hepes, and antibiotics.

Retroviral stocks were prepared by calcium phosphate-mediated cotransfection of 293T cells with a retroviral plasmid or library encoding the gene(s) of interest and the packaging plasmids pCL-Eco and pVSVg. Retroviral supernatants were harvested 48 h posttransfection and filtered through 0.45-μm filters. Virus stocks were used immediately or stored at −80 °C for later use.

Focus-Formation Assay and Recovery of Transforming Clones. To screen the UDv3 library for focus formation, ten 60-mm dishes of C127 cells at 70% confluence were infected with 3 mL of UDv3 virus stock supplemented with 4 μg/mL polybrene (∼850,000 infectious units in total). After 24 h, each dish of cells was trypsinized and replated into two dishes. Cells were incubated at confluence for 14 d with biweekly medium changes to allow focus formation.

Transformed foci were isolated using cloning cylinders, and cells were expanded over five rounds to generate morphologically transformed cell lines. Genomic DNA was isolated from transformed cells using a DNeasy kit (Qiagen), and retroviral inserts were amplified by PCR from the genomic DNA using primers recognizing the invariant portions of the library (24). Amplified

inserts were subcloned into pMSCV-puro, which was then used to transform *E. coli* strain DH5α by electroporation. Plasmid DNA was isolated from randomly chosen colonies, sequenced, and used to generate retrovirus in 293T cells. To identify individual clones with transforming activity, C127 cells were infected with 1 mL of unconcentrated retrovirus in the presence of polybrene. After 24 h, the infected cells were split, and 48 h postinfection, puromycin was added. Morphological transformation was assessed after 7 d. Clones with transforming activity were subjected to a quantitative focus forming assay by using 2–10 μL of virus to infect C127 cells or HFFs in 60-mm dishes as above. The cells were maintained in the absence of drug selection for approximately 2 wk to allow the outgrowth of discrete foci. The cells were then fixed in methanol and stained with a 5% (vol/vol) dilution of a modified Giemsa solution (Sigma-Aldrich) to visualize foci. The number of foci was normalized for virus titer, which was determined by plating dilutions of infected cells into 100-mm dishes and counting the puromycin-resistant colonies that developed.

To determine whether transformation required PDGF receptor activity, C127 cells expressing the empty vector, E5, 12A-5, or 6A-1 were plated at ∼85% confluence in six-well plates and incubated in DMEM-10 containing 20 μM AG1296 in DMSO or an equivalent volume of DMSO. Cells were photographed after 2 d.

IL-3 Independence Assay. The murine PDGF β receptor, the βαβ chimeric receptor (24), and a truncated PDGF β receptor lacking the extracellular domain (TPR) (27) were subcloned into the pLXSN retroviral vector, which harbors the G418 resistance marker. To establish BaF3 cell lines stably expressing each receptor construct or control LXSN lacking a transgene, 2.5×10^6 BaF3 cells in 10 mL of RPMI/IL-3 were infected with 1–2 mL of the appropriate viral stock in the presence of 4 μg/mL polybrene. G418 was added to a final concentration of 1 mg/mL 48 h postinfection, and cells were incubated until mock-infected cells died. The resulting G418-resistant BaF3 cell lines were infected as above with retrovirus expressing the traptamers or E5 from the MSCV-puro vector, or *v-sis* from pBabepuro vector, and then selected in medium containing 1μg/mL puromycin. To assay for IL-3–independent proliferation, 5×10^5 puromycin-resistant cells were washed twice with PBS and then incubated in RPMI containing 1% HI-FBS, 0.05 mM β-ME, and antibiotics. Live cells were counted at various times thereafter.

Immunoprecipitation and Immunoblotting. C127 cells and their transformed derivatives were serum-starved overnight in 100-mm plates at 80% confluence, washed twice with PBS, and lysed in 1 mL of cold radioimmunoprecipitation assay (RIPA)-Mops buffer [20 mM morpholinepropanesulfonic acid (pH 7.0), 150 mM NaCl, 1% Nonidet P-40, 1 mM EDTA, 1% deoxycholic acid, 0.1% SDS] with inhibitors [1× Halt protease and phosphatase inhibitor mixture (Thermo Scientific), 1 mM phenylmethylsulfonylfluoride, and 0.5 mM sodium metavanadate].

To immunoprecipitate PDGF β receptor, anti-PDGF β receptor rabbit antiserum raised against the C-terminal 13 aa of the human PDGF β receptor was added to cell extracts at 5 μL/mg extracted protein. After overnight incubation at 4 °C, extracts were incubated with protein A-Sepharose beads (GE Healthcare) for 1–2 h at 4 °C. The beads were then pelleted, washed, and resuspended in 2× Laemmli sample buffer. Immunoprecipitates were boiled to elute immune complexes and electrophoresed for 1.5 h at 150 V on a 7.5% (wt/vol) polyacrylamide/SDS gel. Samples were transferred for 1.5 h at 100 V to nitrocellulose membranes in transfer buffer [25 mM Tris base, 192 mM glycine, and 20% (vol/vol) methanol] containing 0.1% SDS. Membranes were blocked in 5% (wt/vol) nonfat dry milk in TBST [10 mM Tris·HCl (pH 7.4), 167 mM NaCl, 1% Tween-20] for 1 h and incubated overnight at 4 °C with either anti-phosphotyrosine (PY-100; Cell Signaling), anti-PDGF β receptor, anti-PI3K (Millipore), or anti-SHP2 (BD Biosciences) antibody diluted 1:1,250, 1:250, 1:2,000, and 1:1,000, respectively, in 5% (wt/vol) milk/TBST. Blots were washed five times in TBST and subsequently incubated for 1 h at room temperature with horseradish peroxidase (HRP)-conjugated donkey anti-mouse secondary antibody or HRP–protein A diluted 1:10,000 or 1:8,000, respectively, in 5% (wt/vol) milk/TBST. Blots were then washed and visualized by enhanced chemiluminescence.

C127 cells stably expressing the N-terminally FLAG-tagged UDv3 constructs in 10-cm dishes at 90% confluence were washed twice with PBS and lysed in 0.5 mL per dish of lysis buffer [50 mM Tris·HCl (pH 7.4), 150 mM NaCl, 1 mM EDTA, 1% Triton-X-100] containing protease and phosphatase inhibitors. Extracts were precleared by incubating with protein A–Sepharose beads for 1 h at 4 °C, after which the beads were pelleted and discarded. To immunoprecipitate the FLAG-tagged proteins, 2.5 mg of precleared extract was incubated overnight at 4 °C with an anti-FLAG affinity matrix consisting of the M2 monoclonal antibody, which recognizes the FLAG epitope, conjugated to agarose beads (Sigma). The beads were then washed five times with TBS [50 mM Tris·HCl (pH 7.4), 150 mM NaCl] and incubated on ice with 70 μL of FLAG peptide (Sigma) at 200 ng/μL to elute the FLAG-tagged proteins from the beads. Eluates were mixed with 35 μL of 5× Laemmli sample buffer, boiled, and electrophoresed on a 7.5% or 4–20% (wt/vol) gradient TGX gel (Bio-Rad). Gels were then transferred to 0.2-μm PVDF membranes for 1 h in transfer buffer lacking SDS. Blots were blocked and probed as described above with a 1:250 dilution of the anti-PDGF receptor antiserum or 9 μg/mL of the anti-FLAG M2 monoclonal antibody.

Tumorigenicity Assay. Six-week-old female nude mice (Athymic NCr-nu/nu; National Cancer Institute) were maintained under standard conditions. C127 cells stably expressing empty vector, E5, or a UDv3 traptamer were harvested at 70% confluence by trypsinization, washed three times in PBS, and resuspended in PBS at a final density of 10^8 cells per milliliter. Approximately 10^7 cells (0.1 mL of the cell suspension) were injected s.c. into the left and right lateral flanks of each mouse. Mice were monitored biweekly, and tumor formation was measured in two dimensions using a caliper. Tumor volume was calculated using the formula $V = (\text{length} \times \text{width}^2) \times 0.5$. After 5 wk, mice were killed, and tumors were excised. Tissues were fixed in 4% (vol/vol) paraformaldehyde, embedded in paraffin, sectioned, stained with hematoxylin/eosin, and examined by bright-field microscopy. To prepare extracts from two tumors induced by each traptamer, excised tumors were frozen in liquid nitrogen, ground with a mortar and pestle, and lysed in RIPA buffer.

ACKNOWLEDGMENTS. We thank Xiaopan Yao and Fang-yong Li for statistical analysis, James Noonan for advice, and Jan Zulkeski for assistance in preparing the manuscript. This work was supported by a generous gift from Laurel Schwartz and by National Cancer Institute Grant CA037157.

1. Fisher MA, McKinley KL, Bradley LH, Viola SR, Hecht MH (2011) *De novo* designed proteins from a library of artificial sequences function in Escherichia coli and enable cell growth. *PLoS ONE* 6(1):e15364.
2. Binz HK, Plückthun A (2005) Engineered proteins as specific binding reagents. *Curr Opin Biotechnol* 16(4):459–469.
3. Hilvert D (2013) Design of protein catalysts. *Annu Rev Biochem* 82:447–470.
4. Kries H, Blomberg R, Hilvert D (2013) *De novo* enzymes by computational design. *Curr Opin Chem Biol* 17(2):221–228.
5. Yin H, et al. (2007) Computational design of peptides that target transmembrane helices. *Science* 315(5820):1817–1822.
6. Smith BA, Hecht MH (2011) Novel proteins: From fold to function. *Curr Opin Chem Biol* 15(3):421–426.
7. Urvoas A, Valerio-Lepiniec M, Minard P (2012) Artificial proteins from combinatorial approaches. *Trends Biotechnol* 30(10):512–520.
8. Keefe AD, Szostak JW (2001) Functional proteins from a random-sequence library. *Nature* 410(6829):715–718.
9. Ja WW, et al. (2007) Extension of Drosophila melanogaster life span with a GPCR peptide inhibitor. *Nat Chem Biol* 3(7):415–419.
10. Raffler NA, Schneider-Mergener J, Famulok M (2003) A novel class of small functional peptides that bind and inhibit human alpha-thrombin isolated by mRNA display. *Chem Biol* 10(1):69–79.
11. Fleishman SJ, et al. (2011) Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. *Science* 332(6031):816–821.
12. Jiang L, et al. (2008) *De novo* computational design of retro-aldol enzymes. *Science* 319(5868):1387–1391.
13. Röthlisberger D, et al. (2008) Kemp elimination catalysts by computational enzyme design. *Nature* 453(7192):190–195.
14. Seelig B, Szostak JW (2007) Selection and evolution of enzymes from a partially randomized non-catalytic scaffold. *Nature* 448(7155):828–831.
15. Yamauchi A, et al. (2002) Evolvability of random polypeptides through functional selection within a small library. *Protein Eng* 15(7):619–626.
16. Petti L, Nilson LA, DiMaio D (1991) Activation of the platelet-derived growth factor receptor by the bovine papillomavirus E5 transforming protein. *EMBO J* 10(4):845–855.
17. Talbert-Slagle K, DiMaio D (2009) The bovine papillomavirus E5 protein and the PDGF β receptor: It takes two to tango. *Virology* 384(2):345–351.
18. Edwards APB, Xie Y, Bowers L, DiMaio D (2013) Compensatory mutants of the bovine papillomavirus E5 protein and the platelet-derived growth factor β receptor reveal a complex direct transmembrane interaction. *J Virol* 87(20):10936–10945.
19. Cammett TJ, et al. (2010) Construction and genetic selection of small transmembrane proteins that activate the human erythropoietin receptor. *Proc Natl Acad Sci USA* 107(8):3447–3452.
20. Freeman-Cook LL, et al. (2004) Selection and characterization of small random transmembrane proteins that bind and activate the platelet-derived growth factor β receptor. *J Mol Biol* 338(5):907–920.

21. Freeman-Cook LL, et al. (2005) Specific locations of hydrophilic amino acids in constructed transmembrane ligands of the platelet-derived growth factor β receptor. *J Mol Biol* 345(4):907–921.
22. Ptacek JB, Edwards APB, Freeman-Cook LL, DiMaio D (2007) Packing contacts can mediate highly specific interactions between artificial transmembrane proteins and the PDGFbeta receptor. *Proc Natl Acad Sci USA* 104(29):11945–11950.
23. Scheideman EH, et al. (2012) Transmembrane protein aptamers that inhibit CCR5 expression and HIV coreceptor function. *J Virol* 86(19):10281–10292.
24. Talbert-Slagle K, et al. (2009) Artificial transmembrane oncoproteins smaller than the bovine papillomavirus E5 protein redefine the sequence requirements for activation of the platelet-derived growth factor β receptor. *J Virol* 83:9773–9785.
25. Senes A, Gerstein M, Engelman DM (2000) Statistical analysis of amino acid patterns in transmembrane helices: The GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J Mol Biol* 296(3):921–936.
26. Lehnert U, et al. (2004) Computational analysis of membrane proteins: Genomic occurrence, structure prediction and helix interactions. *Q Rev Biophys* 37(2):121–146.
27. Drummond-Barbosa DA, Vaillancourt RR, Kazlauskas A, DiMaio D (1995) Ligand-independent activation of the platelet-derived growth factor beta receptor: Requirements for bovine papillomavirus E5-induced mitogenic signaling. *Mol Cell Biol* 15(5): 2570–2581.
28. Klein O, et al. (1998) Role of glutamine 17 of the bovine papillomavirus E5 protein in platelet-derived growth factor beta receptor activation and cell transformation. *J Virol* 72(11):8921–8932.
29. Petti LM, Ray FA (2000) Transformation of mortal human fibroblasts and activation of a growth inhibitory pathway by the bovine papillomavirus E5 oncoprotein. *Cell Growth Differ* 11(7):395–408.
30. Goldstein DJ, Andresson T, Sparkowski JJ, Schlegel R (1992) The BPV-1 E5 protein, the 16 kDa membrane pore-forming protein and the PDGF receptor exist in a complex that is dependent on hydrophobic transmembrane interactions. *EMBO J* 11(13): 4851–4859.
31. Horwitz BH, Burkhardt AL, Schlegel R, DiMaio D (1988) 44-amino-acid E5 transforming protein of bovine papillomavirus requires a hydrophobic core and specific carboxyl-terminal amino acids. *Mol Cell Biol* 8(10):4071–4078.
32. Petti LM, Talbert-Slagle K, Hochstrasser ML, DiMaio D (2013) A single amino acid substitution converts a transmembrane protein activator of the platelet-derived growth factor β receptor into an inhibitor. *J Biol Chem* 288(38):27273–27286.
33. Sparkowski J, Anders J, Schlegel R (1994) Mutation of the bovine papillomavirus E5 oncoprotein at amino acid 17 generates both high- and low-transforming variants. *J Virol* 68(9):6120–6123.
34. Kulke R, Horwitz BH, Zibello T, DiMaio D (1992) The central hydrophobic domain of the bovine papillomavirus E5 transforming protein can be functionally replaced by many hydrophobic amino acid sequences containing a glutamine. *J Virol* 66(1): 505–511.
35. Gan SW, Ng L, Lin X, Gong X, Torres J (2008) Structure and ion channel activity of the human respiratory syncytial virus (hRSV) small hydrophobic protein transmembrane domain. *Protein Sci* 17(5):813–820.
36. Guo HC, et al. (2013) Viroporin activity and membrane topology of classic swine fever virus p7 protein. *Int J Biochem Cell Biol* 45(7):1186–1194.
37. Pinto LH, Lamb RA (2006) The M2 proton channels of influenza A and B viruses. *J Biol Chem* 281(14):8997–9000.
38. Satheshkumar PS, Chavre J, Moss B (2013) Role of the vaccinia virus O3 protein in cell entry can be fulfilled by its Sequence flexible transmembrane domain. *Virology* 444(1-2):148–157.
39. Sood CL, Ward JM, Moss B (2008) Vaccinia virus encodes I5, a small hydrophobic virion membrane protein that enhances replication and virulence in mice. *J Virol* 82(20): 10071–10078.