JVI Accepts, published online ahead of print on 15 July 2009
J. Virol. doi:10.1128/JVI.00946-09

Artificial transmembrane oncoproteins smaller than the bovine papillomavirus E5 protein redefine sequence requirements for activation of the platelet-derived growth factor β receptor

Kristina Talbert-Slagle[1], Sara Marlatt[2], Francisco N. Barrera[3], Ekta Khurana[3], Joanne E. Oates[5], Mark Gerstein[3], Donald M. Engelman[3], Ann Dixon[5], and Daniel DiMaio[2,3,4*]

[1]Department of Epidemiology and Public Health, P.O. Box 208034

[2]Department of Genetics, P.O. Box 208005

[3]Department of Molecular Biophysics & Biochemistry, P.O. Box 208114

[4]Department of Therapeutic Radiology, P.O. Box 208040

Yale University School of Medicine, New Haven, CT 06520 USA

[5] Department of Chemistry, University of Warwick

Room B600, Gibbet Hill Road, Coventry UK CV4 7AL

Running Title: Small transmembrane activators of the PDGF receptor

Keywords: transmembrane; oncogene; PDGF receptor; E5 protein; bovine papillomavirus

Word count: Abtract = 190; Text = 8,637

[*]Corresponding author: Tel: 203-785-2684; Fax: 203-785-6765; Email: daniel.dimaio@yale.edu

ABSTRACT

The bovine papillomavirus E5 protein (BPV E5) is a 44-amino acid homodimeric transmembrane protein that binds directly to the transmembrane domain of the PDGF β receptor and induces ligand-independent receptor activation. Three specific features of BPV E5 are considered important for its ability to activate the PDGF β receptor and transform mouse fibroblasts: a pair of C-terminal cysteines, a transmembrane glutamine, and a juxtamembrane aspartic acid. By using a new genetic technique to screen libraries expressing artificial transmembrane proteins for activators of the PDGF β receptor, we isolated much smaller proteins, from 32 to 36 residues, that lack all three of these features yet still dimerize non-covalently, specifically activate the PDGF β receptor via its transmembrane domain, and transform cells efficiently. The primary amino acid sequence of BPV E5 is virtually unrecognizable in some of these proteins, which share as few as seven consecutive amino acids with the viral protein. Thus, small artificial proteins that bear little resemblance to a viral oncoprotein can nevertheless productively interact with the same cellular target. We speculate that similar cellular proteins may exist but have been overlooked due to their small size and hydrophobicity.

INTRODUCTION

Viruses are tiny relative to the cells they infect. The largest animal viruses encode at most a few hundred genes, in contrast to the tens of thousands of genes expressed by the host cell. To overcome the constraints imposed by their small size, viruses use multiple or even overlapping reading frames and alternative splicing to produce more than one protein from single transcripts, build capsids from repeating subunits, harness cellular mechanisms to produce viral products, and express extremely small proteins. Several of these small viral proteins are membrane-anchored, including the 44-residue SH protein of parainfluenza virus 5, which blocks apoptosis of infected cells (16), and the 96-residue M2 protein of influenza A, which forms an ion channel in endosomal membranes of infected cells and is required for infectivity {reviewed in (38)}. One of the best-characterized viral small transmembrane proteins is the E5 protein encoded by bovine papillomavirus type 1 (BPV E5). BPV E5 contains only 44 amino acids and is thus essentially an isolated transmembrane domain (Fig. 1) {reviewed in (47)}. It is the primary translation product from a small open reading frame and is sufficient to cause tumorigenic cell transformation, making it the smallest known autonomous oncoprotein.

BPV E5 induces cell transformation by specifically activating a much larger cellular target, the PDGF β receptor tyrosine kinase (8, 14, 30, 35). The PDGF β receptor, a single-span transmembrane protein of more than 1000 amino acids, is normally activated by the binding of its dimeric ligand, PDGF, to the extracellular domain of the receptor. In contrast, a dimer of the BPV E5 protein binds to the transmembrane domain of two monomers of the PDGF β receptor (5, 8, 13, 34, 45), inducing ligand-independent receptor dimerization, trans-phosphorylation of tyrosine residues in the intracellular catalytic domain of the receptor, and sustained mitogenic signaling (24, 47). This interaction is highly specific; at normal expression levels, BPV E5 does

not bind or activate other receptor tyrosine kinases, not even the closely-related PDGF α receptor (14, 33).

Extensive characterization of BPV E5 showed that its ability to activate the PDGF β receptor requires only a few specific features, in addition to its overall hydrophobic composition: the ability to form covalently-linked homodimers mediated by two C-terminal cysteines, a transmembrane glutamine that is important both for the homodimerization of BPV E5 and for its interaction with the PDGF β receptor, and a juxtamembrane aspartic acid that provides a negative charge necessary for interacting with the PDGF β receptor (17, 20, 21, 43, 44, 46). These essential amino acids, gln17, asp33, cys37, and cys39, are perfectly conserved among the E5 proteins of the other papillomaviruses including the deer papillomavirus E5 protein, which also activates the PDGF β receptor and transforms cells (22). In contrast, the transmembrane sequences of these proteins are otherwise not well conserved. These findings imply that the fibropapillomavirus E5 proteins evolved from a common precursor and that various substitutions of the hydrophobic residues in these proteins have been well-tolerated over evolutionary time, as long as the four essential amino acids were preserved.

Mutational analysis of the BPV E5 protein supports these conclusions, demonstrating that it can tolerate significant changes to its primary amino acid sequence without loss of activity, as long as these essential amino acids are present. Frameshift mutations that introduced or removed residues from the N-terminal sequence of BPV E5 had little effect on its ability to induce cell transformation if the remainder of the protein was translated in-frame (6, 18). Removal of nine C-terminal residues (positions 35-44) of BPV E5 also did not abrogate its activity, as long as the juxtamembrane aspartic acid was preserved and a cysteine was introduced at the terminal position, presumably to allow covalent dimerization (29). In addition, many hydrophobic

substitutions were tolerated in the transmembrane domain, but almost all of the active mutants retained at least the transmembrane glutamine, the aspartic acid, and the C-terminal cysteines (17, 18, 29). Indeed, insertion of a transmembrane glutamine into otherwise inactive proteins with random transmembrane domains often restored transforming activity (23), and a heterologous dimerization domain could compensate for loss of the cysteines (28). By screening libraries that express thousands of small proteins with randomized transmembrane domains, we identified small transmembrane proteins that induced focus formation in murine C127 fibroblasts by activating the endogenous PDGF β receptor (10, 12, 39). These proteins were designed to retain the N-terminal and C-terminal segments of BPV E5, including the cysteines important for dimerization, and all of the active proteins shared at least 50% sequence identity with BPV E5 (11, 39).

In addition to inducing focus formation in C127 cells, BPV E5 can also induce PDGF β receptor-dependent proliferation of murine hematopoietic BaF3 cells, which ordinarily require the cytokine IL-3 for survival and proliferation. Parental BaF3 cells do not express the PDGF β receptor and do not respond to BPV E5, but the E5 protein can induce growth factor independence in BaF3 cells engineered to express the PDGF β receptor (8). Here, we used these cells to isolate a 32-residue protein with a novel transmembrane domain that specifically activated the exogenous PDGF β receptor in BaF3 cells, even though it lacked the transmembrane glutamine, the juxtamembrane aspartic acid, and twelve C-terminal residues, including the two cysteines. Furthermore, high-level expression of certain 36-amino acid proteins with randomized transmembrane domains and no cysteines was sufficient to induce efficient PDGF-receptor-dependent transformation of C127 and BaF3 cells, even though they retained only seven contiguous amino acids from BPV E5. Our results demonstrate that small

transmembrane proteins that differ almost completely from the BPV E5 protein can effectively activate the PDGF β receptor via transmembrane interactions and transform cells. Furthermore, the elements once thought to be essential for this interaction can, in fact, be removed. Thus, by studying artificial transmembrane proteins significantly shorter than the wild-type E5 protein, we have redefined sequence requirements for PDGF β receptor activation and cell transformation.

MATERIALS and METHODS

Library and retroviral clone construction. To construct the LFC5 library, we used a degenerate oligonucleotide with a fixed upstream sequence that encoded residues 9-13 of the wild-type BPV E5 sequence, including a SpeI restriction site, followed by a randomized sequence corresponding to residues 14-30 of the wild-type BPV E5 protein, and finally a nucleotide sequence encoding residues 31-38 of the E5 sequence, with position 33 also randomized. For the randomized codons, position one was an equal mix of A, C, T and G; position two was a mixture of T:A:C in the ratios 5:1:0.1; and position three was an equal mix of C and G. To this degenerate oligonucleotide, we annealed an oligonucleotide that was complementary to the fixed sequence encoding residues 34-38 and that also contained residues 39-44 of the BPV E5 protein as well as a downstream BamHI restriction site. After extension of these oligonucleotides, we amplified them by PCR with short, internal primers and digested the purified PCR products with SpeI and BamHI. The products were subcloned into a pRV-Hyg^R retroviral vector (40) downstream of residues 1-12 of the BPV E5 protein. The predicted composition of the randomized segment of this library is as follows (with percentages): L(30.7), V(20.5), I(10.2), M(10.2), F(10.2), K(2), N(2), stop(2), Y(2), Q(2), H(2), E(2), D(2), S(0.4), T(0.4), P(0.4), A(0.4). Further cloning details are available from the authors upon request (for primer sequences, see Table S1).

To construct the KTS1 library, we used a degenerate oligonucleotide that encoded residues 10 and 11 of the wild-type BPV E5 protein, which we had altered by silent mutation to encode an AvrII restriction site, followed by randomized codons corresponding to positions 12-30 of the wild-type BPV E5 protein, then fixed residues 31 and 32 of wild-type BPV E5, a stop codon at position 33, and a BamHI site for cloning. For randomized codons, position one was a

mix of A:C:G:T in the ratios 1:1:1:0.5, position two was a mixture of A:C:G:T in the ratios 0.1:0.25:0.1:1, and position three was a mix of C and G in the ratios 1:0.1. This composition allows all 20 amino acids at ratios approximating those present in natural transmembrane proteins (26). The predicted composition of the randomized segment of the KTS1 library is as follows: L(20.6), V(19.7), I(17.9), M(1.8), F(8.95), K(0.18), N(1.8), stop(0.09), Y(0.9), Q(0.18), H(1.8), E(0.18), D(1.8), S(1.8), T(4.9), P(4.9), A(4.9). As with the LFC5 library, we synthesized double-stranded products of these randomized nucleotide sequences, amplified them with short, internal primers, and then cloned the library into the pT2H-F13 retroviral plasmid. pT2H-F13 was derived from pRV-Hyg^R; it contains a Kozak consensus sequence in the fixed 5' sequence of the BPV E5 gene and an additional internal AvrII restriction site created by silent mutation. Further cloning details are available from the authors upon request (for primer sequences, see Table S1).

To generate cell lines expressing individual small transmembrane proteins, we cloned the genes into pRV-Hyg^R or pT2H-F13. We also utilized the high-expression retroviral plasmid pMSCVhygro (Clontech, Mountain View, CA, USA) to express some of the individual proteins. Genes encoding the chimeric receptors were cloned into an LXSN plasmid, which confers G418 resistance. Further details are available in Supplementary Information. Mutation of single residues in protein pTM32-1 was performed using the Quick Change Site-Directed Mutagenesis protocol (Stratagene, La Jolla, CA, USA) on pTM32-1 cloned in pRV-Hyg^R, and sequences were confirmed on both strands. See Table S1 for primer sequences.

Cell Lines, Retroviral Infections, and Tissue Culture. BaF3 cells and their derivatives were maintained in RPMI/IL-3 as described (7, 8), and C127 and HEK293T cells were

maintained in Dulbecco's Modified Eagle Medium containing 10% fetal bovine serum (FBS) (DMEM10) (10).

To generate cell lines expressing the murine PDGF β receptor or the βαβ or βKitβ chimeric receptors, we prepared retroviral particles by calcium phosphate cotransfection of HEK293T cells with the LXSN retroviral plasmid encoding the receptor of interest and the pCL-Eco and VSV-G packaging plasmids (Imgenex, San Diego, CA, USA). We harvested retroviral supernatants from these cells at 24, 48, and 72 hours post-transfection, pooled and filtered the retroviral supernatants through 0.45 micron filters (Millipore, Leiden, The Netherlands) and concentrated the retrovirus in Amicon Ultra high-speed centrifugal filtration concentrators (Millipore). We infected parental BaF3 cells with the concentrated retrovirus as follows: $5x10^5$ BaF3 cells in 500 µl of medium per well of a 12-well dish were infected with 500 µl of concentrated virus. We added polybrene at a concentration of 4 µg/ml and incubated at 37°C for four hours. The infected cells were then transferred to a T25 flask containing 9 ml of RPMI/IL-3 with polybrene. After 48 hours, we added G418 at a final concentration of 1 mg/ml. We subcloned the pool of infected, G418-resistant cells by serial dilution in a 96-well plate. We expanded individual clonal lines and tested them for receptor expression, either by immunoblotting (described below) or by determining their ability to survive in the absence of IL-3 and the presence of PDGF-BB (Calbiochem, San Diego, CA, USA). Clonal lines expressing the desired receptor and maintaining a strict requirement for IL-3 were used for the experiments reported here.

To express small transmembrane proteins in BaF3 cell lines, we prepared retroviral particles from HEK293T cells and infected BaF3 cells with up to 500 µl of concentrated virus, as described above. To select for infected cells, the cells were incubated in RPMI/IL-3 + 1 mg/ml

hygromycin. We maintained selected cells in hygromycin and G418 (where applicable) at the concentrations used for selection.

C127 cells at approximately 50% confluence in a 60 mm dish were infected as follows: filtered, concentrated virus was brought to a volume of 1 ml in DMEM10. C127 cells were infected with a 1:20 dilution of this stock in DMEM10 to a final volume of 1 ml (virus expressing BPV E5 was diluted 1:200), and supplemented with 4 μg/ml polybrene. After four hours at 37°C, we aspirated the virus-containing medium and added 4 ml of DMEM10 containing polybrene. After 24 hours, the cells were passaged for focus forming assays and titer, as described below.

IL-3-Independence assays, focus forming assays, viral titering, and recovery of transforming clones. To assess whether infected BaF3 cultures required IL-3, cells were washed twice with 10 ml phosphate-buffered saline (PBS) and resuspended in 10 ml complete RPMI medium lacking IL-3. We maintained the cells in T25 flasks without further passaging and counted viable cells at various time points.

To determine focus-forming ability of individual clones in C127 cells, we trypsinized the infected or mock-infected cells 24 hours post-infection, resuspended the cells in a total of 12 ml DMEM10, and added 4 ml of the cell suspension into each of two 60 mm dishes. Cells were maintained in DMEM10 for 11 days, then fixed and stained with Giemsa (Sigma-Aldrich). To determine titer, we plated dilutions of the remaining cell suspension into 100 mm dishes. After 24 hours, we added hygromycin to the medium (final concentration: 0.3 mg/ml) and maintained the cells in selection medium until mock-infected cells had died. We then fixed and stained the cells, and counted colonies to determine colony-forming units per ml of virus. We normalized

focus forming ability to the number of colony-forming units from a given infection, taking into account the dilution of virus used to infect the cells.

To recover library sequences, we harvested genomic DNA from $2x10^6$ IL-3-independent BaF3 cells using a DNeasy kit (Qiagen, Valencia, CA, USA). We amplified the library sequences from 400 ng of genomic DNA using PCR primers specific to fixed sequences that flanked the randomized transmembrane segments. See Table S1 for primer sequences. Detailed PCR program information is available from the authors upon request. We subcloned the amplified inserts back into the original retroviral vectors used to build the libraries, (pRV-Hyg^R for LFC5; pT2H-F13 for KTS1) and sequenced both strands of individual clones.

PDGF β receptor inhibitor studies. BaF3-βR cells expressing BPV E5 or pTM32-1 were grown in RPMI/IL-3 to a density of approximately 10^6/ml. One hundred thousand cells from each culture were incubated in 10 ml RPMI or RPMI/IL-3 in the presence of 50 μM AG1295 (Calbiochem) or an equivalent volume of dimethyl sulfoxide (DMSO) (the vehicle). C127 cells expressing BPV E5 or pTM36-4 were incubated for four days in DMEM10 containing 50 μm AG1295 or an equivalent volume of DMSO.

Biochemical analyses. For detection of proteins, lysates were prepared from approximately $4x10^7$ BaF3 cells in either RIPA-MOPS (20 mM morpholinepropanesulfonic acid, pH 7.0/150 mM NaCl/1 mM EDTA/1% Nonidet P-40/1% deoxycholate/1% SDS) or CHAPS [15 mM 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate /30 mM NaCl/1 mM EDTA/50 mM Tris-HCl (pH 7.4)] buffer containing protease and phosphatase inhibitors. We then immunoprecipitated the protein of interest essentially as described (7, 8). Briefly, we added 5-7 μl of αPR-A rabbit antiserum raised against the carboxy-terminal 13 amino acids of the human PDGF β receptor to 1 mg of protein extract. After rotating overnight at 4°C, we added 50

μl of protein A sepharose beads and rotated for at least one hour at 4°C. The beads were pelleted and washed with cold NET-N buffer containing 1 mM PMSF as described (8). For C127 cells, extracts were prepared from two or more confluent, 15 cm plates of cells, which were scraped in cold PBS into Eppendorf tubes, washed again with PBS, and then lysed on ice and processed as described (39). The lysates were centrifuged for 30 minutes at 14000 rpm in a tabletop centrifuge at 4°C. Lysates for phosphotyrosine blots were prepared from BaF3-βR cells using CHAPS buffer; all other lysates from BaF3 and C127 cells were made in RIPA-MOPS. We quantitated protein using a bicinchoninic acid (BCA) kit (Pierce/Thermo Scientific, Franklin, MA, USA).

To immunoprecipitate HA-tagged proteins, we added 50 μl of an anti-HA affinity matrix (Roche, Mannheim, Germany) to 1 mg of protein lysate, rotated overnight at 4°C, and proceeded as described for receptor immunoprecipitations. All immunoprecipitated samples were resuspended in 2x Laemmli sample buffer with 200 mM DTT and 5% β-mercaptoethanol.

We electrophoresed the samples at 150 volts for one hour in a 7.5% polyacrylamide gel for PDGF β receptor and phosphotyrosine blots or a 20% polyacrylamide gel for HA blots. Gels were soaked in transfer buffer (39), and then transferred electrophoretically for one hour at 90 volts to either 0.45 micron polyvinylidene difluoride (PVDF) membrane (Biorad, Hercules, CA, USA), for PDGF β receptor and HA blots, or 0.45 micron nitrocellulose (Biorad) for phosphotyrosine blots. We blocked the membranes for one hour at room temperature or at 4°C overnight, in 5% BSA/TBST (TBS with 0.1% [vol/vol] Tween 20) for phosphotyrosine blots or 5% milk/TBST for all others, as described (8, 39).

We immunoblotted for PDGF β receptor as previously described (8). For detection of phosphotyrosine, we used a P-Tyr-100 antibody (Cell Signaling Technology, Danvers, MA, USA) at a 1:1500 dilution in 5% BSA/TBST, and for anti-HA blots, we used a 1:500 dilution of

a mouse monoclonal antibody recognizing the HA epitope tag (a generous gift from Susan Baserga, Yale University) in 5% milk/TBST. We incubated all membranes in primary antibody overnight at 4°C, washed five times in TNET (8) (PDGF β receptor blots) or TBST (all others), and then added protein-A HRP secondary antibody (Amersham, 1:7500 in 5% milk/TBST) to detect PDGF β receptor or donkey anti-mouse HRP to detect phosphotyrosine (1:10,000 in 5% BSA/TBST) and HA (1:5,000 in 5% milk/TBST). Super Signal West Pico or Femto chemiluminescent detection (Pierce/Thermo Scientific) were used to visualize bands.

TOXCAT assay. The transmembrane sequences (positions 8-32) of wild-type pTM32-1, E19L, and T29L mutants were cloned into the TOXCAT chimeric construct, and expressed in *E. coli*. The level of oligomerization was determined via enzymatic CAT activity quantification, employing ^{3}H-chloramphenicol, as described (41). The CAT activity was normalized using the known dimerization properties of GpA variants. Wild-type GpA, a strong dimer in the membrane, was used as a positive control, while the G83I mutant, which forms only a very weak dimer, was used as a negative control (41). The expression level of all constructs was determined by Western blotting, with an antibody which specifically recognizes the maltose binding protein (ZYMED laboratories, South San Francisco, CA, USA). The correct orientation of the construct in the *E. coli* inner membrane was confirmed through a proteinase K sensitivity assay.

Structural Models. *Modeling using CHI.* The computational search strategy to generate models for dimeric transmembrane peptides employing CHI (crystallography and NMR System (CNS) searching of helix interactions) has been described previously (1, 2, 4). Briefly, a pair of canonical alpha helices was constructed from the pTM32-1 sequence (residues W5-W32 or G11-W32, with similar results in both cases), with crossing angles of +35° and -35°, a distance between helices of 10.5 Å and rotation increments of 15°. Molecular dynamics (MD) simulations

were performed using simulated annealing of atomic coordinates. Energy minimization of structures was performed before and after MD simulations, and groups of structures with a backbone RMSD of 1 Å or less were placed into clusters. The side chain of the E19 residue was considered in its negatively charged state.

Molecular dynamics simulations using NAMD. Three simulations were performed with the peptide dimers obtained from CHI modeling inserted in the transmembrane orientation into equilibrated and hydrated 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) bilayer patches containing 64 lipids in each leaflet. The lipid and water molecules overlapping with the peptide were removed. MD simulations were performed using the code NAMD (37) with CHARMM22 protein force field (27), CHARMM27 lipid force field (9), and TIP3P water model (19). Full details of the simulation conditions are in Supplementary Information.

Peptide synthesis, purification, and analysis. Peptides corresponding to the transmembrane domain of E5 ($E5_{TM}$) and pTM36-4 ($pTM36\text{-}4_{TM}$) were synthesized by the Keck Biotechnology Resource Center at Yale University using F-moc chemistry. The sequence of the $E5_{TM}$ peptide was KKKFLGLVAAMQLLLLLFLLLFFLVYWDHK, containing residues F_9-H_{34} from E5 as well as non-native lysine residues to aid solubility and serve as end-caps on the C- and N- termini. The sequence of the $pTM36\text{-}4_{TM}$ peptide was KKKFLGIINLLTLFLITLILIILVFYWDHK, containing the corresponding region of this protein. Both peptides were purified by reversed-phase HPLC using a linear acetonitrile gradient including 0.1% trifluoroacetic acid on a Phenomenex C4 column. The purity of pooled fractions was confirmed by mass spectrometry before lyophilisation.

Cross-linking reactions were carried out using 20 µM peptide dissolved in 10 mM SDS micelles. All samples were prepared in buffer containing 20 mM sodium phosphate and 150 mM

NaCl (pH 8). Bis[sulfosuccinimidyl]suberate (BS_3, Pierce) was used according to the manufacturer's protocol to cross-link the peptide in solution via primary amine groups. The cross-linking reaction was terminated after 30 min by the addition of 1 M Tris-HCl (pH 8). Uncross-linked samples were prepared as controls. Samples were separated by gel electrophoresis using a NuPage 12% Bis-tris gel with the running buffer containing 0.1% SDS (Invitrogen, Paisley, UK), and peptides were visualized by staining with silver nitrate.

RESULTS

Use of a novel screening approach to isolate a 32-amino acid activator of PDGF β receptor.

To identify activators of the PDGF β receptor without relying on their ability to induce focus formation, we employed a novel screening approach with a derivative of BaF3 cells, a mouse hematopoietic pro-B cell line that ordinarily requires IL-3 for growth. Because parental BaF3 cells do not express endogenous PDGF β receptor, they do not grow in the presence of PDGF-BB or BPV E5 unless IL-3 is supplied. For the experiments reported here, we used BaF3 cells engineered to express the murine PDGF β receptor (designated BaF3-βR cells) (8). As reported previously, BaF3-βR cells proliferate in the absence of IL-3 if PDGF-BB is added to the growth media or if they express *v-sis*, a viral homologue of PDGF. Importantly, expression of the wild-type E5 protein also allows BaF3-βR cells to proliferate in the absence of IL-3 (8).

We constructed a library expressing 44-amino acid proteins with randomized transmembrane domains from a retroviral vector. In this library, designated LFC5, the transmembrane region of the wild-type E5 protein was randomized from positions 14 through 30 and at position 33 (Fig. 1). The randomization scheme encoded predominantly hydrophobic amino acids, although approximately two hydrophilic residues, on average, were also encoded per randomized segment. At each randomized position, a stop codon was predicted to occur approximately 2% of the time.

To screen the LFC5 library for activators of the PDGF β receptor, we infected twelve independent cultures of clonal BaF3-βR cells with the library, selected for hygromycin resistance in the presence of IL-3, and then incubated the cells in the absence of IL-3. Cells containing the empty vector died in the absence of IL-3, but cells survived and proliferated in three of the pools infected with the library. We isolated genomic DNA from these pools, amplified the library

inserts using PCR with primers that annealed to the fixed portion of the vector (see Table S1), cloned the PCR products, and sequenced the recovered clones.

From one of the proliferating pools, we isolated only a single sequence, which was predicted to encode a 32-amino acid protein because a stop codon had been inserted at position 33. This protein, which we named pTM32-1 (Fig. 1), had a novel transmembrane sequence and lacked the twelve C-terminal residues of the wild-type E5 protein, including the aspartic acid and two cysteines thought to be required for the activity of the E5 protein and all other known transmembrane activators of the PDGF β receptor.

pTM32-1 induces IL-3 independence by activating the PDGF β receptor.

To determine whether pTM32-1 induced IL-3 independence and whether this activity required the PDGF β receptor, we stably expressed the wild-type E5 protein or pTM32-1 in parental BaF3 cells and in BaF3-βR cells. Infected cells were selected with hygromycin and then tested for growth factor independence (Fig. 2A). As expected, the empty vector did not induce IL-3 independence in either cell line, and the E5 protein induced IL-3 independence in BaF3-βR cells, which express the wild-type PDGF β receptor, but not in the parental BaF3 cells lacking PDGF β receptor expression. Like the E5 protein, pTM32-1 induced IL-3 independence in BaF3-βR cells but not in parental cells, demonstrating that it conferred growth factor independence in a PDGF β receptor-dependent fashion.

We utilized a selective PDGF β receptor kinase inhibitor, AG1295, to determine whether the tyrosine kinase activity of the receptor was required for the growth factor independence induced by pTM32-1 (Fig. 2B). AG1295 did not significantly affect the growth of BaF3-βR cells expressing BPV E5 or pTM32-1 in the presence of IL-3, because the growth of these cells was driven by IL-3 signaling. In contrast, in the absence of IL-3, BaF3-βR cells expressing BPV

E5 or pTM32-1 died in the presence of AG1295, indicating that sustained signaling by the PDGF β receptor was required for proliferation under these conditions.

To test whether pTM32-1 activated the PDGF β receptor, we immunoprecipitated the receptor from whole cell detergent lysates of BaF3-βR cells expressing empty vector or pTM32-1, electrophoresed the immunoprecipitated material, and immunoblotted with antibodies specific for the PDGF β receptor or for phosphotyrosine. Expression of the mature and precursor forms of PDGF β receptor was similar in both cell lines (Fig. 2C, top). The PDGF β receptor was not tyrosine phosphorylated in BaF3-βR cells infected with the empty vector. In contrast, the PDGF β receptor was phosphorylated in response to pTM32-1, indicating that pTM32-1 activated the receptor (Fig. 2C, bottom). This response was most striking for the precursor form of the PDGF β receptor, as we and others have previously noted for the E5 protein {*e.g.*, (3, 35, 44)}. These results provided biochemical evidence that pTM32-1 activated the PDGF β receptor.

To determine whether either hydrophilic residue in the transmembrane segment of pTM32-1, glutamic acid 19 or threonine 29, was important for its biological activity, we constructed leucine substitutions at these positions and tested the activity of the mutated proteins. Mutating the threonine to leucine had no effect on activity (data not shown), but changing the glutamic acid to leucine (E19L) abrogated the ability of the mutant protein to confer IL-3 independence (Fig. 2D). In addition, the E19L mutant no longer induced PDGF β receptor phosphorylation, as shown by immunoblotting (Fig. 2C). Collectively, these biochemical and physiological results demonstrated that pTM32-1 activated the PDGF β receptor in BaF3-βR cells, which resulted in growth factor independence, and suggested that these activities require a transmembrane glutamic acid, although it is possible that other hydrophilic amino acids could substitute for glutamic acid.

pTM32-1 specifically activates the PDGF β receptor by transmembrane interactions.

BPV E5 activates the PDGF β receptor by interacting directly with its transmembrane domain. To determine whether pTM32-1 also required the transmembrane domain of the PDGF β receptor, we utilized chimeric receptors in which the transmembrane domain of the wild-type PDGF β receptor was replaced with the transmembrane domain of either the closely-related PDGF α receptor or the stem cell factor receptor (c-Kit), another receptor tyrosine kinase (Fig. S1). These chimeras retained the extracellular ligand binding domain and the intracellular signaling domain of the PDGF β receptor, but they were unable to respond to the E5 protein because they lacked the transmembrane domain of the PDGF β receptor (10). Genes encoding these chimeric receptors were introduced into BaF3 cells, generating BaF3-βαβ and BaF3-βKitβ cells, respectively. These cells did not proliferate in the absence of growth factors, but they displayed growth factor independence in response to *v-sis*, which binds to the ligand binding domain retained in the chimeras, indicating that the chimeric receptors were expressed and functional despite the foreign transmembrane domain (Fig. 3).

We expressed the E5 protein or pTM32-1 in parental BaF3, BaF3-βR, BaF3-βαβ, or BaF3-βKitβ cells and assayed the growth of these cells in medium lacking IL-3. The E5 protein and pTM32-1 were inactive in parental BaF3 cells and in cells expressing the chimeric receptors, but they induced growth factor independence in cells expressing wild-type PDGF β receptor (Fig. 3). The empty vector was inactive in all cell types, as expected. Thus, the ability of pTM32-1 to induce IL-3 independence required the transmembrane domain of the PDGF β receptor. Taken together, our results show that pTM32-1 is, in effect, a 32-amino acid isolated transmembrane domain whose biological activity is dependent on the integrity of the PDGF β receptor transmembrane domain.

Dimerization and molecular modeling of pTM32-1

To determine whether the wild-type and mutant versions of pTM32-1 formed dimers in membranes like other transmembrane activators of the PDGF β receptor, we employed the TOXCAT assay (41). This assay utilizes constructs in which the transmembrane domain (residues L8-W32) of pTM32-1, E19L, or T29L is flanked by a maltose binding domain and by the oligomerization-dependent, DNA-binding domain of the ToxR transcriptional activator. We expressed these constructs in *E. coli* harboring a reporter gene in which chloramphenicol acetyltransferase (CAT) expression is driven by the ToxR-responsive *ctx* promoter (10, 41). If the transmembrane domain oligomerizes in the bacterial inner membrane, then the reconstituted ToxR transcriptional activator will activate the *ctx* promoter and drive expression of CAT. In this assay, CAT expression is proportional to the strength of the transmembrane interactions (41). As controls, we used glycophorin A (GpA), a transmembrane helix that forms a high-affinity dimer, and the GpA mutant, G83I, which dimerizes weakly. The fusion construct containing the transmembrane domain of pTM32-1 and its mutants induced high CAT activity, 1.6 times higher than that induced by wild-type GpA and similar to that induced by the transmembrane domain of the E5 protein (Fig. 4A and data not shown). Western blotting demonstrated that all constructs were expressed at a similar level (data not shown). These results indicate that pTM32-1 is capable of forming a stable oligomer in membranes, presumably a dimer, and that the mutations E19L and T29L do not reduce oligomerization.

We used the CHI molecular dynamics simulation protocol, as in previous studies (10, 46), to generate plausible structural models of the pTM32-1 homodimer. The structural calculations were performed on predicted transmembrane residues G11-W32 of pTM32-1. Three models were generated, as shown in Figure 4B, left: two left-handed coiled-coils (Models

1 and 2), and one right-handed coiled-coil (Model 3). The plots of interaction energy (Fig. 4B, right) show the predicted energetic contribution of each residue to the homodimeric interface in the corresponding model. The predicted interfaces of the three models differ significantly. Of note, the glutamic acid at position 19, which the TOXCAT experiment showed is not important for dimerization, is predicted to be a key energetic contributor to the dimerization interface only in the left-handed coiled-coil predicted in model 2. Glutamic acid 19 is oriented away from the interface in the other two models. Therefore, model 2 is inconsistent with the ability of the E19L mutant to form a strong dimer.

The CHI program optimizes the interactions of the helices relative to each other *in vacuo*, but it does not incorporate the molecular constraints and protein-lipid interactions intrinsic to a lipid bilayer environment. Additionally, it considers the side chain E19 as negatively charged, but it is reasonable to assume that this charged state would be unfavorable within the hydrophobic core of the membrane. To assess the stability of the three structures generated by the CHI program in a lipid milieu *in silico*, we performed simulations with the peptides inserted in DMPC bilayer patches using the program NAMD. The root mean square deviation (RMSD) values for backbone atoms shown in Figure 4C represent the deviation from the initial backbone coordinates during the course of the simulations (model 1: 2.63 ± 0.95 Å, model 2: 2.13 ± 0.59 Å and model 3: 1.33 ± 0.39 Å). A low RMSD value along with low standard deviation, as is the case for model 3, indicate higher stability of the initial structure in the lipid environment. Based on the RMSD values and the observation that model 2 requires Glu19 for stabilization of the homodimeric interface, we propose that model 3 is the most plausible structure for pTM32-1. Thus, unlike BPV E5 and several other small, transmembrane activators of the PDGF β receptor

that we previously isolated (10, 12), these data suggest that homodimeric pTM32-1 forms a right-handed coiled-coil.

Truncated BPV E5 protein binds to and activates the PDGF β receptor.

The ability of pTM32-1 to activate the PDGF β receptor was surprising because we and others had previously reported that the full-length E5 protein requires the C-terminal cysteine residues for transforming activity (17, 29, 42). In our previous screens of small, transmembrane protein libraries, we used a low expression vector and selected for focus formation in C127 fibroblasts, which endogenously express PDGF β receptor. Now we are using retroviral vectors that express small transmembrane proteins at much higher levels than vectors used in the past (data not shown), and pTM32-1 was isolated because it induced growth factor independence in hematopoietic cells expressing exogenous PDGF β receptor. We hypothesized that high-level expression of an E5 protein lacking the cysteines might dimerize non-covalently in cells and induce growth factor independence because the transmembrane domain of BPV E5 has intrinsic ability to dimerize, as assessed by TOXCAT experiments and biophysical studies (10, 32).

Therefore, we determined whether over-expression of a truncated version of the BPV E5 protein that lacked the C-terminal cysteines but retained the wild-type transmembrane domain bound and activated the PDGF β receptor in BaF3-βR cells and induced growth factor independence. We introduced a stop codon at position 34 in the wild-type E5 sequence to generate a clone that expressed the first 33-amino acids of the E5 protein (Fig. 1). We chose to keep the aspartic acid at position 33 in this mutant, because prior mutational analysis demonstrated the importance of this residue in the full-length E5 protein for activation of the PDGF β receptor (20). We also added an N-terminal HA epitope tag in order to visualize this protein biochemically because the C-terminal truncation removed the epitope recognized by the

anti-E5 antibody. The resulting protein, named HA/$E5_{33}$, was expressed in parental BaF3 cells, BaF3-βR cells, and BaF3-βKitβ cells and assayed for its ability to support IL-3 independent growth. Like the full-length E5 protein, HA/ $E5_{33}$ was inactive in parental BaF3 and BaF3-βKitβ cells, but it induced IL-3 independence in cells expressing wild-type PDGF β receptor (Fig. 5A). All of the infected BaF3-βKitβ cell lines grew in the presence of PDGF-BB, demonstrating that the chimeric receptors were expressed and functional (data not shown). These results indicated that a truncated BPV E5 protein lacking the C-terminus can functionally interact with the transmembrane domain of the wild-type PDGF β receptor, but not the c-kit transmembrane domain, and induce growth factor independence. To determine whether HA/$E5_{33}$ induced tyrosine phosphorylation of the PDGF β receptor, we immunoprecipitated the receptor from BaF3-βR cells expressing HA/$E5_{33}$, the full-length E5 protein, or an empty vector control, and immunoblotted for the receptor and phosphorylated tyrosine. The wild-type PDGF β receptor was expressed in all transduced cell lines, and it was tyrosine phosphorylated in BaF3-βR cells expressing HA/$E5_{33}$ or BPV E5, but not in cells containing the empty vector (Fig. 5B). Thus, HA/$E5_{33}$, like pTM32-1 and BPV E5, specifically activated the PDGF β receptor via transmembrane interactions.

To determine whether the truncated E5 protein or pTM32-1 was able to activate the PDGF β receptor in the cell system we traditionally use to measure E5-mediated cell transformation, we expressed wild-type BPV E5, HA/$E5_{33}$, pTM32-1, and an empty retroviral vector control in C127 cells and assayed for focus formation. The control vector was devoid of transforming activity, while HA/$E5_{33}$ and pTM32-1 exhibited very low levels of focus forming activity, approximately 2% and 0.2%, respectively, the level of the wild-type E5 protein (Table 1, Fig. 6A). Although there was background tyrosine phosphorylation of the mature form of the

PDGF β receptor in untransformed C127 cells, both HA/E5$_{33}$ and pTM32-1 induced tyrosine phosphorylation of the precursor form of the PDGF β receptor in C127 cells, as determined by immunoprecipitation of the receptor from whole cell lysates and immunoblotting for tyrosine phosphorylation (Fig. 6B). Thus, removal of the C-terminal segment of the E5 protein markedly inhibited but did not eliminate focus forming activity in C127 cells, although this truncated protein still induced tyrosine phosphorylation of the PDGF β receptor.

By immunoprecipitating with a monoclonal antibody that recognizes the HA epitope, followed by immunoblotting with the same antibody, we demonstrated that the truncated E5 protein was expressed in infected BaF3-βR and C127 cells (data not shown). In addition, we used co-immunoprecipitation to determine whether HA/E5$_{33}$ formed a stable complex with the PDGF β receptor in C127 cells. Whole cell detergent lysates were immunoprecipitated with anti-HA affinity matrix beads, separated by SDS-PAGE, and immunoblotted using an antibody recognizing the PDGF β receptor. As shown in Figure 6C, the HA antibody immunoprecipitated the precursor form of the PDGF β receptor (predicted molecular weight of approximately 200 kDa) from cells expressing HA/E5$_{33}$, but not from cells expressing the untagged E5 protein. In the reciprocal experiment, immunoprecipitation of the PDGF β receptor followed by immunoblotting with the anti-HA antibody also demonstrated a stable interaction between the PDGF β receptor and HA/E5$_{33}$ (data not shown). Collectively, these data indicated that HA/E5$_{33}$ and pTM32-1 activated the PDGF β receptor sufficiently to induce its phosphorylation and to transform BaF3-βR cells, but these extremely small proteins were markedly defective in their ability to induce focus formation in C127 cells. Strikingly, the truncated E5 protein that lacked the C-terminal segment still bound and activated the PDGF β receptor in C127 cells.

The C-terminal domain of the E5 protein is not required for focus formation in mouse fibroblasts.

Unlike the truncated HA/E5$_{33}$ and pTM32-1 proteins, the full-length E5 protein efficiently induces focus formation in C127 cells. To determine whether it was possible to identify small transmembrane proteins that lacked the C-terminal domain of the E5 protein yet efficiently transformed both BaF3-βR cells and mouse C127 fibroblasts, we constructed and screened a new library designed to encode small transmembrane proteins lacking this domain. In this library, designated KTS1, residues 14 to 30 of the wild-type E5 sequence were randomized to encode primarily hydrophobic amino acids, and a fixed stop codon was inserted at position 33 (Fig. 1). We also appended a hexahistidine tag at the amino terminus, resulting in a library that encoded proteins with a total size of 36 amino acids. The library was introduced by retroviral infection into BaF3-βR cells in the presence of IL-3. After hygromycin selection, IL-3 was removed from the growth medium, genomic DNA was isolated from cells that proliferated in the absence of exogenous growth factors, and inserts were recovered by PCR and sequenced.

The sequences of the transmembrane domains of two of the recovered clones, pTM36-3 and pTM36-4, differed markedly from one another and from the transmembrane sequences of the E5 protein and pTM32-1 (Fig. 1). Expression of pTM36-3 or pTM36-4 induced growth factor independence in BaF3-βR cells, but not in BaF3-βKitβ cells, demonstrating that these clones functionally interacted with the transmembrane domain of the PDGF β receptor (Fig. 7A). Furthermore, in the absence of IL-3, BaF3-βR cells expressing pTM36-3 or pTM36-4 died in the presence of the PDGF β receptor kinase inhibitor AG1295 (data not shown), confirming that PDGF β receptor signaling is required to induce growth factor independence. Thus, diverse

transmembrane proteins lacking the C-terminus of the E5 protein can activate the PDGF β receptor and thereby induce growth factor independence in BaF3-βR cells.

Remarkably, unlike pTM32-1 and HA-$E5_{33}$, pTM36-3 and pTM36-4 induced substantial focus formation in C127 cells (Fig. 6A). Indeed, after normalization for retroviral titer, pTM36-4 induced foci at >30% the level of the 44-amino acid wild-type E5 protein (Table 1), a 150-fold increase over pTM32-1, even though it too lacked the C-terminal amino acids of the native viral protein and shared only seven consecutive amino acids with the transmembrane domain of the wild-type E5 protein. As shown in Figure 7B, the kinase inhibitor AG1295 caused a reversion of the transformed morphology of C127 cells expressing BPV E5 or pTM36-4, demonstrating that sustained PDGF receptor signaling was required for this phenotype. In addition, phosphotyrosine blotting demonstrated that these small transmembrane proteins induced tyrosine phosphorylation of the precursor form of the PDGF β receptor in transformed C127 cells (Fig. 7C). Thus, certain transmembrane domains do not require the C-terminal portion of the E5 protein to induce PDGF β receptor activation, growth factor independence, and efficient focus formation.

To determine whether the hydrophobic segment of pTM36-4 was able to form a dimer, we synthesized a peptide consisting of pTM36-4 amino acids 13 to 36 (corresponding to E5 amino acids 9 to 32). This peptide was reconstituted in SDS micelles, and its ability to dimerize was assessed by electrophoresis. As shown in Figure 8, the pTM36-4 peptide migrated as a monomer in the absence of a chemical cross-linking agent. However, if the peptide was treated with cross-linker prior to electrophoresis, a clear fraction migrated as a sharp band at the position of a homodimer. The appearance of this band and not a ladder or high molecular weight smear suggests that this is a specific dimer and not a higher order oligomer or an aggregate. A substantial fraction of a control E5 transmembrane peptide migrated as a dimer in the presence or

absence of cross-linker. This experiment indicated that the pTM36-4 transmembrane domain has weak intrinsic homodimerization potential. Attempts to use co-immunoprecipitation to show complex formation between pTM36-4 and the PDGF β receptor were unsuccessful. If pTM36-4 and the PDGF β receptor form a complex, it is not stable in our lysis conditions or during biochemical analysis, perhaps because of the relative instability of the pTM36-4 dimer.

DISCUSSION

By conducting a genetic screen in BaF3 cells engineered to express the PDGF β receptor, we isolated pTM32-1, a 32-residue artificial transmembrane protein that specifically activated the PDGF β receptor to induce growth factor independence. pTM32-1 is the shortest transforming protein yet isolated, and unlike other known transmembrane activators of the PDGF β receptor, it lacked the entire C-terminus including both cysteines and the aspartic acid, was almost devoid of focus forming activity, and appeared to dimerize non-covalently and form a right-handed coiled-coil. These findings indicate that there are multiple transmembrane sequence motifs and configurations that can drive PDGF β receptor activation. These results led us to reassess the role of the C-terminus of the E5 protein itself. We demonstrated that an HA-tagged, truncated E5 ($HA/E5_{33}$), which lacks the two cysteines and nine other C-terminal residues, still bound and activated the PDGF β receptor and induced growth factor independence in BaF3-βR cells. $HA/E5_{33}$ also induced focus formation in C127 cells at markedly reduced levels compared to the wild-type protein. We attribute the ability of this truncated protein to transform cells to our current use of retroviral vectors that yield much higher expression levels than the vectors we used previously.

To determine whether the E5 C-terminus was required for efficient focus formation, we screened a library of randomized transmembrane proteins that lacked this segment. We identified a 36-residue protein, pTM36-4, which efficiently induced focus formation in C127 cells as well as growth factor independence in BaF3-βR cells. Other than the six-residue histidine tag and the initiating methionine, pTM36-4 consisted of only 29 amino acids and shared only seven consecutive amino acids with the E5 protein, which are not required for focus formation (6, 18). Therefore, essentially the entire amino acid sequence of BPV E5 can be

583 changed, or even partially removed, without eliminating high-level transforming activity or
584 specificity for the transmembrane domain of the PDGF β receptor. Our findings corroborate a
585 previous report showing that a 34-residue E5 protein with only a single cysteine was still capable
586 of inducing foci, although the role of PDGF β receptor was not assessed in that study (29), and
587 contrast to an older report that the carboxyl-terminus has independent mitogenic activity (15).
588 Do pTM32-1 and pTM36-4 activate the PDGF β receptor by using similar mechanisms
589 as those used by the E5 protein? The cysteines and glutamine 17 in the E5 protein are
590 important for homodimerization of the E5 monomers, which is thought to generate two identical
591 binding sites for the receptor. This allows the viral protein to bind the transmembrane domains
592 of two PDGF β receptor molecules simultaneously, inducing receptor dimerization and trans-
593 phosphorylation. In addition, glutamine 17 and aspartic acid 33 appear to form specific contacts
594 with amino acids in the receptor, but they can be replaced by structurally similar amino acids
595 (20, 21). The small artificial proteins reported here lack glutamine 17, aspartic acid 33, and the
596 cysteines. Nevertheless, pTM32-1, pTM36-4, and the E5 protein share important features: they
597 are predominantly hydrophobic, extremely small, likely to form alpha helical, homodimeric
598 coiled-coils in membranes, contain one or two transmembrane hydrophilic amino acids, and
599 require the transmembrane domain of the PDGF β receptor for activity. Therefore, we think it is
600 likely that pTM36-4 and pTM32-1 use variations of the strategies employed by the E5 protein to
601 bind to the transmembrane region of the PDGF β receptor and induce receptor dimerization and
602 transphosphorylation. For example, the hydrophilic residues in pTM32-1 and pTM36-4 may
603 form essential hydrogen bonds with amino acids in the transmembrane domain of the PDGF β
604 receptor, a requirement that would account for the defect caused by the glutamic acid mutation in
605 pTM32-1. However, because we have not demonstrated that these proteins form a complex with

606 the receptor, it remains possible that the E5 protein and the proteins reported here employ
607 different mechanisms to activate the PDGF β receptor.

608 Our results suggest that tiny hydrophobic proteins with apparently unrelated sequences
609 can bind the PDGF β receptor by using specific hydrophilic residues that form hydrogen bonds
610 or electrostatic contacts with the receptor, as well as hydrophobic side chains that generate a
611 three-dimensional hydrophobic molecular surface that binds a complementary surface of the
612 transmembrane domain of the receptor. These motifs may be difficult to recognize if they are
613 generated by residues that are not contiguous in the primary amino acid sequence of these small
614 proteins or if they are formed by amino acids from both monomers of dimeric proteins, as we
615 proposed for the wild-type E5 protein (46). Further experiments are required to determine the
616 mechanism of PDGF β receptor activation by these new proteins and will shed insight into
617 alternative molecular strategies to activate the receptor from within the membrane.

618 Transmembrane activators of the PDGF β receptor display distinct activity profiles.
619 Wild-type BPV E5, pTM36-3, and pTM36-4 efficiently induced focus formation and growth
620 factor independence, whereas pTM32-1 and HA-$E5_{33}$ induced growth factor independence but
621 were markedly defective for focus formation. We reported earlier that some small
622 transmembrane proteins that efficiently induce foci are severely impaired in their ability to
623 activate specific transient mitogenic signaling pathways (39). In addition, some point mutants of
624 BPV E5 induce PDGF β receptor tyrosine phosphorylation and enhanced growth of human
625 diploid fibroblasts, but not focus formation, differences that appear to be due to activation of
626 specific PDGF receptor-initiated intracellular signaling pathways (31, 36). These results suggest
627 that activation of the PDGF β receptor is not an all-or-nothing phenomenon, but rather more
628 nuanced, with different small transmembrane proteins eliciting different cellular outcomes,

perhaps by activating different signaling pathways. We also note that the short transmembrane proteins reported here preferentially activate the precursor, presumably intracellular, form of the PDGF β receptor. This tendency was previously observed for the wild-type E5 protein but is more pronounced for these shorter proteins, and may reflect different primary cellular localization of these proteins or subtly different interactions with the PDGF β receptor.

The development of BaF3 cells as a platform for screening libraries of small transmembrane proteins for activators of the exogenous PDGF β receptor liberates us from relying on focus formation in mouse fibroblasts and will allow the recovery of activators, and possibly inhibitors, of a variety of exogenous transmembrane targets. Furthermore, libraries of truncated small transmembrane proteins that cannot undergo covalent dimerization will express monomeric proteins that could display various biological activities.

Our results hold promise for the design and synthesis of biologically active, transmembrane peptides based on the sequences recovered from genetic libraries. Not only are the proteins described here significantly shorter than the full-length E5 protein, but they lack cysteine residues that might induce aggregation. It may be possible to add such peptides to cells where they can insert into membranes, bind their targets, and exert biological activity. Such peptides would also be useful for structural analysis because of their more favorable chemical properties than proteins containing multiple cysteines. Finally, because of the relatively simple structures adopted by transmembrane helices, they may serve as templates for the design of peptidomimetic compounds with important research or clinical applications.

Viruses have responded to the evolutionary pressures of their size by using small transmembrane proteins to modulate cellular targets. Here, we have followed their example and constructed extremely small transmembrane proteins that modulate a much larger cellular target,

652 thereby exerting a tremendous effect on cell behavior. Because up to 30% of all cellular proteins
653 are thought to have membrane-spanning helices (25), a substantial fraction of the biochemical
654 functions of cells could be modulated by similar mechanisms. It seems likely that cells express
655 similar, small transmembrane proteins as a powerful and energy-efficient mechanism to regulate
656 their biochemical activities. Such proteins would not have been detected in standard genomics
657 searches or biochemical analyses and may represent elusive protein analogues of microRNAs.

658 **Acknowledgments**

659 We thank Lisa Freeman-Cook, Tobin Cammett, and Gregory Korbel for help with library
660 design and construction and construction of chimeric receptors, Emily Freed for assistance with
661 the isolation of pTM32-1, and Jan Zulkeski for assistance in preparing this manuscript. The
662 molecular dynamics simulations using NAMD were performed at the Yale Center for High
663 Performance Computation in Biology and Biomedicine. F.N.B. was supported by a postdoctoral
664 fellowship from the Fundación Alfonso Martin Escudero. J.O. and A.M.D. were supported by
665 Cancer Research UK (CR-UK) grant number C21449/A6926 to A.M.D. This work was
666 supported by a grant from the NIH to DD (CA37157).

REFERENCES

1. **Adams, P. D., I. T. Arkin, D. M. Engelman, and A. T. Brunger.** 1995. Computational searching and mutagenesis suggest a structure for the pentameric transmembrane domain of phospholamban. Nature Struct. Biol. **2:**154-162.
2. **Adams, P. D., D. M. Engelman, and A. T. Brunger.** 1996. Improved prediction of the structure of the dimeric transmembrane domain of glycophorin A obtained through global searching. Proteins: Struct. Funct. Genet. **26:**257-261.
3. **Adduci, A. J., and R. Schlegel.** 1999. The transmembrane domain of the E5 oncoprotein contains functionally discrete helical faces. J. Biol. Chem. **274:**10249-10258.
4. **Brunger, A. T., P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. L. Warren.** 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta. Crystallogr. D. Biol. Crystallogr. **54:**905-921.
5. **Cohen, B. D., D. J. Goldstein, L. Rutledge, W. C. Vass, D. R. Lowy, R. Schlegel, and J. T. Schiller.** 1993. Transformation-specific interaction of the bovine papillomavirus E5 oncoprotein with the platelet-derived growth factor receptor transmembrane domain and the epidermal growth factor receptor cytoplasmic domain. J. Virol. **67:**5303-5311.
6. **DiMaio, D., D. Guralski, and J. T. Schiller.** 1986. Translation of open reading frame E5 of bovine papillomavirus is required for its transforming activity. Proc Natl Acad Sci U S A **83:**1797-801.
7. **Drummond-Barbosa, D., and D. DiMaio.** 1997. Virocrine transformation. Biochim. Biophys. Acta. **1332:**M1-17.
8. **Drummond-Barbosa, D. A., R. R. Vaillancourt, A. Kazlauskas, and D. DiMaio.** 1995. Ligand-independent activation of the platelet-derived growth factor beta receptor: requirements for bovine papillomavirus E5-induced mitogenic signaling. Mol. Cell. Biol. **15:**2570-2581.
9. **Feller, S. E., D. Yin, R. W. Pastor, and A. D. MacKerell.** 1997. Molecular dynamics simulation of unsaturated lipid bilayers at low hydration: parameterization and comparison with diffraction studies. Biophys. J. **73:**2269-2279.
10. **Freeman-Cook, L., A. M. Dixon, J. B. Frank, Y. Xia, L. Ely, M. Gerstein, D. M. Engelman, and D. DiMaio.** 2004. Selection and characterization of small random transmembrane proteins that bind and activate the platelet-derived growth factor β receptor. J. Mol. Biol. **338:**907-920.
11. **Freeman-Cook, L. L., and D. DiMaio.** 2005. Modulation of cell function by small transmembrane proteins modeled on the bovine papillomavirus E5 protein. Oncogene **24:**7756-7762.
12. **Freeman-Cook, L. L., A. P. B. Edwards, A. M. Dixon, K. E. Yates, L. Ely, D. M. Engelman, and D. DiMaio.** 2005. Specific locations of hydrophilic amino acids in constructed transmembrane ligands of the platelet-derived growth factor β receptor. J. Mol. Biol. **345:**907-921.
13. **Goldstein, D. J., T. Andresson, J. J. Sparkowski, and R. Schlegel.** 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:**4851-4859.

14. **Goldstein, D. J., W. Li, L.-M. Wang, M. A. Heidaran, S. A. Aaronson, R. Shinn, R. Schlegel, and J. H. Pierce.** 1994. The bovine papillomavirus type 1 E5 transforming protein specifically binds and activates the beta-type receptor for platelet-derived growth factor but not other tyrosine kinase-containing receptors to induce cellular transformation. J. Virol. **68:**4432-4441.

15. **Green, M., and P. M. Loewenstein.** 1987. Demonstration that a chemically synthesized BPV1 oncoprotein and its C-terminal domain function to induce cellular DNA synthesis. Cell **51:**795-802.

16. **He, B., G. Y. Lin, J. E. Durbin, R. K. Durbin, and R. A. Lamb.** 2001. The SH integral membrane protein of the paramyxovirus Simian Virus 5 is required to block apoptosis in MDBK cells. J. Virol. **75:**4068-4079.

17. **Horwitz, B. H., A. L. Burkhardt, R. Schlegel, and D. DiMaio.** 1988. 44-amino-acid E5 transforming protein of bovine papillomavirus requires a hydrophobic core and specific carboxyl-terminal amino acids. Mol. Cell. Biol. **8:**4071-8.

18. **Horwitz, B. H., D. L. Weinstat, and D. DiMaio.** 1989. Transforming activity of a 16-amino-acid segment of the bovine papillomavirus E5 protein linked to random sequences of hydrophobic amino acids. J. Virol. **63:**4515-9.

19. **Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein.** 1983. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. **79:**926-935.

20. **Klein, O., D. Kegler-Ebo, J. Su, S. Smith, and D. DiMaio.** 1999. The bovine papillomavirus E5 protein requires a juxtamembrane negative charge for activation of the platelet-derived growth factor beta receptor and transformation of C127 cells. J. Virol. **73:**3264-3272.

21. **Klein, O., G. W. Polack, T. Surti, D. Kegler-Ebo, S. O. Smith, and D. DiMaio.** 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:**8921-8932.

22. **Kulke, R., and D. DiMaio.** 1991. Biological properties of the deer papillomavirus E5 gene in mouse C127 cells: growth transformation, induction of DNA synthesis, and activation of the platelet-derived growth factor receptor. J. Virol. **65:**4943-9.

23. **Kulke, R., B. H. Horwitz, T. Zibello, and D. DiMaio.** 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:**505-11.

24. **Lai, C. C., C. Henningson, and D. DiMaio.** 2000. Bovine papillomavirus E5 protein induces the formation of signal transduction complexes containing dimeric activated platelet-derived growth factor β receptor and associated signaling proteins. J. Biol. Chem. **275:**9832-9840.

25. **Lehnert, U., Y. Xia, T. E. Royce, C. S. Goh, Y. Liu, A. Senes, H. Yu, Z. L. Zhang, D. M. Engelman, and M. Gerstein.** 2004. Computational analysis of membrane proteins: genomic occurrence, structure prediction and helix interactions. Q. Rev. Biophys. **37:**121-146.

26. **Liu, Y., D. M. Engelman, and M. Gerstein.** 2002. Genomic analysis of membrane protein families: abundance and conserved motifs. Genome Biol. **3(10):**54.1-54.12.

27. **MacKerell, A. D., D. Bashford Jr., M. Bellott, R. L. Dunbrack, J. D. Evanseck Jr., M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K.**

Kuczera, F. T. K. Lau, C. Mattos, S. W. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. D. Reiher, B. Roux III, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B. **102:**3586-3616.

28. **Mattoon, D., K. Gupta, J. Doyon, P. J. Loll, and D. DiMaio.** 2001. Identification of the transmembrane dimer interface of the bovine papillomavirus E5 protein. Oncogene **20:**3824-3834.
29. **Meyer, A. N., Y.-F. Xu, M. K. Webster, A. S. Smith, and D. J. Donoghue.** 1994. Cellular transformation by a transmembrane peptide: structural requirements for the bovine papillomavirus E5 oncoprotein. Proc. Natl. Acad. Sci. USA **91:**4634-4638.
30. **Nilson, L. A., and D. DiMaio.** 1993. Platelet-derived growth factor receptor can mediate tumorigenic transformation by the bovine papillomavirus E5 protein. Mol Cell Biol **13:**4137-45.
31. **Nilson, L. A., R. L. Gottlieb, G. W. Polack, and D. DiMaio.** 1995. Mutational analysis of the interaction between the bovine papillomavirus E5 transforming protein and the endogenous beta receptor for platelet-derived growth factor in mouse C127 cells. J. Virol. **69:**5869-5874.
32. **Oates, J., M. Hicks, T. R. Dafforn, D. DiMaio, and A. M. Dixon.** 2008. *In vitro* dimerization of the bovine papillomavirus E5 protein transmembrane domain. Biochemistry **47:**8985-8992.
33. **Petti, L., and D. DiMaio.** 1994. Specific interaction between the bovine papillomavirus E5 transforming protein and the beta receptor for platelet-derived growth factor in stably transformed and acutely transfected cells. J. Virol. **68:**3582-3592.
34. **Petti, L., and D. DiMaio.** 1992. Stable association between the bovine papillomavirus E5 transforming protein and activated platelet-derived growth factor receptor in transformed mouse cells. Proc. Natl. Acad. Sci USA **89:**6736-6740.
35. **Petti, L., L. A. Nilson, and D. DiMaio.** 1991. Activation of the platelet-derived growth factor receptor by the bovine papillomavirus E5 transforming protein. EMBO J. **10:**845-855.
36. **Petti, L. M., E. C. Ricciardi, H. J. Page, and K. A. Porter.** 2008. Transforming signals resulting from sustained activation of the PDGFβ receptor in mortal human fibroblasts. J. Cell Sci. **121:**1172-1182.
37. **Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten.** 2005. Scalable molecular dynamics with NAMD. J. Comp. Chem. **26:**1781-1802.
38. **Pinto, L., and R. A. Lamb.** 2006. The M2 proton channels of influenza A and B viruses (Minireview). J. Biol. Chem. **281:**8997-9000.
39. **Ptacek, J. B., A. P. B. Edwards, L. L. Freeman-Cook, and D. DiMaio.** 2007. Packing contacts can mediate highly specific interactions between artificial transmembrane proteins and the PDGF beta receptor. Proc. Natl. Acad. Sci USA **104:**11945-11950.
40. **Riese, D. J., II, and D. DiMaio.** 1995. An intact PDGF signaling pathway is required for efficient growth transformation of mouse C127 cells by the bovine papillomavirus E5 protein. Oncogene **10:**1431-1439.
41. **Russ, W. P., and D. M. Engelman.** 1999. TOXCAT: a measure of transmembrane helix association in a biological membrane. Proc. Natl. Acad. Sci. USA **96:**863-868.

804 42. **Schapiro, F., J. Sparkowski, A. Adduci, F. A. Suprynowicz, R. Schlegel, and S.**
805 **Grinstein.** 2000. Golgi alkalinization by the papillomavirus E5 oncoprotein. J. Cell. Biol.
806 **148:**305-315.
807 43. **Schlegel, R., M. Wade-Glass, M. S. Rabson, and Y.-C. Yang.** 1986. The E5
808 transforming gene of bovine papillomavirus encodes a small hydrophobic protein.
809 Science **233:**464-467.
810 44. **Sparkowski, J., M. Mense, M. Anders, and R. Schlegel.** 1996. E5 oncoprotein
811 transmembrane mutants dissociate fibroblast transforming activity from 16-kilodalton
812 protein binding and platelet-derived growth factor receptor binding and phosphorylation.
813 J. Virol. **70:**2420-2430.
814 45. **Staebler, A., J. H. Pierce, S. Brazinski, M. A. Heidaran, W. Li, R. Schlegel, and D. J.**
815 **Goldstein.** 1995. Mutational analysis of the beta-type platelet-derived growth factor
816 receptor defines the site of interaction with the bovine papillomavirus type 1 E5
817 transforming protein. J. Virol. **69:**6507-6517.
818 46. **Surti, T., O. Klein, K. Aschheim, D. DiMaio, and S. O. Smith.** 1998. Structural models
819 of the bovine papillomavirus E5 protein. Proteins **33:**601-612.
820 47. **Talbert-Slagle, K., and D. DiMaio.** 2009. The bovine papillomavirus E5 protein and the
821 PDGF β receptor: It takes two to tango. Virology **384:**345-351.
822

FIGURE LEGENDS

Figure 1. Amino acid sequences of small transmembrane proteins and libraries. The wild-type bovine papillomavirus E5 protein sequence is shown in bold, as are the BPV E5 residues deliberately retained in the other small transmembrane proteins. All sequences are aligned to the fixed tryptophan at position 5 in BPV E5. Epitopes for antibody recognition are underlined. Residues randomized in small transmembrane protein libraries are represented by X's.

Figure 2. pTM32-1 specifically activates the PDGF β receptor. (A) BaF3 cells without exogenous receptor (dashed lines, open shapes) or with PDGF β receptor (solid lines, closed shapes) expressing empty retroviral vector (circle), BPV E5 (triangle), or pTM32-1 (square), were incubated in medium lacking IL-3. Live cells were counted on the indicated days. **(B)** BaF3-βR cells expressing BPV E5 or pTM32-1 were grown in the presence or absence of IL-3 as indicated and were treated with either DMSO (filled bars) or the PDGF β receptor kinase inhibitor AG1295 (open bars). Live cells were counted on day four. **(C)** PDGF β receptor was immunoprecipitated from BaF3-βR lysates expressing either empty vector, pTM32-1, or the E19L mutant. Blots were probed for PDGF β receptor expression (βR) or phosphorylated tyrosine (PY). The mature (M) and precursor (P) forms of the PDGF β receptor are indicated. **(D)** BaF3-βR cells expressing BPV E5 (triangle), pTM32-1 (square), or the pTM32-1 E19L mutant (circle), were incubated in medium lacking IL-3. Live cells were counted on indicated days. Results shown in all panels are representative of at least three independent experiments.

Figure 3. Induction of IL-3 independence by pTM32-1 requires the transmembrane domain of the PDGF β receptor. Parental BaF3 cells (open bars), BaF3-βR (filled bars), BaF3-

βαβ (hatched bars), or BaF3-βKitβ (gray bars) were infected with the empty retroviral vector, or with viruses expressing wild-type BPV E5, pTM32-1, or the *v-sis* oncogene. Cells were plated in medium lacking IL-3, and live cells were counted after five days. Results shown are representative of multiple independent experiments.

Figure 4. Dimerization and molecular modeling of pTM32-1. (A) The graph shows normalized CAT activity induced by fusion proteins with the indicated wild-type and mutant transmembrane domains in a TOXCAT assay, which is proportional to the strength of the transmembrane interactions. Mean and standard deviations of two independent experiments, done in triplicate, are shown. GpA, glycophorin A; G83I, glycophorin A glycine to isoleucine mutation at position 83. **(B) Left panel.** Helical backbone of three models of the pTM32-1 homodimer predicted by CHI molecular dynamics simulation. Glutamic acid at position 19 is shown in orange. **Right panel.** The plots show the interaction energy of the amino acids predicted to form the homodimer interface for each model. The sequence of the predicted transmembrane domain of pTM32-1 is shown at bottom. **(C)** Root mean square deviation (RMSD) of the backbone atoms from the initial coordinates for the three homodimer models during NAMD molecular dynamics simulations. Model 1, black; Model 2, red; Model 3, green.

Figure 5. Ha/E5$_{33}$ activates the PDGF β receptor in BaF3 cells. (A) Parental BaF3 cells (open bars), BaF3-βR cells (filled bars), or BaF3-βKitβ cells (gray bars) stably expressing either empty retroviral vector, wild-type BPV E5, or HA/E5$_{33}$ were plated in medium lacking IL-3, and live cells were counted after five days. Results shown are representative of three independent experiments. **(B)** PDGF β receptor was immunoprecipitated from BaF3-βR lysates expressing

either empty vector, BPV E5, or HA/$E5_{33}$. Blots were probed for PDGF β receptor expression (βR) or phosphorylated tyrosine (PY). The mature (M) and precursor (P) forms of the PDGF β receptor are indicated.

Figure 6. Truncated E5 protein induces low-level focus formation, forms a stable complex with the PDGF β receptor in C127 cells, and induces receptor phosphorylation. (A) C127 cells were infected with high-titer stocks of the empty retroviral vector or viruses expressing the indicated small transmembrane protein and incubated for 11 days. **(B)** PDGF β receptor was immunoprecipitated from lysates of C127 cells expressing empty vector, BPV E5, HA/$E5_{33}$, or pTM32-1. Blots were probed for the receptor (βR) or phosphorylated tyrosine (PY). The mature (M) and precursor (P) forms of the PDGF β receptor are indicated. **(C)** Lysates of C127 cells expressing empty vector, BPV E5, or HA/$E5_{33}$ were immunoprecipitated with antibody recognizing the HA epitope and probed for the PDGF β receptor. Molecular weight marker is shown for reference.

Figure 7. pTM36-3 and pTM36-4 specifically activate the PDGF β receptor. (A) BaF3-βR (solid lines) and BaF3-βKitβ (dashed lines) expressing BPV E5 (squares), pTM36-3 (triangles), or pTM36-4 (circles) were incubated in medium lacking IL-3. Live cells were counted on indicated days. Results shown are representative of three independent experiments. **(B)** C127 cells stably transformed with the E5 protein or pTM36-4 were incubated for four days in medium containing DMSO (top) or AG1295 (bottom), and photographed by phase contrast microscopy. **(C)** PDGF β receptor was immunoprecipitated from lysates of C127 cells expressing empty vector, BPV E5, pTM36-3, or pTM36-4. Blots were probed for the receptor (βR) or

892 phosphorylated tyrosine (PY). The mature (M) and precursor (P) forms of the PDGF β receptor
893 are indicated.
894
895 **Figure 8. Dimerization of pTM36-4 peptide.** Synthetic peptides corresponding to the
896 transmembrane domains of pTM36-4 and BPV E5 were solubilized in 10 mM SDS, subjected to
897 crosslinking with BS_3, as indicated, and electrophoresed in bis-tris gels. The position of the
898 pTM36-4 peptide dimer is indicated by the arrowhead. The sizes of molecular weight markers
899 (in kDa) is shown on right.
900

Table 1.

	Foci/5,000 CFU[a]	% activity[b]
Vector	0	0
BPV E5	165	100
HA/$E5_{33}$	2.6	1.6
pTM32-1	0.3	0.2
pTM36-3	20.2	12.2
pTM36-4	52.2	31.6

[a]Average number of foci from two dishes of infected cells, normalized to viral titer
[b]Percent focus forming activity relative to wild-type BPV E5