Biologically Active Ultra-Simple Proteins Reveal Principles of Transmembrane Domain Interactions

Ross S. Federman [1], Anna-Sophia Boguraev [2,†], Erin N. Heim [2,†] and Daniel DiMaio [2,3,4,5]

*1 - **Department of Immunobiology,** Yale School of Medicine, PO Box 208011, New Haven, CT 06520-8011, USA*
*2 - **Department of Genetics,** Yale School of Medicine, PO Box 208005, New Haven, CT 06520-8005, USA*
*3 - **Department of Therapeutic Radiology,** Yale School of Medicine, PO Box 208040, New Haven, CT 06520-8040, USA*
*4 - **Department of Molecular Biophysics & Biochemistry,** Yale School of Medicine, PO Box 208024, New Haven, CT 06520-8024, USA*
*5 - **Yale Cancer Center,** PO Box 208028, New Haven, CT 06520-8028, USA*

***Correspondence to Daniel DiMaio:** Department of Genetics, Yale School of Medicine, PO Box 208005, New Haven, CT 06520-8005, USA. daniel.dimaio@yale.edu*
https://doi.org/10.1016/j.jmb.2019.07.009

Abstract

Specific interactions between the helical membrane-spanning domains of transmembrane proteins play central roles in the proper folding and oligomerization of these proteins. However, the relationship between the hydrophobic amino acid sequences of transmembrane domains and their functional interactions is in most cases unknown. Here, we use ultra-simple artificial proteins to systematically study the sequence basis for transmembrane domain interactions. We show that most short homopolymeric polyleucine transmembrane proteins containing single amino acid substitutions can activate the platelet-derived growth factor β receptor or the erythropoietin receptor in cultured mouse cells, resulting in cell transformation or proliferation. These proteins displayed complex patterns of activity that were markedly affected by seemingly minor sequence differences in the ultra-simple protein itself or in the transmembrane domain of the target receptor, and the effects of these sequence differences are not additive. In addition, specific leucine residues along the length of these proteins are required for activity, and the positions of these required leucines differ based on the identity and position of the central substituted amino acid. Our results suggest that these ultra-simple proteins use a variety of molecular mechanisms to activate the same target and that diversification of transmembrane domain sequences over the course of evolution minimized off-target interactions.

Introduction

Transmembrane (TM) proteins comprise roughly 30% of the eukaryotic cell proteome and usually contain one or more hydrophobic, helical membrane-spanning domains. These TM domains (TMDs) often interact with one another in multi-pass TM proteins and multi-protein complexes to determine the functional folded state and association of protein subunits. These interactions are dictated primarily by amino acid side-chains that participate in van der Waals packing interactions requiring precise molecular arrangement or in electrostatic interactions and hydrogen bonding [1–4]. However, due to the difficulty in purifying and analyzing TMDs, there is less structural information to guide our molecular understanding of these interactions compared with interactions between soluble proteins [5]. In addition, lipids, which vary among different subcellular membranes, presumably play important roles in TMD interactions, further complicating efforts to understand TMD folding and structure [6,7].

Despite the overall structural and chemical similarity of most TMDs, their amino acid sequences vary [8], and the relationship between the sequence and function of TMDs remains elusive. Several motifs involved in TM protein interactions have been identified [1]. The GxxxG motif provides

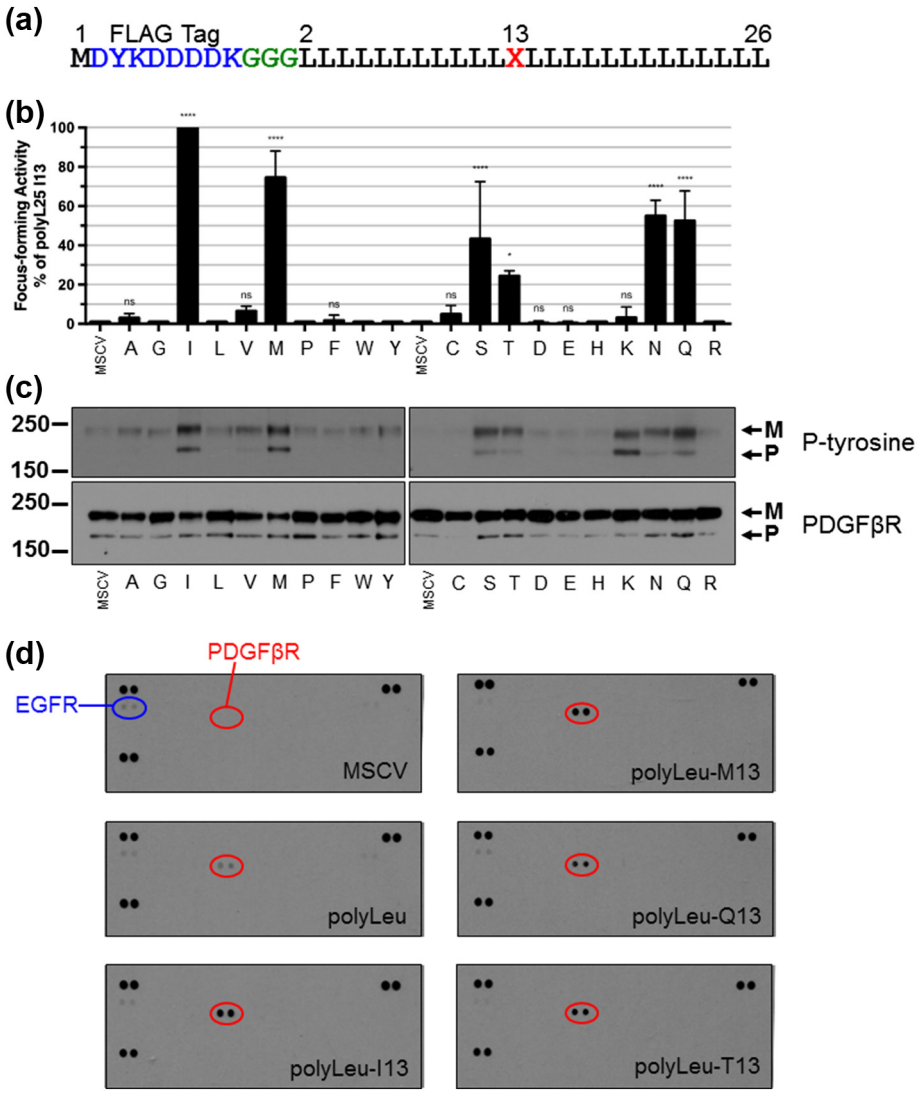

Fig. 1. Polyleucine traptamers with amino acid substitutions at position 13 activate the PDGFβR. (a) Ultra-simple traptamer sequence in single-letter amino acid code, where red "X" denotes one of the 20 standard amino acids, blue residues comprise the FLAG epitope tag, and green residues comprise a flexible linker. (b) C127 cells were infected with empty vector (MSCV) or retroviruses expressing each of the polyleucine traptamers with different residues substituted at position 13. Cellular transformation was quantitated by counting foci after 3 weeks in culture in three independent experiments for each traptamer. Results were normalized for viral titer and are shown as a percentage of positive control polyLeu-I13, plus and minus standard error of the mean. Significance from ordinary one-way ANOVA Dunnett's multiple comparison test: ****$p < .0001$, *$p < .05$, ns = not significant for each sample compared to cells infected with MSCV. (c) Extracts were prepared from C127 cells expressing MSCV vector or an ultra-simple traptamer with the indicated amino acid at the 13th position. Extracts were immunoprecipitated with anti-PDGFβR antibody and blotted for phosphotyrosine (upper) or PDGFβR (lower). M and P indicate mature and precursor forms of the receptor, respectively. (d) Extracts were prepared from HFFs expressing MSCV or an ultra-simple traptamer with the indicated amino acid at the 13th position. Phospho-RTK array membranes were incubated with cell extracts, probed with anti-phosphotyrosine, and visualized by chemiluminescence. The paired spots representing PDGFβR are circled in red on all arrays. EGF receptor is circled in blue on the empty vector membrane. The dark pairs of spots in the corners are standards.

a fundamental framework for one mode of helix homodimerization determined by packing interactions, and other common packing motifs have recently been described and validated [3,9–12]. Similarly, hydrogen bonds and salt bridges play important roles in maintaining the architecture of TM complexes, for example, in the assembly of the T-cell receptor [13–15]. Unbiased *in vivo* screens have identified strongly interacting helices from libraries expressing artificial TMDs with

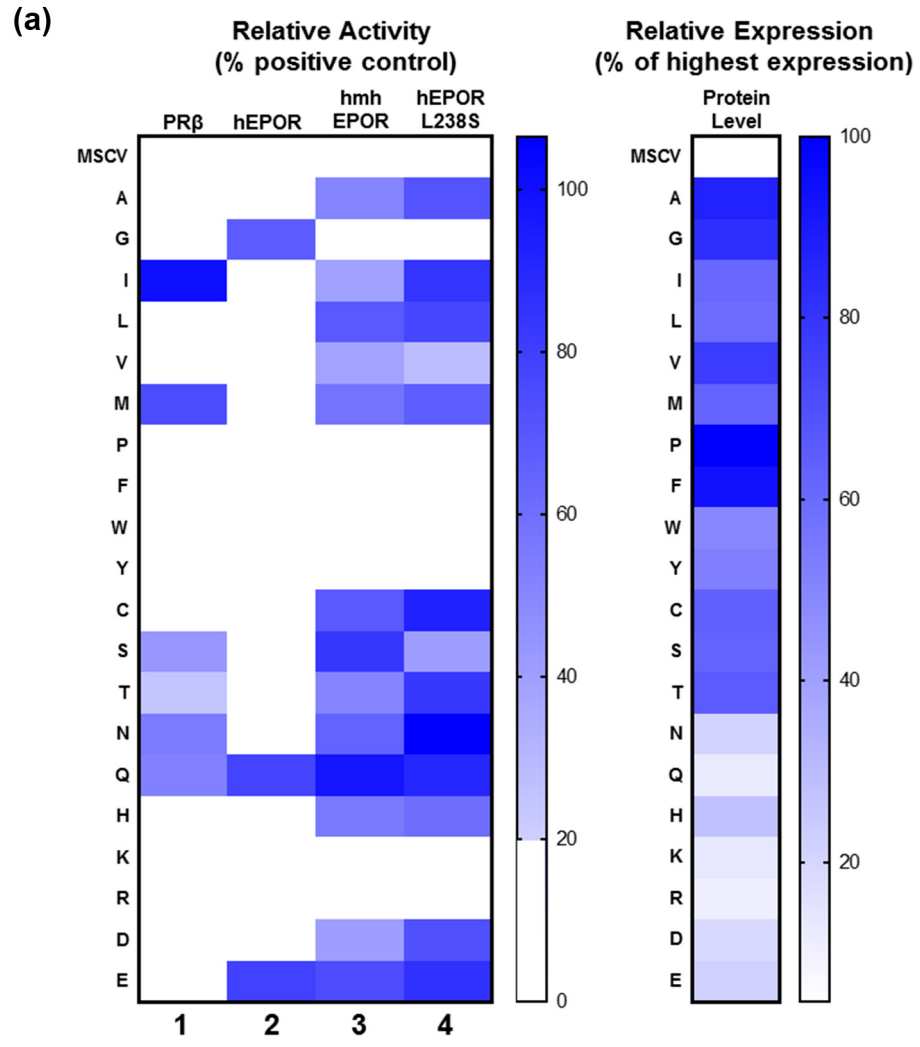

Fig. 2. Polyleucine traptamers with amino acid substitutions at position 13 activate the human and mouse EPORs. (a) BaF3 cells expressing the hEPOR, the hEPOR bearing the TMD of the mEPOR (hmh), or the hEPOR L238S TMD mutant were retrovirally infected to express ultra-simple traptamers with each of the 20 standard amino acids at position 13. The heat maps summarize the relative activity (left panels) and expression (right panels) of these traptamers. All values are shown as a percentage of expression of polyLeu-P13 (the highest value) in the case of expression levels, or as percentage of positive control (MSCV plus 0.06 units/mL EPO for EPORs; polyLeu-I13 for PDGFβR (PRβ)) in activity, where darker blues indicate values closer to the positive control, and white boxes indicate a value of zero or 20% or less in activity measurements. Primary data for hEPOR/hmhEPOR activity shown in Fig. S3. Activity in PDGFβR cells is the focus forming activity in C127 cells from Fig. 1b. (b) TMD sequences of the mouse PDGFβR, hEPOR, and mEPOR. Residues in red in the mEPOR sequence differ between mouse and human forms of the receptor.

randomized sequences [16–18], but these experiments do not provide a clear explanation for why certain sequences interact but others do not. It is clear that the basis for these interactions is more complex than can be explained through the known motifs alone [17].

Single-pass TMDs are often viewed as static structures, simply anchoring proteins within the

membrane, but they can undergo highly specific protein–protein interactions to mediate complex formation and affect protein activity within the lipid bilayer [1,4]. This is illustrated by the interaction between the E5 oncoprotein of bovine papillomavirus (BPV) and the platelet-derived growth factor β receptor (PDGFβR) [19]. Only 44 residues, BPV E5 is essentially a freestanding TMD that specifically interacts with the TM and juxtamembrane segment of the PDGFβR, acting like a set of clamps to drive productive dimerization of the PDGFβR in the absence of PDGF, resulting in transformation of cultured fibroblasts [20–28].

Using E5 as a platform for novel protein discovery, we developed an approach to screen libraries encoding short artificial TM proteins with randomized hydrophobic segments for their ability to affect cell behavior [29,30]. These artificial proteins are termed traptamers, for *tra*nsmembrane *p*rotein *aptamers*. We used this approach to isolate novel artificial oncogenes encoding transforming proteins as short as 29 amino acids and bearing no sequence similarity to E5 or other known proteins [31]. We also isolated LIL traptamers that activate the PDGFβR or the erythropoietin receptor (EPOR) [32,33]. These traptamers are only 26 amino acids long and comprise specific sequences of leucine and isoleucine residues following an initiating methionine. Remarkably, the ability of one LIL traptamer to distinguish between mouse and human forms of the EPOR can be toggled by changing the placement of a single side-chain methyl group at a few key positions within the traptamer [32]. We also discovered that a polyleucine TMD with a single leucine-to-isoleucine mutation at the 13th position (but not at neighboring positions) activated the PDGFβR and transformed mouse C127 fibroblasts [33]. We term these polyleucine proteins "ultra-simple traptamers." We reasoned that systematic analysis of the biological activity of ultra-simple traptamers might provide new insight into the principles governing TMD interactions.

Several studies have employed similar polyleucine TMDs to examine TM helix homo-dimerization and found that leucine residues contribute to packing interactions along interacting helical interfaces and that replacing interfacial leucines with asparagines at certain positions strengthens these helix–helix interactions through hydrogen bond formation in bacterial membranes or detergent micelles [17,34–38]. Here, we show that most ultra-simple traptamers can cooperate with different receptors to induce cell transformation or growth factor independence in mammalian cells. Activity toward different receptors is dependent on the identity and the position of the amino acid substitution in polyleucine, and amino acid differences between TMDs of closely related target receptors cause marked differences in how receptors respond to these proteins. Furthermore, structurally diverse amino acids at the same position in a traptamer could confer the same activity, different patterns of activity were often observed for the same substitution at neighboring positions in the traptamer, and specific leucine-to-isoleucine substitutions in flanking sequences affected activity. Overall, although the precise rules that determine the specificity of TM interactions appear to be complex, our studies of ultra-simple TM proteins have revealed several fundamental principles that appear to govern some of these interactions.

Results

Polyleucine traptamers with different amino acid substitutions at position 13 activate the PDGFβR

Because expression of a 25-residue polyleucine TMD with a single isoleucine at the 13th position activated the PDGFβR, we determined whether expression of polyleucine with other single amino acid substitutions at position 13 displayed similar activity. We constructed 20 artificial genes, each encoding an N-terminal FLAG tag and a triple-glycine linker followed by 11 leucines, one of the 20 standard amino acids, and 13 additional leucines, for a total length of 37 residues (Fig. 1a). We numbered these proteins such that position "2" refers to the first leucine of the polyleucine stretch, so these mutants had substitutions at position 13. We introduced these genes into C127 cells by retroviral transduction and assessed their ability to induce the appearance of transformed foci, a marker of PDGFβR activation. In addition to isoleucine (I, Ile), which displayed the highest activity, methionine (M, Met), serine (S, Ser), threonine (T, Thr), glutamine (Q, Gln), or asparagine (N, Asn) at position 13 resulted in focus formation (Figs. 1b and 2a, column 1). We then immunoprecipitated the PDGFβR from extracts harvested from serum starved cells and immunoblotted for phosphotyrosine. As shown in Fig. 1c, there was a strong correlation between focus formation and PDGFβR phosphorylation, with the exception of lysine (K, Lys), where phosphorylation was observed in the absence of focus formation. We assume that polyLeu-K13 does not induce phosphorylation of the necessary PDGFβR tyrosines or generate the proper overall PDGFβR dimer conformation to activate downstream signaling responsible for cell transformation. Both the slower migrating mature form of the PDGFβR and the faster migrating precursor form with immature carbohydrates were phosphorylated in response to several of the active traptamers, suggesting that traptamer-induced signaling occurs at both the plasma membrane and in internal membranes. These results demonstrate that

several chemically diverse amino acids can confer biological activity when inserted into the middle of a polyleucine traptamer.

To confirm that the biological activity of the traptamers was dependent on the PDGFβR, we tested the ability of the active traptamers to confer growth factor independence in murine BaF3 cells, which lack endogenous PDGFβR and have a strict interleukin-3 (IL-3) requirement for growth and survival. Removing IL-3 from medium leads to rapid cell death, but BaF3 cells expressing PDGFβR can proliferate in the absence of IL-3 if soluble PDGF is added to the culture medium or if a small TM protein is expressed that activates the PDGFβR (Fig. S1) [39]. We individually expressed each of the six active polyleucine traptamers in parental BaF3 cells lacking the PDGFβR and in BaF3 cells harboring wild-type PDGFβR. We then removed IL-3 and assessed IL-3-independent growth. All six traptamers conferred growth factor independence in cells expressing the wild-type PDGFβR but not in parental BaF3 cells (Figs. S2a and b). These results confirm the necessity of the PDGFβR for the activity of these ultra-simple proteins. In contrast, the traptamers did not cooperate with chimeras of the PDGFβR with its TMD replaced by the TMD of the related PDGF α receptor (βαβ, which shares 52% sequence identity to PDGFβR TMD) or the TMD of the c-Kit receptor (another single-span receptor tyrosine kinase [RTK] with an unrelated TMD), neither of which responds to the BPV E5 protein) (Fig. S2b). These results suggest the importance of the receptor TMD in traptamer action, a topic we consider in great detail with the EPOR, as described in later sections.

There are few common characteristics among the residues that activated the PDGFβR. The long hydrophilic side-chains of glutamine and asparagine share similar chemical properties including the ability to form hydrogen bonds but are very different from the hydrophobic side-chains of isoleucine and methionine, which must rely on specific packing interactions. Similarly, serine and threonine have relatively small polar side-chains with a hydroxyl group, while the side-chains of the other active residues are much larger. These differences suggest that these proteins may use different mechanisms to activate the PDGFβR.

We next determined whether these simple proteins activated other RTKs. In our previous studies, we found that polyLeu-I13 activated the PDGFβR and not a panel of other RTKs [33]. Here, to determine the specificity of traptamers that induced transformation of human foreskin fibroblasts (HFFs), which express a wide repertoire of RTKs [33], we used phospho-RTK arrays to simultaneously interrogate tyrosine phosphorylation of 49 different RTKs. As shown in Fig. 1d, polyLeu-M13, -Q13, and -T13 induced tyrosine phosphorylation of the PDGFβR compared to control cells, with little to no phosphorylation of the other tested RTKs above background (Fig. 1d). Thus, these traptamers appear to be specific for the PDGFβR. We observed low levels of phosphorylation of the PDGFβR in cells expressing polyleucine lacking a substituted residue compared with cells expressing vector only, indicating that polyleucine induces a level of PDGFβR tyrosine phosphorylation that is below the threshold necessary to lead to transformation in C127 cells.

Polyleucine traptamers with different amino acid substitutions at position 13 activate the human and mouse EPORs

To determine whether these simple proteins activated targets other than PDGFβR, we tested their ability to cooperate with the human or mouse erythropoietin receptor (EPOR). We previously isolated and characterized more complex traptamers that activate the human EPOR (hEPOR) [32,40,41]. The EPOR is a cytokine receptor that lacks intrinsic kinase activity and signals through the Jak/STAT pathway. We individually expressed the 20 polyleucine traptamers with a substitution at position 13 in BaF3 expressing the hEPOR and tested the ability of the cells to grow in the absence of IL-3. Three traptamers (polyLeu-E13, -G13, and -Q13) conferred growth factor independence (Fig. 2a, column 2; Fig. S3). PolyLeu-E13 and -G13 did not cooperate with the PDGFβR, while polyLeu-Q13 cooperated with both PDGFβR and hEPOR. Isoleucine, methionine, serine, and threonine, which generated traptamers that activated the PDGFβR, were inactive with the hEPOR. The residues that allow hEPOR activation have chemically distinct side-chains. Thus, polyleucine traptamers with single substitutions at position 13 can activate unrelated receptors, and for the most part, different amino acids confer activity with different targets.

The TMDs of the PDGFβR and the hEPOR are not related, complicating analysis of their response to the traptamers. Therefore, we next compared the activity of the traptamers on two similar target receptors, namely, the human and mouse forms of the EPOR, which differ at only three residues within their TMD (Fig. 2b). To ensure that any differences observed were due to sequence differences in the TMDs and not to differences elsewhere in the receptors, we used a chimera of the hEPOR with its TMD replaced with that of the mouse EPOR (hmhEPOR). Strikingly, 13 of 20 traptamers with substitutions at position 13 cooperated with the hmhEPOR to confer growth factor independence, compared with only three that cooperated the hEPOR (Fig. 2a, columns 2 and 3). Thus, despite their high sequence similarity, the human and mouse EPOR TMDs respond to traptamers in markedly distinct fashions, with many more traptamers

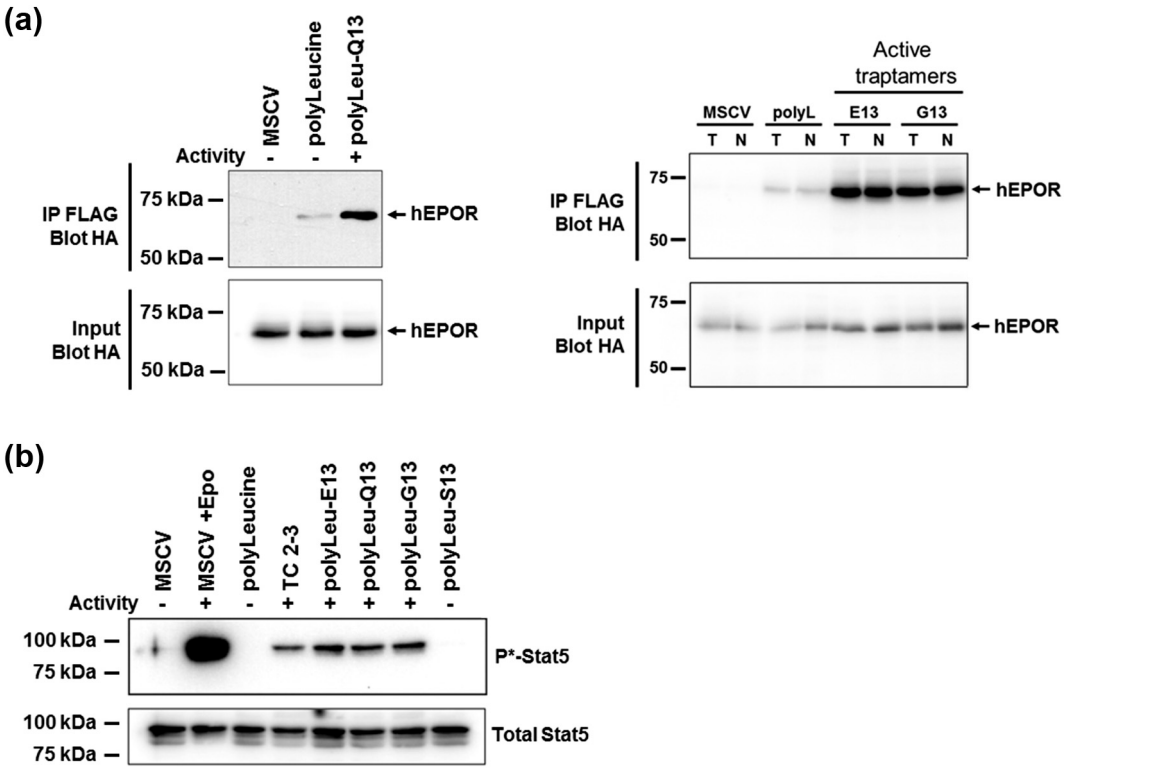

Fig. 3. Biochemical analysis of hEPOR activation by polyleucine traptamers. (a) Extracts were prepared from BaF3 cells expressing the hEPOR together with MSCV, polyleucine, or polyLeu with the indicated substitution at position 13. Lysates were subjected to SDS-PAGE and blotted for HA-tag on the hEPOR (input*).* Traptamers were immunoprecipitated using anti-FLAG magnetic beads and then electrophoresed and blotted as above (IP*).* Lysates were prepared in 1% Triton X-100 (left panel) or in 1% Triton X-100 (T) or in 1% NP40 (N) (right panel). (b) BaF3 cells expressing the hEPOR together with MSCV or different ultra-simple traptamers or positive control traptamer TC2–3 were starved of IL-3 for 3 h. MSCV plus EPO sample was then acutely treated with 0.6 units/mL EPO for 5 min. Extracts were prepared and proteins were separated by SDS-PAGE and blotted for phosphorylated STAT5 (top panel), stripped, and re-probed for total STAT5 (bottom panel).

showing activity with the hmhEPOR. The threshold for activating the hmhEPOR is not simply lower than that of the hEPOR, because polyLeu-G13 displayed strong activity for the hEPOR but not the hmhEPOR. Notably, the polyleucine backbone alone without a substituted amino acid conferred growth factor independence in cells expressing hmhEPOR.

Of the three amino acids that differ between the hEPOR and mEPOR TMDs, the difference in leucine and serine side-chains at position 238 is most dramatic. Therefore, we tested the ability of the traptamers to cooperate with a hEPOR mutant containing this single substitution, L238S, and found that the set of traptamers that cooperated with this mutant receptor was essentially identical to that which cooperated with hmhEPOR (Fig. 2a, columns 3 and 4), indicating that a single difference in the chemically complex EPOR TMD accounts for the different response of hEPOR and hmhEPOR to the traptamers.

The different activity displayed by the traptamers with the two forms of the EPOR implied that traptamer activity did not merely reflect traptamer expression levels. To verify this, we measured expression of the traptamers by staining cells with anti-FLAG antibody followed by flow cytometry. All traptamers were expressed above empty vector background, with no clear correlation between activity and expression levels (Fig. 2a, right panels). For example, although polyleucine traptamers with proline (P) or phenylalanine (F) at position 13 were highly expressed, they were inactive with all tested receptors. These results indicate that the relative activity of these traptamers did not merely reflect their expression levels. In control experiments, the hEPOR, hmhEPOR, and hEPOR L238S supported similar levels of IL-3 independent proliferation in response to soluble EPO (Fig. S4a). In addition, none of the 20 traptamers with substitutions at position 13 induced growth of parental BaF3 cells lacking PDGFβR or EPOR, indicating that traptamer activity requires receptor expression (Fig. S4b).

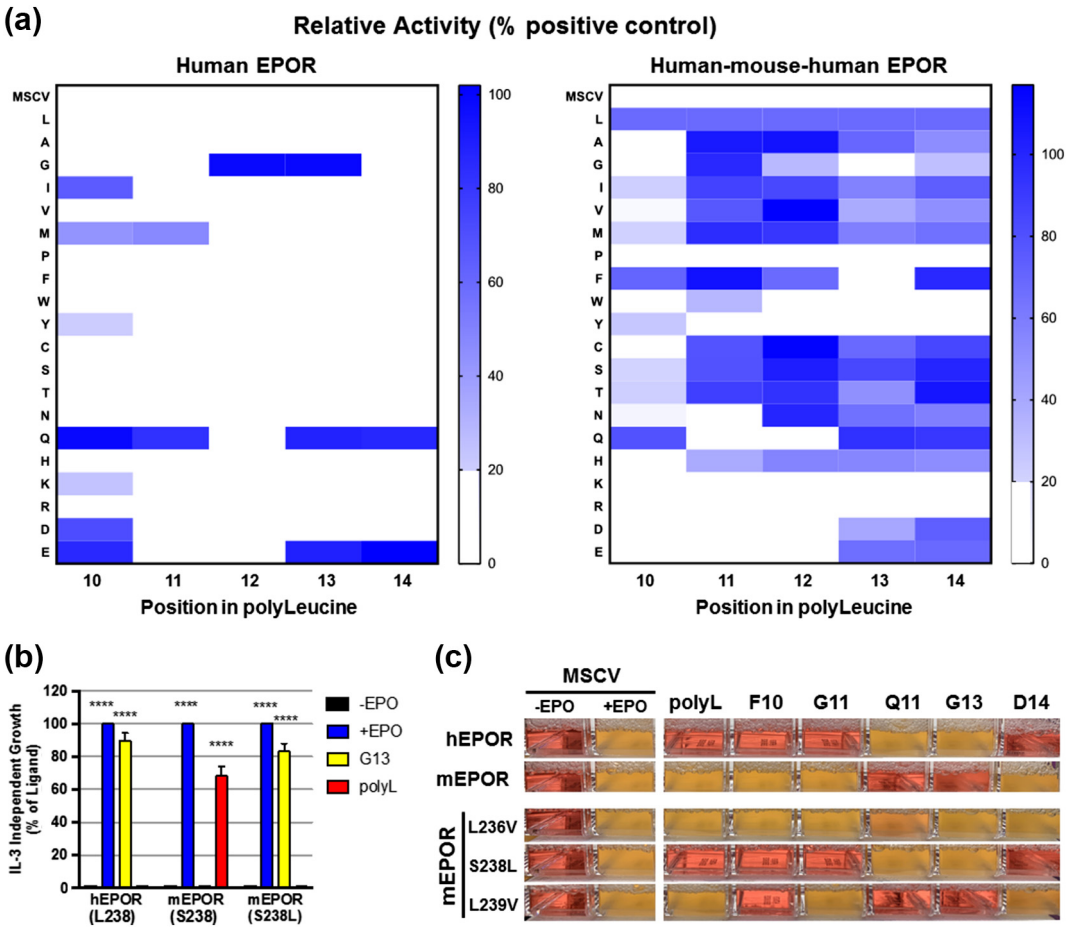

Fig. 4. Polyleucine traptamers with an amino acid substitution at several individual positions activate the EPOR. (a) Heat maps show ability of ultra-simple traptamers to confer growth factor independence in cells expressing the hEPOR (left) or hmhEPOR (right). Numbers on the bottom indicate the position of the substitution within the polyleucine stretch, and letters along the left indicate the amino acid present at the positions within the polyLeu TMD. All values are given as a percent of cells incubated in soluble EPO (0.06 units/mL). Darker values represent greater activity. Primary data provided in Fig. S3. An arbitrary cutoff was set at 20%. Activity was assessed by counting live cells 4 days after IL-3 removal. (b) IL-3 independence assay in BaF3 cells expressing hEPOR, mEPOR, or mEPOR S238L TM mutant together with MSCV, polyleucine or polyLeu-G13. MSCV plus 0.06 units/mL was used as positive control. Activity was assessed by counting live cells 4 days after IL-3 removal. Significance from ordinary one-way ANOVA Dunnett's multiple comparison test: ****$p < .0001$ for each sample compared to MSCV in the absence of EPO. (c) Representative IL-3 independence assay in BaF3 cells expressing hEPOR, mEPOR, or mEPOR TM mutant L236V, S238L, or L239V together with MSCV, polyleucine, polyLeu-G11, -Q11, -G13, -F10, or -D14. MSCV plus 0.06 units/mL EPO was used for positive control. Pictures of flasks show the relative growth after 8 days without IL-3. Color change from pink to yellow reflects the pH in dense cell cultures, illustrating traptamer activity.

Because polyleucine cooperated with the hmhEPOR, we determined how the number of leucines affected activity in the absence of amino acid substitutions. Polyleucine traptamers containing the FLAG tag and linker followed by 24 or 25 leucines were most active in cooperating with the hmhEPOR, with activity tapering off for fewer or more leucines (Fig. S5). None of these traptamers containing exclusively leucine cooperated with the hEPOR. The traptamers contained an N-terminal FLAG tag followed by a GGG linker (DYKDDDDKGGG). The FLAG tag was not sufficient for activity, as many of the constructs were not active. To determine if the FLAG tag was required for activity with the EPOR, we replaced it in several active traptamers with the GGG linker alone, an HA tag followed by the linker (YPYDVPDYAGGG), or the sequence DDDAGGG. GGG and YPYDVPDYAGGG did not support activity, whereas DDDAGGG behaved similarly to the FLAG tag (Fig. S6). These results suggest that the net negative charge on the tag in some way facilitates productive interaction with the EPOR. We expect that the traptamers are type II in orientation because it is

unlikely that the highly-charged N-terminal tag would pass through the membrane.

Biochemical analysis of the interactions between ultra-simple traptamers and hEPOR

To test whether substitutions that activated polyLeu also allowed complex formation with the hEPOR, we immunoprecipitated traptamers from detergent lysates with antibody recognizing the FLAG tag, followed by SDS-PAGE and blotting for HA tag on the EPOR. Some basal association was observed for polyleucine, but the signal was increased for the active clones polyLeu-Q13, -E13, and -G13 (Fig. 3a). We also assayed hEPOR signaling through its downstream second messenger, STAT5. STAT5 was heavily phosphorylated in cells acutely treated with EPO. In the absence of EPO, there was little to no signal for cells expressing empty vector (MSCV), polyleucine, or polyLeu-S13, none of which cooperate with the hEPOR. In contrast, all three active polyleucine traptamers (polyLeu-E13, -G13, and -Q13) induced phosphorylation of STAT5, as did TC 2–3, a traptamer with a chemically complex sequence that activates the hEPOR [40] (Fig. 3b). These results suggest that ultra-simple traptamers act *via* physical association with the hEPOR and activation of the canonical EPOR STAT5 signaling pathway, as was described previously with more complex traptamers [32,41]. While polyleucine associates with the receptor at low level, there is no STAT5 phosphorylation. Thus, activating substitutions in the ultra-simple traptamers appear to convert a basal, non-productive interaction to one that confers a signaling-competent state to the hEPOR to drive cell proliferation.

Single amino acid substitutions at several positions within polyleucine confer activity

To determine whether substitutions at positions other than 13 yielded activity, we tested the activity of polyleucine traptamers with single residue substitutions of all 20 standard amino acids at positions 10 to 14. Most substitutions yielded traptamers that cooperated with the hmhEPOR, hEPOR, or both (Fig. 4a).

Most strikingly, far more traptamers activated the hmhEPOR than the hEPOR, showing that the difference of three amino acids in the EPOR TMDs leads to pronounced differences in how the EPOR responds to traptamers and confirming that the hmhEPOR is much more prone to activation by simple TM proteins. However, a few traptamers such as polyLeu-Q11 or polyLeu-E10 displayed activity with the hEPOR but not hmhEPOR. Some traptamers, such as polyleucine with glutamine at several positions, were active in cells expressing either receptor, but not in parental BaF3 cells lacking EPOR, indicating that traptamer activity was not due to their ability to cooperate with a receptor in BaF3 cells other than the EPOR (Fig. S4c). Furthermore, because most traptamers cooperated with at least one EPOR, the differences in traptamer activity are unlikely due to their differential ability to traffic to and remain stably expressed within membranes. The different specificities displayed by the traptamers also suggest that EPOR activation is not the result of non-specific aggregation of traptamers.

Substitutions at the other positions did not recapitulate the patterns observed for position 13, although the patterns for positions 13 and 14 were similar, as were the patterns for positions 11 and 12. Similar residues did not necessarily behave the same way. For example, aspartic acid and glutamic acid at position 10 both conferred growth factor independence with the hEPOR, yet at positions 13 and 14, glutamic acid and not aspartic acid was active. Strikingly, the removal of a leucine side-chain (i.e., a glycine substitution) led to intriguing, position-dependent activity. A traptamer containing glycine at position 10 was not active with either receptor; at position 11, glycine allowed strong cooperation with the hmhEPOR but no activity with the hEPOR; at position 12, glycine led to activity with both EPORs; and glycine at position 13 conferred strong activity with the hEPOR and was inactive with the hmhEPOR. Even though all flanking residues were fixed as leucine, these unpredictable patterns of activity of traptamers with the same substitution at neighboring positions suggest that the rules that underlie these interactions and their biological outcomes are complex.

Because each substituted residue is flanked by multiple leucines, it seemed plausible that these simple proteins could adjust up or down within the membrane such that substitutions at neighboring positions would result in similar patterns of activity. Instead, as noted above for glycine, the same substitution at neighboring positions could have markedly different consequences. Most notably, there were well-defined “gaps” in activity: for example, polyLeu-Q12 and -F13 are inactive with the hEPOR and hmhEPOR, respectively, while these substitutions at neighboring positions conferred activity.

We used hierarchical clustering to analyze the patterns of activity of the traptamers with the hmhEPOR (Fig. S7). Traptamers with glutamine substitutions occupied their own branch of the tree. Other traptamers were divided further into sets that work well at most positions and those that work at one or two positions or not at all. For many cases, residues with similar side-chains clustered together. For example, aspartate and glutamate, asparagine and histidine, and serine and cysteine, pairs of amino acids with chemically similar side-chains, clustered together, as did threonine and isoleucine,

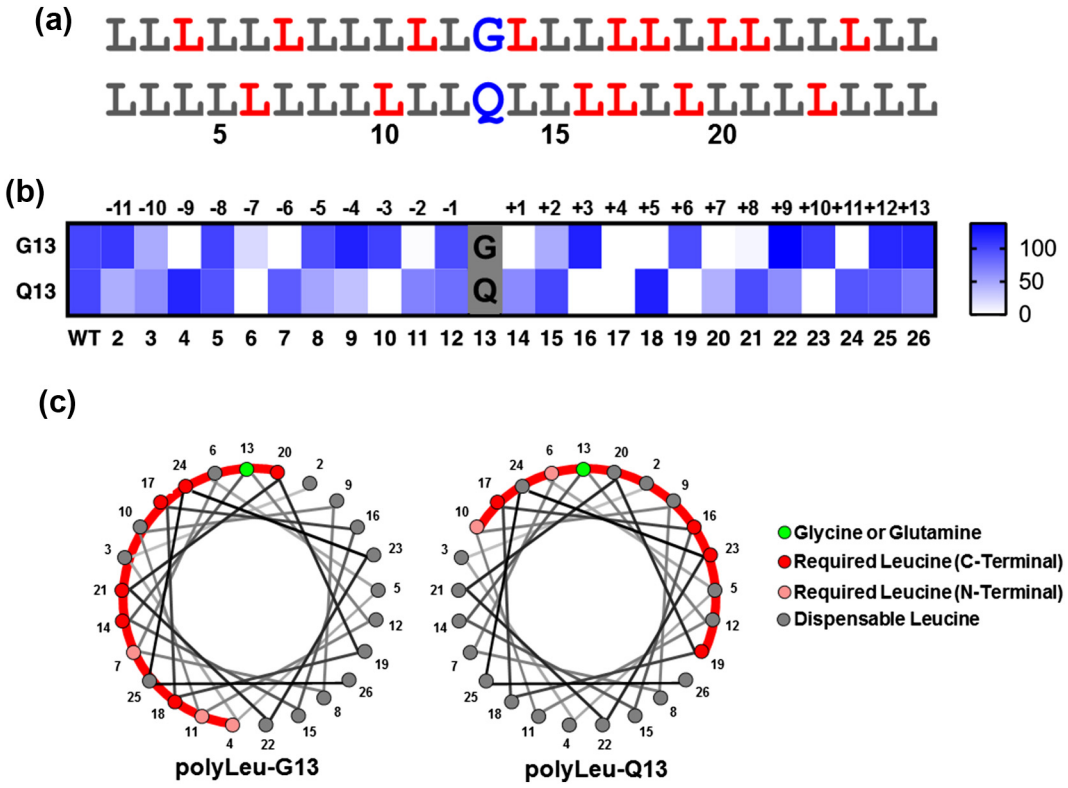

Fig. 5. Identification of flanking leucines required for activity with hEPOR. (a) Relative activity of polyLeu-G13 and -Q13 and mutants with individual leucine-to-isoleucine mutations at every position in the polyleucine stretch. Red leucines indicate positions where ability to induce IL-3 independence in BaF3/hEPOR cells was lost (<5% wild-type) for mutants containing a leucine-to-isoleucine mutation at that position. (b) The activity of wild-type traptamers is set at 100% in dark blue, and the activity of each mutant relative to wild-type is shown in shades of blue. White boxes represent a complete lack of activity. Numbers on bottom show absolute positions of leucine residues; numbers on top show positions relative to position 13. Activity was assessed by counting live cells 4 days after IL-3 removal in multiple independent experiments. (c) Helical wheel diagrams summarizing the results in panel a for a canonical alpha helix. Position 13 is in green, pink dots represent required leucines that are located upstream of the substituted amino acid, and red dots represent those that are downstream. Red arcs denote the regions of the helix where required leucines are located.

each of which contains a beta-branched side-chain methyl group. Interestingly, valine, another beta-branched amino acid whose side-chain is isosteric with threonine, clustered with alanine and not with threonine and isoleucine, highlighting the difficulty in predicting activity based on the structure of the amino acid side-chain.

Mapping differences in EPOR TMDs that confer specificity

The ability of polyleucine to activate hmhEPOR indicates that certain substitutions in polyleucine *inhibit* the productive interaction between polyleucine and this receptor. In contrast, polyleucine was inactive with the hEPOR, and a relatively limited set of substitutions in polyleucine allowed it to cooperate with the hEPOR. Most striking, the removal of a single leucine side-chain in polyLeu-G13 eliminated activity with the hmhEPOR and conferred activity with the hEPOR. We hypothesized that this difference was due to the packing interactions between the large amino acid Leu238 on the hEPOR and the small amino acid Gly13 on the traptamer, or between the small amino acid Ser238 on the hmhEPOR and Leu13 on the traptamer. To test this hypothesis, we determined the ability of polyleucine and polyLeu-G13 to cooperate with wild-type mEPOR or mEPOR S238L. As shown in Fig. 4b, polyleucine cooperated with the wild-type mEPOR (containing serine at position 238) but not S238L, whereas polyLeu-G13 cooperated with the mEPOR S238L mutant and not wild-type mEPOR. These results strongly suggest that specific packing interactions between position 238 in the EPOR TMD and position 13 of the traptamer are crucial for activity.

We next determined which sequence differences in the TMD of the EPOR dictated the activity of traptamers with substitutions at other positions. We examined a set of traptamers with

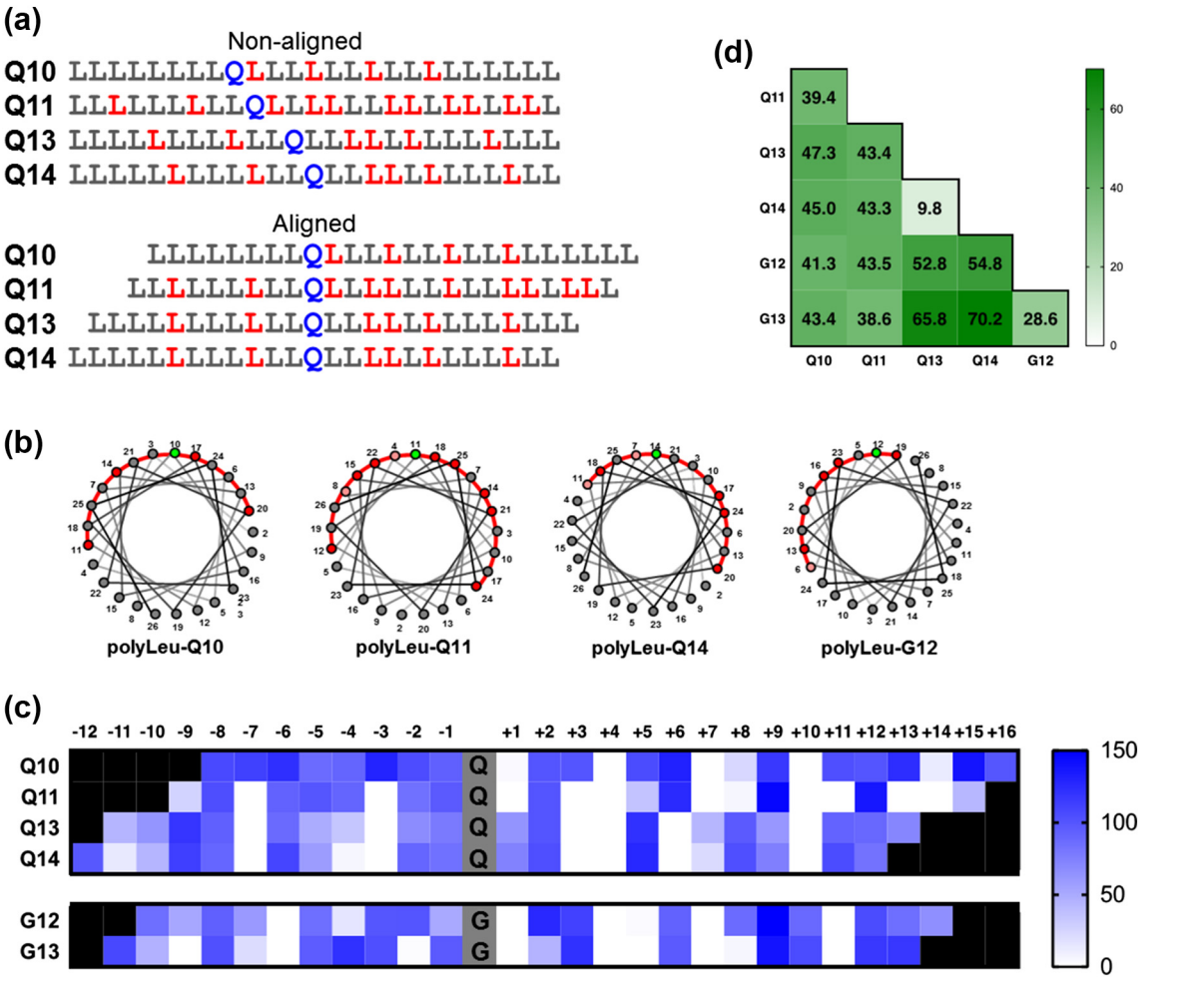

Fig. 6. Leucine-isoleucine mutagenesis of polyleucine traptamers with glutamine or glycine at neighboring positions. (a) Red leucines indicate positions where the ability to induce IL-3 independence in BaF3/hEPOR cells was lost, as in Fig. 5a. Bottom panel shows sequences aligned according to glutamine. (b) Helical wheel diagrams as in Fig. 5c. (c) Activity of the indicated polyLeu traptamers with individual leucine-to-isoleucine mutations, relative to the cognate wild-type traptamer, as in Fig. 5b. Sequences are aligned to the glutamine or glycine, with numbers denoting position relative to glutamine or glycine. (d) Table showing the similarity in activity profiles for each pair of traptamers aligned to the glutamine or glycine. The difference at each position between a mutant with wild-type activity and one with no activity is set as 100. Numbers show difference in the activity averaged over all positions for the indicated pair. Results are also color-coded in shades of green where the most different pairs are darker, with completely identical patterns set at white.

substitutions at various positions that cooperated with only the hEPOR or the hmhEPOR and tested their activity with mEPOR mutants containing individual mouse-to-human substitutions in the mEPOR TMD, L236V, S238L, and L239V (Fig. 4c). These traptamers displayed the same activity with the hEPOR and with mEPOR S238L, showing that the identity of a single key TM amino acid that differs between the two EPOR forms can define how these receptors respond to traptamers with diverse substitutions. However, the other mEPOR point mutants responded differently than either wild-type EPOR. The mEPOR L236V mutation cooperated with all tested traptamers to confer IL-3 independence, whereas the mEPOR L239V mutation failed to cooperate with polyLeu-F10, -Q11, or -G13. Thus, the S238L mutation has the predominant effect in determining which traptamers cooperate with which receptor, and the individual mutations in the receptor TMD do not have an additive effect in defining the differences between hEPOR and hmhEPOR. Taken together, these data show that traptamer activity is dependent on both the position of the substitution in the traptamer and the identity of the substituted amino acid. In addition, single amino acid

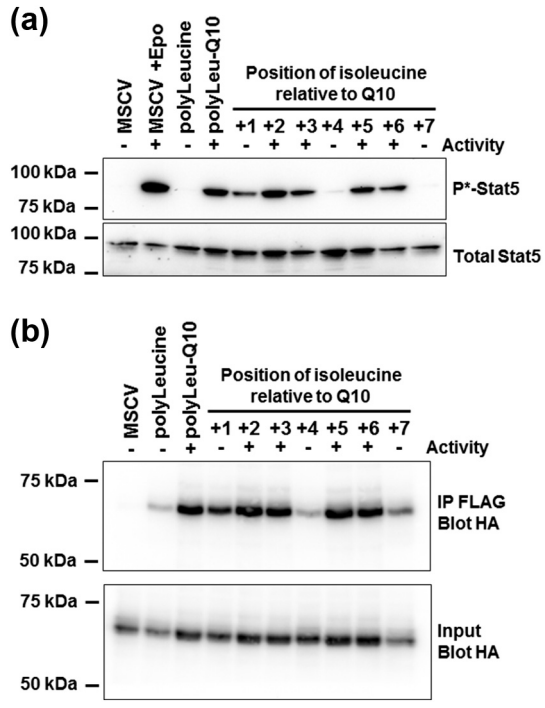

Fig. 7. Biochemical consequences of leucine-to-isoleucine mutations. (a) Cell extracts were prepared from BaF3 cells expressing the hEPOR together with MSCV, polyleucine, polyLeu-Q10, or polyLeu-Q10 with leucine-to-isoleucine mutations at positions +1 to +7 relative to position 10. Co-IP and western blotting was carried out as described in Fig. 3a. (b) Extracts were prepared from BaF3 cells expressing the hEPOR together with MSCV, polyleucine, polyLeu-Q10, or polyLeu-Q10 with leucine-to-isoleucine mutations at positions +1 to +7. STAT5 phosphorylation was determined as described as in Fig. 3b. Activity indicates whether or not conditions confer IL-3-independent growth.

differences between receptor TMDs can lead to drastic differences in their response to traptamers.

Scanning mutagenesis identifies required leucine residues flanking the substituted amino acid

To determine if the leucines flanking the substituted amino acid were important for activity, we conducted mutational analysis of representative traptamers. We first replaced each leucine (positions 2–12 and 14–26) in polyLeu-Q13 and -G13 with an isoleucine, one position at a time, and tested the ability of these mutants to confer growth factor independence in cells expressing hEPOR. Leucine and isoleucine differ by only the placement of a single side-chain methyl group. For both traptamers, isoleucine mutations at several positions markedly inhibited activity (Fig. 5a and b). Importantly, the required leucine positions were located along the lengths of both traptamers, suggesting that key interactions between traptamer and receptor take place throughout the depth of the membrane, or that a mutation from leucine-to-isoleucine at key positions affects the overall helical structure in a way that abolishes the effect of the central substituted amino acid. Interestingly, the L10I and L13Q mutations separately conferred activity (Fig. 4a), but the combination of these two substitutions was inactive. Thus, combining mutations in the same traptamer does not result in a simple additive effect.

For the most part, the required leucines were at different positions in the two traptamers, indicating that the required positions are dictated by the identity of the central substituted amino acid. For polyLeu-Q13 and -G13, the required leucines localize to discrete faces of the helix with little overlap, shown *via* helical wheel diagrams (Fig. 5c). Furthermore, Q13 lies near the middle of the required leucines in this representation, whereas G13 lies at one end of the required leucines. Taken together, these results provide further evidence that polyLeu-G13 and -Q13 do not activate the hEPOR in the same fashion. Rather, the specific arrangement of helices within the TM complexes differs between the two traptamers, reflecting the difference between glutamine and glycine, which introduce hydrogen bonding options and greater structural flexibility, respectively.

We used the same approach to identify required leucines for traptamers with glutamine or glycine substituted at neighboring positions. We examined polyleucine traptamers with glutamine at positions 10, 11, 13, or 14, and glycine at positions 12 or 13. We did not analyze polyLeu-Q12 because it is inactive. In all cases, the required leucines were spread along the length of the traptamer and were confined to a discrete helical face (Fig. 6a and b). Strikingly, although the absolute positions of the required leucines differed for polyLeu-Q13 and -Q14, most of the required leucines were at the same positions relative to the location of the glutamine, which is most clearly seen when the sequences are aligned to the glutamine (Fig. 6a and c). When similarly aligned, polyLeu-Q11 shared some required leucines with polyLeu-Q13 and -Q14 and others with -Q10, but the requirements for polyLeu-Q10 were distinct from polyLeu-Q13 and -Q14. This indicates that leucines are required at a fixed distance from the glutamine within a narrow range of glutamine positions, but when the glutamine is present at distal positions, different patterns emerge. We also note that none of the N-terminal flanking leucines are required for polyLeu-Q10, while two N-terminal leucines (−3 and −7 relative to the glutamine) were required for -Q11, -Q13, and -Q14. Similarly, for polyLeu-G12 and -G13, most required leucines were at the same positions relative to the location of the glycine (Fig. 6c).

To quantify the differences between the traptamers based on the locations of the required leucines,

we aligned each pair of traptamers to the glutamine or glycine and then determined the difference in activity caused by isoleucine mutations, averaged across all positions (Fig. 6d). As expected, polyLeu-Q13 and -Q14 were the most similar, followed by polyLeu-G12 and -G13. Thus, polyleucine traptamers with the same amino acid as its central substituted residue behaved more similarly than those with a different amino acid substituted, and for traptamers with a central glutamine, traptamers with substitutions at neighboring positions behaved more similarly than those with substitutions at more distal positions.

Effect of leucine-to-isoleucine mutations on the interactions between ultra-simple traptamers and EPOR

Finally, we used co-IP and Western blotting to determine how leucine-to-isoleucine mutations in polyLeu-Q10 affected STAT5 activation and physical association of traptamers with the hEPOR. In most cases, IL-3 independence correlated well with STAT5 phosphorylation (Fig. 7a), except for polyLeu-Q10I11, which did not confer IL-3 independence, but did cause minimal STAT5 phosphorylation. When assayed by co-IP, there was low basal association between the hEPOR and polyleucine, as noted in Fig. 3, and robust association with polyLeu-Q10 and with the polyLeu-Q10 mutants that induce growth factor independence (Fig. 7b). The inactive leucine-to-isoleucine polyLeu-Q10 mutants associate to varying degrees with the hEPOR, with polyLeu-Q10I11 displaying the strongest association, consistent with its ability to induce STAT5 phosphorylation. Finally, there is no obvious correlation between activity and expression levels of polyLeu-G13 and several of its isoleucine mutants (Fig. S8). For example, polyLeu-G13I7, -G13I14, -G13I17, and -G13I18 are expressed well as assessed by flow cytometry but inactive. These results suggest that the leucine-to-isoleucine changes are directly affecting the ability of each traptamer to activate the receptor, independent of their effect on traptamer expression.

Discussion

Protein–protein interactions are most often studied in a sequence context established over millions of years of evolution. Thus, these studies unavoidably examine only a miniscule fraction of potential interacting sequences. We decided to simplify the system in order to interrogate the sequence basis for productive TMD interactions. TMD interactions are often mediated by simple helical configurations confined to short sequences of contiguous amino acids, eliminating complications due to complex protein folding and long-range interactions. To further simplify the system, we standardized the sequence context by analyzing proteins consisting of polyleucine with only one or two amino acid substitutions. We examined the biological consequence of all possible substitutions at each tested position to systematically explore how sequence dictates biological activity. These experiments identified several principles that appear to govern the association between these ultra-simple TMDs, which are masked by the sequence complexity of naturally occurring TMDs.

Remarkably, most ultra-simple short TM proteins consisting of homopolymeric leucine linked to an epitope tag displayed biological activity in mammalian cells, even though these proteins were tested for their ability to productively interact with only three targets. Analysis of the effect of amino acid substitutions in more than 200 ultra-simple TM proteins provided new insights into the sequence basis for TM protein activity. First, dissimilar amino acids at the same position in a simple protein can activate the same target, implying that these ultra-simple proteins use a variety of mechanisms to activate their target receptors, dictated by the different chemical and structural properties of single amino acids. The activity of some ultra-simple traptamers containing solely hydrophobic amino acids in the TMD provides further evidence that packing interactions are sufficient to confer activity, consistent with prior studies [3,42]. Second, the effect of the same substitutions at neighboring positions in TMDs can differ, even though the substitutions are in essentially identical polyleucine sequence contexts. Thus, the precise locations of amino acids within membranes determine their roles in protein–protein interactions, and activity cannot be simply explained by adjusting the protein segment "up" or "down" in the membrane. In addition, leucine-to-isoleucine mutations along the length of the polyleucine TMD can affect activity, and the identity of the amino acid at single centrally located positions dictates which positions can tolerate mutations. Flanking sequences can also affect the stability or specificity of interactions established by the GxxxG motif in more complex TMDs [43,44]. These observations imply that even in this simplified system, interactions are not strictly local but rather are distributed on discrete helical faces along the length of the TMD, suggesting that packing interactions throughout the TMD are important for activity even in ultra-simple proteins containing a single strongly polar amino acid that confers activity. Finally, the activity of a polyleucine traptamer with two mutations cannot be predicted simply by the effect of the constituent single mutations.

Another striking finding was the markedly different response of the human and mouse forms of the EPOR to the ultra-simple traptamers. These different

responses are a consequence of sequence differences in the TMDs of the receptors. Far more proteins activate the mEPOR, primarily due to a single serine residue in its TMD, which is also a crucial determinant for spleen focus-forming virus gp55-P to activate the mEPOR [45]. Several single amino acid substitutions at different positions reduce or abolish the activity of polyleucine toward the hmhEPOR, whereas substitution of certain residues at several positions in polyleucine leads to cooperation with the hEPOR. Some of the substitutions that imparted activity toward the hEPOR abolished activity toward the hmhEPOR (e.g., glycine at position 13). As was the case for mutations in the ultra-simple traptamers, the effects of substitutions in the more complex target TMD are not simply additive.

Overall, our results show that amino acid substitutions in interacting TMDs often have unpredictable effects on the ability of these domains to functionally interact, even in this very simple system where the effect of each amino acid at different positions can be compared within the same leucine context. Thus, it is not surprising that it has been so difficult to determine universally applicable rules that specify TMD interactions by studying naturally-occurring, far more complex TMDs.

Our results also challenge the notion of "conservative" mutations in TMDs. A leucine-to-isoleucine mutation in the PDGFβR TMD can block the interaction with the BPV E5 protein, and Deber and Stone present biophysical evidence that leucine and isoleucine in TMDs are not equivalent, in part due to differential interactions with lipids, consistent with earlier mutational analysis of the glycophorinA TMD [46–48]. However, these prior studies do not indicate how often such mutations cause a phenotype. In a systematic analysis of LIL traptamers, we showed that many single leucine/isoleucine mutations affected traptamer activity or specificity [32,33]. Here, we showed that even in ultra-simple polyleucine traptamers, single leucine, isoleucine, and valine substitutions act differently. This is most notable in the isoleucine scanning mutagenesis of six different ultra-simple traptamers, which revealed that the great majority of mutants displayed a measurable phenotype, implying that the differences in the shape or rotameric freedom of leucine and isoleucine side-chains can often affect TMD interactions.

The unusual chemistry of these ultra-simple traptamers undoubtedly influences their activity. Like the fatty acid tails in membrane lipids, the amino acid side-chains of polyleucine proteins are composed almost exclusively of hydrocarbons, although the structure, size, and dynamics of lipids obviously differ greatly from hydrophobic α-helical peptides. It is possible that the active ultra-simple traptamers displace the lipids that associate with receptor TMDs in such a fashion that stabilizes the receptor in an active oligomeric state or conformation. In addition, multiple traptamer molecules may cluster around each receptor, and/or some may form highly transient or dynamic complexes with their target TMDs, instead of forming a specific, stable, well-defined TM complex, as E5 does with the PDGFβR [22]. However, because many traptamers cooperate with only one receptor or another, they do not appear to act by non-specific aggregation- or localization-based mechanisms. The diverse substitutions that activate ultra-simple traptamers suggest that the stoichiometry, oligomeric state, stability, dynamics, or precise physical arrangement of TM helices may differ for different traptamers. Further biochemical and biophysical studies are required to understand the structural basis for these interesting findings.

Our results also have implications for the evolution of TMDs. Evolution has presumably weeded-out most TMDs that interact with many other TMDs. However, roughly two-thirds of polyleucine proteins with single substitutions showed some level of activity with one or more of the tested targets, implying that such simple proteins are able to interact promiscuously with many different TMDs, a perilous situation for maintaining order in the cell. When an additional isoleucine substitution was introduced into an active polyleucine protein, the number of proteins that activated the hEPOR fell by roughly one-third. On the other hand, we estimate that only 1 in ~10–20,000 more complex LIL traptamers consisting of random sequences of leucine and isoleucine displayed activity when tested with the PDGF β receptor or hEPOR [32,33]. This general relationship between the amino acid sequence complexity of a TMD and its likelihood of cooperating with a specific target suggests that more complex TMDs present in naturally occurring proteins likely exhibit activity toward a very restricted set of targets. Therefore, extant TMDs may have evolved to their current complex sequences in part to limit potential off-target interactions, or, in some cases, to prevent binding to other TMDs altogether.

Materials and Methods

Construction of ultra-simple traptamers

DNA encoding the ultra-simple traptamers was obtained from Integrated DNA Technologies (IDT) in the form of double-stranded DNA GeneBlocks. Oligonucleotides were designed with several bases at the 5′ end followed by an Xho1 restriction site, a methionine start codon, followed by codons encoding a FLAG epitope tag, a triple-glycine linker, then

25 leucines (or with single amino acid codon substitutions at single positions, or at one position with a leucine-to-isoleucine codon at a second position), 2 tandem stop codons, an EcoR1 restriction site, and several more random bases at the 3′ end (shown in Fig. 1a). In all cases, codon optimization was performed using IDT's online codon optimization tool. Ultra-simple traptamers with different tags or of different lengths were designed and ordered in an analogous fashion. Double-stranded DNA was reconstituted with water to a final concentration of 10 ng/μL and incubated at 50 °C for 20 min following IDT protocols. DNA (100 ng) was digested with 1 μL each of Xho1 and EcoR1 and 2 μL Cutsmart buffer (New England Biolabs) in 20 μL for 37 °C for 3 h, and then purified (QIAquick MinElute PCR purification Kit). Five microliters of purified digest was incubated with 150 ng of pMSCVpuro digested with EcoR1 and Xho1, 1 μL T4 DNA ligase (New England Biolabs) at either 15 °C overnight or room temperature for 1 h. One microliter of the ligation mixture was used to transform chemically competent DH10β *Escherichia coli* (Invitrogen) and the entire transformation mixture was plated on LB plates containing ampicillin, which were incubated overnight at 37 °C. Individual colonies were picked and used to inoculate cultures in LB containing 100 mg/mL ampicillin. Plasmid DNA was extracted with a MiniPrep kit (Machery-Nagel) and sequenced.

Vectors, cell lines, and culture conditions

C127 mouse fibroblasts, human foreskin fibroblasts (HFFs), and HEK293T cells were maintained in Dulbecco's modified Eagle medium (DMEM) with 10% fetal bovine serum (FBS), 2% Hepes buffer, and 1% penicillin/streptomycin at 37 °C. BaF3 Cells were maintained in RPMI 1640 with 10% heat inactivated FBS, WEHI-3B cell-conditioned medium (source of IL-3), 1% penicillin/streptomycin, 0.06 mM β-mercaptoethanol at 37 °C, 4 mM L-glutamine, and 0.06 mM β-mercaptoethanol.

The PDGFβR, βαβ PDGFR chimera, and c-kit genes were previously described [33] and expressed from the retroviral vector LXSN (Mouse Moloney murine sarcoma virus) carrying a neomycin resistance marker. HA-tagged human and mouse EPORs and all mutant forms of both receptors expressed from pMSCVneo (Clontech) were previously described [32].

Retroviruses were produced in 293T cells. Briefly, 2–2.5 × 10^6 HEK293T cells were plated in 10-cm dishes in 6 mL DMEM. After 24 h, medium was replaced with 5 mL fresh DMEM. For each transfection, 6 μg pCL-Eco (Imgenex) [49], 4 μg pVSV-g (Clontech) [50], and 10 μg of the retroviral expression plasmid of interest were added to 250 μL Optimem medium (ThermoFisher). Ten microliters of Lipofectamine (ThermoFisher) was added to a separate aliquot of 250 μL Optimem. After 5 min at 22 °C, Optimem/Lipofectamine mixture was added to the Optimem/DNA mixture and incubated for 30 min at 22 °C. Five hundred microliters of combined mixture was then added drop-wise to each 10-cm dish of HEK293T cells. Cells were incubated for 2–3 days at 37 °C, and the supernatant was harvested, filtered through a 0.45-μm filter (Millipore, Danvers, MA), and either used immediately or aliquoted and stored at −80 °C.

BaF3 cell lines expressing various receptors were generated *via* retroviral transduction. Briefly, 500,000 BaF3 cells were infected with 1 mL of retrovirus expressing the receptor of interest (or 1 mL DMEM for mock infection) in 5 mL complete RPMI-1640 medium containing 4 μg/mL polybrene final concentration. After 24 h at 37 °C, 5 mL complete RPMI-1640 medium containing 2 mg/mL G418 was added for a final G418 concentration of 1 mg/mL in 10 mL. Cells were cultivated in drug until mock-infected cells died (5–6 days). To express traptamers, 500,000 BaF3 cells expressing various receptors were retrovirally infected as above. After 24 h, 5 mL complete RPMI-1640 medium containing 2 mg/mL G418 and 2 μg/mL puromycin was added for a final concentration of 1 mg/mL G418 and 1 μg/mL puromycin in 10 mL total. Cells were cultured until mock-infected cells died (3–4 days).

Focus-forming assays

Six-centimeter dishes of C127 cells at ~70% confluence in DMEM-10 were infected with 3 μL of a retrovirus and 4 μg/mL polybrene. After 24 h, the cells were divided evenly into two new 6-cm dishes and incubated at confluence for 3 weeks with medium changes every 2–3 days. Foci were visualized by fixation in methanol followed by staining with a 5% Giemsa solution (Sigma-Aldrich). To determine viral titer, 1% of the infected cells were plated on 10-cm dishes and incubated in DMEM-10 medium containing 1.25 μg/mL puromycin. After mock-infected cells died (~7 days), colonies of puro resistant virus-infected cells were stained with Giemsa as above and counted. The number of transformed foci was normalized to virus titer.

IL-3 independence assays

For IL-3 independence assays, 5 × 10^5 BaF3 cells expressing various traptamer and receptor combinations were washed in PBS three times to remove IL-3. Cell pellets were resuspended in 10 mL RPMI-10 medium lacking IL-3. Live cells were counted at various days after IL-3 removal. Unless otherwise noted, relative cell counts

presented were determined at day 4 after IL-3 removal. Cells expressing a receptor and empty MSCVpuro were used as negative control. Cells expressing MSCVpuro were treated with 0.06 units/mL recombinant human EPO or 100 ng/mL human SCF (for c-kit) as positive controls. For PDGFβR of βαβ PDGFR cells, MSCVpuro expressing v-sis (viral homolog of soluble PDGF) were used as positive controls. For clarity, in all graphs, values for samples with no growth were arbitrarily set at 1%. To determine the similarities between the mutational profiles of pairs of traptamers containing single leucine-to-isoleucine mutations, each pair was aligned to the central glutamine or glycine. Then, at each position present in both traptamers, the absolute value of the difference in activity caused by a leucine-to-isoleucine substitution at that position was determined (with 100 representing the difference between the activity of a wild-type traptamer and a totally defective mutant). The average difference across all assessable positions for each pair is shown in Fig. 6d.

Immunoprecipitation and immunoblotting

C127 cells or HFFs stably expressing a traptamer or empty MSCVpuro were established by infecting cells in 6-cm dishes with 1 mL of retrovirus in 2 mL of medium containing 4 μg/mL polybrene. After 24 h, cells were passaged into 10-cm dishes and then selected in medium containing 1 μg/mL puromycin until the mock-infected cells died. After passage, cells at 80%–90% confluence were serum-starved overnight, washed twice with PBS, and lysed in 500 μL of cold FLAG lysis buffer [50 mM Tris–HCl (pH 7.4); 150 mM NaCl; 1% Triton X-100; 1 mM EDTA]. All lysis buffers contained Halt protease and phosphatase inhibitors (Thermo Scientific) and 1 mM phenylmethylsulfonyl fluoride. Sodium metavanadate (0.5 mM) was added to the buffer for phospho-tyrosine analysis. All lysates were incubated on ice for 20 min, followed by centrifugation at 14,000 rpm for 30 min at 4 °C. Total protein concentration of the supernatants was determined using a bicinchoninic acid (BCA) protein assay kit (ThermoFisher).

For immunoprecipitation of the PDGFβR, lysates containing equal amounts of extracted proteins (typically 400 μg) were incubated overnight at 4 °C with 5 μL anti-PDGFβR rabbit antiserum per milligram extracted protein, followed by addition of 50 μL of protein A Sepharose beads, during rotating incubation for 1 h at 4 °C. For immunoprecipitation of FLAG-tagged traptamers, extracts were incubated for 6 h at 4 °C with 60 μL of the anti-FLAG M2 magnetic beads (Sigma-Aldrich, M8823) per sample according to the manufacturer's instructions. For PDGFβR, immunoprecipitates were washed three times in NETN buffer [100 mM NaCl, 20 mM Tris·HCl (pH 8.0), 0.5 mM EDTA, 0.5% Nonidet P-40] supplemented with 1 mM PMSF and 0.5 mM sodium metavanadate, resuspended in 2× Laemmli sample buffer, boiled, and electrophoresed for 1.5–2 h at 150 V on a 7.5%, 10% or 20% (wt/vol) polyacrylamide/SDS gel.

For immunoprecipitation of FLAG-tagged traptamers, pellets of BaF3 cells expressing traptamers and EPOR constructs were lysed in ~500 μL FLAG lysis buffer (as above). Lysates were incubated with ~60 μL Anti-FLAG magnetic beads for ~6 h at 4 °C. Beads were washed 3× in ice-cold TBST then resuspended in 2× Laemmli sample buffer. Samples were incubated for 5 min at 80 °C prior to polyacrylamide gel electrophoresis.

Samples were transferred for 1 h at 100 V to 0.45 μm nitrocellulose (to detect total or phosphorylated PDGFβR or hEPOR) in Tris/glycine transfer buffer [25 mM Tris, 192 mM glycine, and 20% (vol/vol) methanol]. Transfer buffer supplemented with SDS to 0.1% was used for anti-PDGFβR immunoblotting. 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 2 h at 22 °C and incubated overnight at 4 °C with antiphosphotyrosine antibody (PY1000; Cell Signaling 8954S), horseradish peroxidase (HRP)-conjugated anti-HA antibody (Cell Signaling Technology 2999S), which were diluted 1:1400, 1:500, and 1:1000, respectively, in 5% (wt/vol) milk/TBST. Blots were washed 5× in TBST and incubated for 1 h at room temperature with 1:10,000 in 5% (wt/vol) milk/TBST HRP-conjugated donkey anti-mouse antibody (Jackson ImmunoResearch) for antiphosphotyrosine blots or 1:7000 HRP-protein A (GE Healthcare; for anti-PDGF receptor blots). Blots were washed and visualized by enhanced chemiluminescence (SuperSignal West Pico Chemiluminescent Substrate, Thermo Scientific). HRP-conjugated anti-HA (Cell Signaling) was used at 1:1000 for EPOR detection after overnight incubation similarly and detected after 5× washes in TBST.

For the phospho-RTK arrays, HFFs stably expressing MSCVpuro, or polyleucine, polyLeu-I13, -M13, -Q13, or -T13 were grown to confluence in 100 mm dishes and serum-starved overnight. All steps thereafter were performed as described in the Human Phospho-RTK Array Kit (Proteome Profiler, R&D Systems). Briefly, the cells were lysed in the provided lysis buffer supplemented with HALT protease inhibitors, 1 mM PMSF, and 0.5 mM sodium metavanadate. Lysates were then incubated with the phospho-RTK array membranes overnight at 4 °C. After washing, the membranes were incubated with the provided anti-phospho-tyrosine-HRP detection antibody for 2 h, washed, and then visualized with the provided Chemiluminescence Reagent Mix.

For STAT5 phosphorylation, cells expressing various traptamers and EPOR constructs were

grown to dense 10 mL cultures. Cells were pelleted on ice in the presence of 15 μL Halt Phosphatase inhibitor (ThermoFisher) and 75 μL sodium metavanadate for 10 min at 1500 RPM at 4 °C. Cell extracts were prepared and electrophoresed as described above using 20–30 μg of total protein. After transfer to 0.2 μm nitrocellulose membranes and blocking in 5% milk in TBST, blots were incubated overnight 4 °C with 1:1000 Anti-PhosphoSTAT5 (Cell Signaling Technology 9351S) in 5% milk in TBST. Blots were washed, incubated with secondary antibody 1:8000, washed again, and visualized using enhanced chemiluminescence. Blots were then stripped with stripping buffer (ThermoFisher), blocked, and re-probed with 1:1000 Anti-Total STAT5 (Cell Signaling Technology 94,205). Quantification of relative band intensity was carried out using ImageJ software.

Flow cytometry

Approximately 1×10^6 BaF3 cells expressing various traptamers were washed 3× in PBS and fixed and permeabilized with the Cytofix/Cytoperm kit (BD Biosciences) following the manufacturer's protocol. Briefly, cell pellets were resuspended in 500 μL Cytofix/Cytoperm buffer and incubated for 30 min at 4 °C. Cells were washed 2× in Perm/Wash buffer and resuspended in Perm/Wash buffer containing 2 μg/mL APC conjugated Anti-FLAG (PerkinElmer AD0059F). Cells were then washed 2× with Perm/Wash buffer and resuspended in 300 μL PBS. Samples were analyzed on a Stratedigm S1000EX and visualized using FlowJo.

Acknowledgments

We thank Don Engelman, Aaron Ring, Lisa Petti, and Anne Edwards for helpful discussions; Sophia Chen, Aldo Salazar Morales, Claire Victoria Ong, and Rebecca Arteaga for their assistance in cloning constructs; Ashish Shelar for help with graphics; and Jan Zulkeski for assistance in preparing this manuscript. This work was supported by a grant (CA037157) from the National Cancer Institute to D.D.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2019.07.009.

Received 23 May 2019;
Received in revised form 2 July 2019;
Accepted 3 July 2019
Available online 10 July 2019

Keywords:
traptamer;
BPV E5 protein;
BaF3 cells;
erythropoietin receptor;
helix

†These authors contributed equally to this work.

Abbreviations used:
TM, transmembrane; TMD, transmembrane domain; EPOR, erythropoietin receptor; hEPOR, human erythropoietin receptor; mEPOR, membrane erythropoietin receptor; PDGFβR, platelet-derived growth factor beta receptor; IL-3, interleukin-3; BPV, bovine papillomavirus.

References

[1] D.T. Moore, B.W. Berger, W.F. DeGrado, Protein–protein interactions in the membrane: sequence, structural, and biological motifs, Structure. 16 (2008) 991–1001.
[2] T.H. Walther, A.S. Ulrich, Transmembrane helix assembly and the role of salt bridges, Curr. Opin. Struct. Biol. 27 (2014) 63–68.
[3] M. Mravic, J.L. Thomaston, M. Tucker, P.E. Solomon, L. Liu, W.F. DeGrado, Packing of apolar side chains enables accurate design of highly stable membrane proteins, Science. 363 (2019) 1418–1423.
[4] D. Langosch, I.T. Arkin, Interaction and conformational dynamics of membrane-spanning protein helices, Protein Sci. 18 (2009) 1343–1358.
[5] K. Bugge, K. Lindorff-Larsen, B.B. Kragelund, Understanding single-pass transmembrane receptor signaling from a structural viewpoint—what are we missing? FEBS J. 283 (2016) 4424–4451.
[6] K. Gupta, J.A.C. Donlan, J.T.S. Hopper, P. Uzdavinys, M. Landreh, W.B. Struwe, et al., The role of interfacial lipids in stabilizing membrane protein oligomers, Nature. 541 (2017) 421–424.
[7] E. Henrich, O. Peetz, C. Hein, A. Laguerre, B. Hoffmann, J. Hoffmann, et al., Analyzing native membrane protein assembly in nanodiscs by combined non-covalent mass spectrometry and synthetic biology, eLife. 6 (2017).
[8] H. Naveed, R. Jackups Jr., J. Liang, Predicting weakly stable regions, oligomerization state, and protein–protein interfaces in transmembrane domains of outer membrane proteins, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 12735–12740.
[9] F. Cymer, A. Veerappan, D. Schneider, Transmembrane helix–helix interactions are modulated by the sequence context and by lipid bilayer properties, Biochim. Biophys. Acta 2012 (1818) 963–973.
[10] G.R. Dieckmann, W.F. DeGrado, Modeling transmembrane helical oligomers, Curr. Opin. Struct. Biol. 7 (1997) 486–494.
[11] K.R. MacKenzie, D.M. Engelman, Structure-based prediction of the stability of transmembrane helix–helix interactions: the sequence dependence of glycophorin A dimerization, Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 3583–3590.
[12] W.P. Russ, D.M. Engelman, The GxxxG motif: a framework for transmembrane helix–helix association, J. Mol. Biol. 296 (2000) 911–919.
[13] M.E. Call, J. Pyrdol, M. Wiedmann, K.W. Wucherpfennig, The organizing principle in the formation of the T cell receptor–CD3 complex, Cell. 111 (2002) 967–979.

[14] P. Cosson, S.P. Lankford, J.S. Bonifacino, R.D. Klausner, Membrane protein association by potential intramembrane charge pairs, Nature. 351 (1991) 414–416.
[15] B.H. Honig, W.L. Hubbell, Stability of "salt bridges" in membrane proteins, Proc. Natl. Acad. Sci. U. S. A. 81 (1984) 5412–5416.
[16] M.E. Call, K.W. Wucherpfennig, J.J. Chou, The structural basis for intramembrane assembly of an activating immunoreceptor complex, Nat. Immunol. 11 (2010) 1023–1029.
[17] E. Li, W.C. Wimley, K. Hristova, Transmembrane helix dimerization: beyond the search for sequence motifs, Biochim. Biophys. Acta 2012 (1818) 183–193.
[18] E. Wallin, G. Von Heijne, Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms, Protein Sci. 7 (1998) 1029–1038.
[19] L. Petti, D. DiMaio, 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 (1994) 3582–3592.
[20] A. Burkhardt, D. DiMaio, R. Schlegel, Genetic and biochemical definition of the bovine papillomavirus E5 transforming protein, EMBO J. 6 (1987) 2381–2385.
[21] R. Schlegel, M. Wade-Glass, M.S. Rabson, Y.-C. Yang, The E5 transforming gene of bovine papillomavirus encodes a small hydrophobic protein, Science. 233 (1986) 464–467.
[22] Karabadzhak AG, Petti LM, Barrera FN, Edwards APB, Moya-Rodriguez A, Polikanov YS, et al. Two transmembrane dimers of the bovine papillomavirus E5 oncoprotein clamp the PDGF ß receptor in an active dimeric conformation. Proc Natl Acad Sci U S A 2017;114:E7262-E71.
[23] V.M. Nappi, J.A. Schaefer, L.M. Petti, Molecular examination of the transmembrane requirements of the platelet-derived growth factor beta receptor for a productive interaction with the bovine papillomavirus E5 oncoprotein, J. Biol. Chem. 277 (2002) 47149–47159.
[24] C.C. Lai, C. Henningson, D. DiMaio, Bovine papillomavirus E5 protein induces oligomerization and trans-phosphorylation of the platelet-derived growth factor beta receptor, Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 15241–15246.
[25] L.A. Nilson, D. DiMaio, Platelet-derived growth factor receptor can mediate tumorigenic transformation by the bovine papillomavirus E5 protein, Mol. Cell. Biol. 13 (1993) 4137–4145.
[26] L. Petti, D. DiMaio, Stable association between the bovine papillomavirus E5 transforming protein and activated platelet-derived growth factor receptor in transformed mouse cells, Proc. Natl. Acad. Sci. U. S. A. 89 (1992) 6736–6740.
[27] L. Petti, L.A. Nilson, D. DiMaio, Activation of the platelet-derived growth factor receptor by the bovine papillomavirus E5 transforming protein, EMBO J. 10 (1991) 845–855.
[28] D.J. Goldstein, T. Andresson, J.J. Sparkowski, R. Schlegel, 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 (1992) 4851–4859.
[29] K. Talbert-Slagle, D. DiMaio, The bovine papillomavirus E5 protein and the PDGF ß receptor: it takes two to tango. (Minireview), Virology. 384 (2009) 345–351.
[30] L.L. Freeman-Cook, D. DiMaio, Modulation of cell function by small transmembrane proteins modeled on the bovine papillomavirus E5 protein, Oncogene. 24 (2005) 7756–7762.
[31] K.M. Chacon, L.M. Petti, E.H. Scheideman, V. Pirazzoli, K. Politi, D. DiMaio, De novo selection of oncogenes, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) E6–E14.
[32] L. He, H. Steinocher, A. Shelar, E.B. Cohen, E.N. Heim, B.B. Kragelund, et al., Single methyl groups can act as toggle switches to specify transmembrane protein–protein interactions, eLife. 6 (2017), e27701.
[33] E.N. Heim, J.L. Marston, R.S. Federman, A.P. Edwards, A.G. Karabadzhak, L.M. Petti, et al., Biologically active LIL proteins built with minimal chemical diversity, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) E4717–E4725.
[34] C. Choma, H. Gratkowski, J.D. Lear, W.F. DeGrado, Asparagine-mediated self-association of a model transmembrane helix, Nature Struct. Biol. 7 (2000) 161–166.
[35] F.X. Zhou, M.J. Cocco, W.P. Russ, A.T. Brunger, D.M. Engelman, Interhelical hydrogen bonding drives strong interactions in membrane proteins, Nat. Struct. Biol. 7 (2000) 154–160.
[36] F.X. Zhou, H.J. Merianos, A.T. Brunger, D.M. Engelman, Polar residues drive association of polyleucine transmembrane helices, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 2250–2255.
[37] R. Gurezka, R. Laage, B. Brosig, D. Langosch, A heptad motif of leucine residues found in membrane proteins can drive self-assembly of artificial transmembrane segments, J Biol Chem 274 (1999) 9265–9270.
[38] W. Ruan, V. Becker, U. Klingmuller, D. Langosch, The interface between self-assembling erythropoietin receptor transmembrane segments corresponds to a membrane-spanning leucine zipper, J. Biol. Chem. 279 (2004) 3273–3279.
[39] D.A. Drummond-Barbosa, R.R. Vaillancourt, A. Kazlauskas, D. DiMaio, Ligand-independent activation of the platelet-derived growth factor beta receptor: requirements for bovine papillomavirus E5-induced mitogenic signaling, Mol. Cell. Biol. 15 (1995) 2570–2581.
[40] T.J. Cammett, S.J. Jun, E.B. Cohen, F.N. Barrera, D.M. Engelman, D. DiMaio, Construction and genetic selection of small transmembrane proteins that activate the human erythropoietin receptor, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 3447–3452.
[41] E.B. Cohen, S.J. Jun, Z. Bears, F.N. Barrera, M. Alonso, D.M. Engelman, et al., Mapping the homodimer interface of an optimized, artificial, transmembrane protein activator of the human erythropoietin receptor, PLoS One 9 (2014), e95593.
[42] J.B. Ptacek, A.P. Edwards, L.L. Freeman-Cook, D. DiMaio, Packing contacts can mediate highly specific interactions between artificial transmembrane proteins and the PDGFbeta receptor, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 11945–11950.
[43] A.K. Doura, F.J. Kobus, L. Dubrovsky, E. Hibbard, K.G. Fleming, Sequence context modulates the stability of a GxxxG-mediated transmembrane helix–helix dimer, J. Mol. Biol. 341 (2004) 991–998.
[44] S.M. Anderson, B.K. Mueller, E.J. Lange, A. Senes, Combination of Cα-H hydrogen bonds and van der Waals packing modulates the stability of GxxxG-mediated dimers in membranes, J. Am. Chem. Soc. 139 (2017) 15774–15783.
[45] S.N. Constantinescu, X. Liu, W. Beyer, A. Fallon, S. Shekar, Y.I. Henis, et al., Activation of the erythropoietin receptor by the gp55-P viral envelope protein is determined by a single amino acid in its transmembrane domain, EMBO J. 18 (1999) 3334–3347.

[46] C.M. Deber, T.A. Stone, Relative role(s) of leucine versus isoleucine in the folding of membrane proteins, Pept. Sci. 111 (2018), e24075.
[47] A.P. Edwards, Y. Xie, L. Bowers, D. DiMaio, Compensatory mutants of the bovine papillomavirus E5 protein and the platelet-derived growth factor beta receptor reveal a complex direct transmembrane interaction, J. Virol. 87 (2013) 10936–10945.
[48] M.A. Lemmon, D.M. Engelman, Specificity and promiscuity in membrane helix interactions, Q. Rev. Biophys. 27 (1994) 157–218.
[49] W.S. Pear, G.P. Nolan, M.L. Scott, D. Baltimore, Production of high-titer helper-free retroviruses by transient transfection, Proc. Natl. Acad. Sci. U. S. A. 90 (1993) 8392–8396.
[50] D.S. Ory, B.A. Neugeboren, R.C. Mulligan, A stable human-derived packaging cell line for production of high titer retrovirus/vesicular stomatitis virus G pseudotypes, Proc Natl Acad Sci U S A 93 (1996) 11400–11406.