A Fusion Protein Platform for Analyzing Tethered Agonism in the Adhesion Family of G Protein-Coupled Receptors

Andrew N. Dates[1], Daniel T.D. Jones[1], Meredith A. Skiba[1], Maggie M. Burruss[1], Andrew C. Kruse[1], and Stephen C. Blacklow[1,2,*]

[1]Department of Biological Chemistry and Molecular Pharmacology, Blavatnik Institute, Harvard Medical School, Boston, MA 02115, USA

[2]Department of Cancer Biology, Dana Farber Cancer Institute, Boston, MA 02215, USA

[*]Correspondence: stephen_blacklow@hms.harvard.edu

Abstract

Adhesion G Protein-Coupled Receptors (aGPCRs) are a distinct family of transmembrane proteins that carry out numerous cellular functions in multicellular organisms. During maturation, aGPCRs undergo autoproteolysis to generate non-covalently associated N-terminal and C-terminal fragments (NTF and CTF, respectively). Signaling is induced when a tethered agonist at the N-terminus of the CTF is released from or unmasked by the NTF to access its binding site within the seven transmembrane domain of the CTF. Here, we show that the tethered sequence does not require a free N-terminus to signal and exploit this finding to create a fusion protein platform for interrogation of structure-function relationships in aGPCRs. We analyze tethered agonism in CD97, a protein implicated in immune cell function and cancer cell invasiveness, and demonstrate the platform's generality by creating constitutively active forms of several other aGPCRs. This platform should find wide applicability in analyzing signals transmitted by tethered agonists in aGPCRs.

Introduction

As the largest class of membrane proteins in the human genome, G protein-coupled receptors (GPCRs) are central modulators of physiology.[1] The 33 Adhesion G Protein-Coupled Receptors (aGPCRs) comprise the second-largest family of GPCRs (behind the rhodopsin family) and influence development, cell maturation, and tissue maintenance through roles in myelination, synapse formation, immune cell migration, auditory perception, and other critical processes.[2–9] Despite their abundance in the human genome, aGPCRs are relatively understudied compared to other GPCR classes,[10] and not currently targeted by therapeutics.

The aGPCRs have a unique modular domain organization, with three functional units (Fig. 1). The adhesion module typically consists of a series of protein interaction domains responsible for ligand binding.[11] Between the adhesion module and the membrane is a regulatory region for signaling, called the GPCR Autoproteolysis-Inducing (GAIN) domain. The third module consists of the seven-transmembrane (7TM) GPCR domain, which couples to intracellular effectors to execute biological outputs. A distinctive feature of the GAIN domain is that it undergoes autocatalytic cleavage during maturation. This auto-processing event divides the aGPCR into two subunits: an N-terminal fragment (NTF) containing the adhesion module and most of the GAIN domain and a C-terminal fragment (CTF) containing a highly conserved residual "stalk" from the GAIN domain followed by the 7TM domain and cytoplasmic tail.[12,13] The stalk is comprised of ~15-25 residues, with the first seven amino acids being highly conserved across family members and largely hydrophobic in nature. The first seven stalk residues are a sequestered β-strand in the GAIN domain and this sequence functions as a tethered agonist (TA) upon removal of the NTF, which has been postulated to be triggered upon the delivery of force by bound ligand to the adhesion module.[14–22]

An atomic-level view of tethered agonism in aGPCRs has recently emerged from cryo-electron microscopic (cryo-EM) studies of seven CTFs from different aGPCRs.[23–26] These studies define a binding pose of the TA in the orthosteric site and the conformational changes imposed on the GPCR to facilitate transition to the active, G protein-bound state. The TA adopts a conserved U-shaped conformation when buried within the orthosteric pocket and is connected to the first transmembrane domain (TM1) by a poorly conserved linker which comprises the remainder of the stalk.

66 Recombinant CTF proteins expressed with an added N-terminal methionine residue (M-
67 CTF) have been used to evaluate tethered agonism and the effects of specific TA residues on
68 aGPCR activity.[14] However, the M-CTF approach for use in comparing the signaling activity of
69 different mutants of a given receptor can be challenging, in part because the dynamic range of
70 the M-CTF-induced signal is variable among receptors, and in part because M-CTF constructs
71 lack N-terminal signal sequences, and as a result the amount of protein expressed (and the extent
72 to which the N-terminal methionine is retained) can vary among the mutants being tested.[14,16,20]
73 One facet of aGPCR activation that remains unclear is whether liberation of a native, free
74 N-terminus is strictly required for tethered agonism, or whether intramolecular agonism can
75 occur when the TA is preceded by protein sequences or domains. The recent cryo-EM structures
76 of aGPCR CTFs suggest the possibility of accommodating an N-terminal extension.[23–26] The
77 addition of short polypeptide sequences to the N-terminal end of the TA however, has had
78 opposing effects on signaling depending on the identities of receptor and fusion sequence tested,
79 suggesting that there may be variation in the capacity of individual receptors to accommodate an
80 N-terminal extension. [15,16,22,27,28] The divergence in these observations might be reconciled with a
81 simple, modular, and generalizable assay system to evaluate structure-function relationships in
82 aGPCRs.

83 Here, we assessed whether a free N-terminus is required for tethered agonism using
84 CD97 (ADGRE5) as a biologically important, prototypical example of an aGPCR.[29,30] CD97 is a
85 critical regulator of dendritic cell homeostasis in the spleen, where it senses its natural ligand
86 CD55 on red blood cells;[31–35] it is also highly expressed in many carcinomas, localizing to the
87 invasion front of tumors, and mediating invasion and migration phenotypes in multiple cancer
88 cell lines.[36–39] In constructing fusion proteins that place an unmasked TA of CD97 between an
89 inert N-terminal domain and the 7TM domain of CD97, we found that these fusion proteins are
90 potent, autonomously active molecules that require an intact TA for their signaling activity. We
91 then used this fusion protein approach to investigate structure-function relationships in CD97 and
92 adapted this strategy to create a platform for the study of aGPCR activity with high signal-to-
93 noise across a variety of aGPCRs and coupled G proteins. This work provides a broadly
94 applicable platform for the evaluation of molecular mechanisms of aGPCR agonism and
95 autoinhibition as well as the identification of G protein coupling profiles for this class of GPCRs.

96

97 **Results**

98 *Development of a robust assay for the signaling activity of the CD97 CTF*. To establish the
99 signaling activity of CD97 using a reference method, we used reporter gene assays,[40] comparing
100 the basal activity of full-length CD97 (CD97 FL) with that of the CD97 M-CTF and that of the
101 CD97 M-CTF with the stalk sequence deleted (designated M-SL, for "stalkless"). Each of these
102 proteins was expressed with a C-terminal FLAG epitope tag to facilitate detection by Western
103 blot or flow cytometry using an anti-FLAG antibody (Fig. 1). Assays were performed using a
104 serum response element (SRE) reporter, a cAMP Response Element (CRE) reporter, a Nuclear
105 Factor of Activated T Cells Response Element (NFAT-RE) reporter, and a Serum Response
106 Factor Response Element (SRF-RE) reporter (Extended Data Fig. 1); each reporter was validated
107 using ligand-induced stimulation of the human β2 adrenergic (β2AR) and muscarinic M3
108 receptors (M3R) as controls (Extended Data Fig. 2a-d). When CD97 FL is expressed and tested
109 in the reporter assay, it results in low basal activity with the SRE and SRF-RE reporters (Fig. 2
110 a,b). In contrast, expression of the CD97 M-CTF gives rise to constitutive activity with the SRE
111 and SRF-RE reporters (Fig. 2 a,b), and no detectable difference with the NFAT-RE and CRE
112 reporters (Fig. 2 c,d). The M-SL protein, which lacks the stalk sequence, shows negligible basal
113 signaling activity with all reporters, confirming that the activity of M-CTF in the SRE and SRF-
114 RE reporter assays depends on the presence of the TA. Because the steady-state amount of the
115 M-SL protein was lower than that of its M-CTF counterpart (Extended Data Fig. 2 e,f), the
116 amount of transfected DNA was titrated to achieve similar total protein amounts as judged by
117 Western blot (Extended Data Fig. 2 g), and a dose-response analysis for the two constructs was
118 also performed in the reporter assay (Extended Data Fig. 2 h). The results show that the M-SL
119 signaling activity remains silent by comparison to M-CTF across a range of comparable protein
120 amounts. These data are in agreement with previous experiments using the SRE assay and
121 establish that CD97 also exhibits a stalk-dependent signal in the SRF-RE assay.[41,42]

122 We next sought to assemble a quiescent form of the CTF in which signaling could be
123 induced by proteolysis, conceptually analogous to the PAR1 fusion of ADGRL3 that can be
124 switched on with thrombin cleavage.[16] In our design, we created a fusion protein in which the
125 tobacco etch virus protease (TEV) cleavage sequence directly links the C-terminus of an inert,
126 thermostable protein domain to the N-terminus of the CD97 CTF. For initial studies, we chose
127 maltose-binding protein (MBP) as a fusion headpiece, creating a construct we refer to as MBP-

128 CTF (Fig. 1). This design positions the N-terminal serine of the CD97 stalk sequence as the P1'
129 residue of the TEV cleavage site, enabling the removal of the fusion sequence to generate an
130 intact CTF identical to that produced by autocatalytic cleavage of the full-length protein *in vivo*.
131 We expressed this construct behind the IL2 signal sequence to facilitate consistent expression
132 and surface delivery among different CD97 variants. We also included an HA epitope tag at the
133 N-terminal end for easy cell-surface detection by flow cytometry and western blot.

134 Remarkably, we found that the MBP-CTF fusion protein and M-CTF possess similar
135 maximal constitutive signaling activity in the SRE assay, and that the DNA encoding the MBP-
136 CTF fusion was ~30-fold more potent than M-CTF, even *without* addition of TEV to catalyze
137 release of the N-terminal fusion domain (Fig. 3 a). As judged by Western blot, the MBP-CTF
138 fusion was produced in greater amounts than the M-CTF when equivalent quantities of DNA
139 were transfected, but this difference in abundance does not fully account for its increased
140 potency (Fig. 3 b). When the stalk sequence was deleted, the robust constitutive signaling
141 activity was abrogated, confirming the dependence of the reporter assay signal on the presence of
142 the stalk sequence in the fusion protein (Extended Data Fig. 3 a). Flow cytometry also showed
143 that the surface abundance of the MBP-CTF protein was much lower than that of CD97 FL or
144 the MBP-SL fusion, even though the total amount of MBP-CTF was comparable to that of the
145 two other proteins (Extended Data Fig. 3 b,c), suggesting that the active species of CD97 is
146 rapidly internalized in these cells (or conceivably is signaling from an intracellular
147 compartment[43,44]).

148 We tested whether a range of different N-terminal fusions resulted in comparable
149 constitutive signaling by the CD97 CTF. Indeed, sequences ranging from thermostable protein
150 domains such as T4-lysozyme and BRIL to entities as short as a 3xFLAG epitope sequence all
151 produce robust signals when tested in the SRE reporter assay (Fig. 3 c, Extended Data Fig. 4 a).
152 We then co-expressed the MBP-CTF with a TEV protease variant optimized for activity within
153 the secretory pathway to determine whether cleavage to generate the CTF would increase the
154 reporter signal.[45] At all amounts of transfected TEV-expressing plasmid tested, the MBP-CTF
155 fusion was efficiently cleaved to populate a ~34 kDa band corresponding to the native CTF. The
156 reporter assay signal remained comparable to the uncleaved MBP-CTF after cleavage and did
157 not increase after MBP removal, suggesting that the MBP fusion does not influence GPCR
158 signaling by the TA (Extended Data Fig. 4 b,c). The intrinsic signaling activity of our N-terminal

159 fusion proteins nicely complements recent structural studies, which show that the TA sequence
160 in other aGPCRs occupies the orthosteric site of the 7TM domain and appears to have an exit
161 vector that will accommodate additional N-terminal sequences extending into the extracellular
162 space.[23–26]
163
164 *CD97 couples selectively to G12/13 heterotrimers*. Reporter assay signals are a signaling readout
165 downstream of heterotrimeric G protein activity and do not unambiguously identify the G protein
166 relied upon for the signal (Extended Data Fig. 1). We used our CD97 MBP-CTF fusion protein
167 to define the G protein coupling profile of CD97 more precisely by testing its activity inducing
168 the SRE reporter in an engineered HEK293T cell line lacking six Gα subunits: $G\alpha_s$ short, $G\alpha_s$
169 long, $G\alpha_q$, $G\alpha_{11}$, $G\alpha_{12}$, and $G\alpha_{13}$, referred to as HEKΔ6.[46] After dose optimization for each of the
170 Gα subunits and control studies with a receptor of known coupling profile (Extended Data Fig.
171 5),[47,48] we then transfected these cells with the CD97 MBP-CTF, the SRE reporter, and each of a
172 series of six Gα proteins, and recorded the signaling activity of the CD97 MBP-CTF associated
173 with each co-transfected Gα subunit. Signaling occurred only when $G\alpha_{12}$ or $G\alpha_{13}$ were co-
174 transfected with the CD97 MBP-CTF in the HEKΔ6 cells (Fig. 4), giving a CD97 coupling
175 profile in agreement with other recent studies.[31,41] These data also indicate that MBP-CTF
176 molecules can be used generally to identify Gα coupling profiles of receptors for which the
177 coupling partners are not known.
178
179 *Using the MBP-CTF platform to probe structure-function relationships in tethered agonism and*
180 *TA masking.* The TA lies laterally across the orthosteric pocket in recent cryo-EM structures of
181 other CTFs, [23–26] features which are preserved in a homology model of the CD97 CTF (Fig. 5
182 a,b). To probe structure-function relationships in the stalk of CD97, we systematically mutated
183 stalk residues in the CD97 MBP-CTF context and evaluated the effects on activity, mutating
184 residues in positions P1'-9' (531-539) individually to either alanine or lysine. For residues that
185 are alanine in the wild-type protein, we substituted glycine for alanine in the mutational scan
186 instead. Only $F533^{P3'}$ and $L536^{P6'}$ were intolerant of either substitution in the activity assay,
187 suggestive that both the size of the amino acid and hydrophobicity are necessary for activation at
188 these positions (Fig. 5 c). The $M537A^{P7'}$ mutation exhibited activity like that of the wild-type
189 protein, as did the $A534G^{P4'}$ and $A538G^{P8'}$ substitutions; however, signaling was abrogated by

190 the M537K$^{P7'}$, A534K$^{P4'}$, and A538K$^{P8'}$ mutations, indicating that either a small or non-polar
191 side chain is required at each of these positions. S531$^{P1'}$, S532$^{P2'}$, and H539$^{P9'}$ each tolerated
192 substitution to either alanine or lysine, as did I535$^{P5'}$, a highly conserved residue in the center of
193 the TA, which remarkably retained ~40% of wild-type activity even when mutated to lysine. We
194 also tested distal stalk residues in positions P10'-15' in triple alanine mutants and saw no effects
195 on activity (Extended Data Fig. 6 a). Note that the amount of mutant protein expressed was
196 comparable to that of the wild-type protein for all mutations tested except for F533A$^{P3'}$ and
197 L536A$^{P6'}$, where the steady state protein amount was lower (Extended Data Figure 6 b,c). The
198 observed effects of mutations to the stalk sequence align nicely with the homology model of the
199 CD97 CTF, in that buried hydrophobic residues are indispensable for agonism whereas residues
200 less important for agonism are tolerant of alanine/glycine mutation and tolerant of lysine
201 substitutions only when the position is exposed to solvent (Fig. 5 d).
202 We also determined the length of flexible tether needed between MBP and the TA to
203 permit constitutive agonism. All proteins tested in this experiment were present in similar
204 amounts, as judged by Western blot (Extended Data Fig. 7 a). Excision of the TEV cleavage
205 sequence and S531$^{P1'}$ had no effect on the activity of the MBP-fused CTF, whereas the
206 additional removal of S532$^{P2'}$ resulted in a reduction in activity to that of the full-length protein
207 with the TA sequestered by the GAIN domain (Fig. 6 a–c). This decrease in activity resulted
208 from steric interference by the MBP fusion, because reinsertion of the TEV cleavage site
209 between MBP and F533$^{P3'}$ restored activity to that of the parental MBP-TEV-CTF protein (Fig. 6
210 b,c).
211 In addition, we used our MBP-fusion strategy to determine the minimal portion of the
212 GAIN domain sufficient for TA sequestration in CD97. We inserted a series of increasing-length
213 segments of the GAIN domain between the TA and the TEV cleavage site of the MBP-CTF
214 fusion protein (Extended Data Fig. 7 b). Only the full-length GAIN insertion spontaneously
215 underwent autoproteolysis upon synthesis (Extended Data Fig. 7 c), and only this sequence
216 restrained the constitutive signaling activity of the MBP-CTF molecule (Fig. 6 d,e), showing that
217 a natively folded GAIN domain is the minimal requirement both for aGPCR processing and for
218 autoinhibition of TA-mediated aGPCR signaling.
219

220 *The MBP-CTF fusion platform is a generalizable strategy for robust analysis of aGPCR*
221 *signaling.* Lastly, we evaluated the generalizability of the MBP-fusion platform as a tool to study
222 aGPCR signaling properties with a panel of diverse aGPCRs. First, we compared FL, M-CTF,
223 and MBP-CTF forms of GPR56 (ADGRG1), an aGPCR which signals through $G_{12/13}$ and which
224 is evolutionarily and structurally distant from CD97 (Fig. 7 a,b) [14,23]. When 30 ng of plasmid
225 DNA expressing GPR56 was used, the MBP-CTF protein strongly stimulated the SRE and SRF-
226 RE reporters and modestly activated the NFAT-RE reporter, whereas neither the M-CTF protein
227 nor the FL protein produced a reporter signal substantially above background (Fig. 7 c–f,
228 Extended Data Fig. 8 a). Increasing the dose of transfected DNA for the M-CTF protein did
229 induce a reporter signal (Extended data Fig. 8 b). The amount of M-CTF protein produced was
230 also much lower than that of the MBP-CTF even at higher DNA doses, again highlighting the
231 advantage of using the MBP-CTF platform for reliable protein expression (Extended Data Fig. 8
232 c).
233 We curated a panel of additional MBP-CTFs corresponding to eight previously studied
234 aGPCRs (Fig. 7 g).[14,15,17,22,49–51] Our platform replicated cAMP-stimulation (as CRE reporter
235 activity) by ADGRF1, ADGRG2, ADGRG5, ADGRG6, and ADGRD1 and showed improved
236 signal amplitude by as much as 50-fold above previous analyses. We also confirmed stimulation
237 of the SRE reporter by ADGRF1, ADGRL3, ADGRG3, and ADGRV1. Every receptor construct
238 tested was reliably expressed at the amount of DNA used for transfection in our reporter assays
239 (Extended Data Fig. 8 c). ADGRL3 and ADGRD1, both of which were inactivated by fusion to
240 the PAR1 N-terminus, were highly active in at least one of the two reporters. In addition, our
241 scan uncovered three novel signals which have not been reported to date: induction of the CRE
242 by ADGRG3 and induction of the SRE by ADGRG5 and ADGRD1. These findings highlight the
243 utility of our platform to identify the downstream pathways accessible to individual receptors,
244 and demonstrate its general utility as a modular assay system to probe the activity and coupling
245 biology of diverse receptors in the aGPCR family.
246
247 **Discussion**
248 Here, we report the development of a modular and highly robust platform for evaluating
249 the signaling profiles of aGPCR family proteins. In our system, we fused MBP with a linker
250 containing a TEV cleavage site to the TA sequence at the amino terminal end of the aGPCR

251 signaling subunit (CTF). Engineered receptors with this domain organization are highly potent
252 and autonomously active in reporter gene assays. The signal rapidly reaches saturation, which is
253 maintained over a wide dose range, buffering the effects of small variations in protein abundance
254 when comparing mutants. This approach also provides advantages over the use of synthetic
255 peptides to activate 7TM domains *in trans* because it is more cost-effective and is insensitive to
256 variation in the solubility of the peptide sequence, which can confound interpretation of results
257 with synthetic peptide addition.[52] The signaling platform has a simple, modular design, is easy to
258 use, is amenable to implementation in engineered cells for determining G protein subtype
259 specificity, and both nicely complements and can be used in conjunction with other methods for
260 probing aGPCR activity, such as [^{35}S]GTPγS binding assays and proximity based recruitment
261 assays.[14,16]

262 We demonstrated the utility of our signaling platform by probing structure-function
263 relationships with the CD97 MBP-CTF. We defined its G protein subtype-specificity, probed the
264 functional roles of individual stalk residues, identified the closest allowable N-terminal fusion to
265 the TA, and determined the minimal sequence derived from the GAIN domain sufficient for
266 autoproteolysis and autoinhibition. We also showed that the structural requirements for
267 autoinhibition are more rigorous than simply blocking the N-terminal end of the stalk peptide
268 with a preceding polypeptide sequence. While other fusion constructs can inhibit autonomous
269 CTF activity,[16,27,28] our findings suggest that there are features of these fusion sequences which
270 interfere with autonomous activation, likely through some form of TA sequestration. In line with
271 this logic, we showed that, for the case of the GAIN domain, only an intact domain sequence
272 possesses such properties. These conclusions support the notion that two discrete states of the
273 TA exist, both of which involve the deep sequestration of its hydrophobic sequence within
274 another protein domain.

275 Across the wide variety of mutational variants tested in our studies of CD97, our platform
276 generated consistent amounts of expressed protein except for the F533A$^{P3'}$ and L536A$^{P6'}$ loss-of-
277 function mutations, which showed a modest decrease in total protein amounts. Mutation of these
278 residues to lysine had no detectable effect on protein expression (Extended Data Fig. 6 a,b). We
279 speculate that disabling nonpolar mutations expose the TA as a hydrophobic neoepitope, which
280 targets the protein for degradation by quality control machinery in the secretory pathway,

281 whereas substitution of a charged residue such as lysine is not detected as an exposed
282 hydrophobic surface and thus avoids a degradative fate.
283 We also showed that our platform should be applicable to signaling across other proteins
284 in the aGPCR family by testing representative examples across six aGPCR subfamilies. In doing
285 so, we uncovered three coupling activities not previously reported, indicating the utility of our
286 platform in identifying all accessible downstream pathways of a given receptor. The expression
287 advantages conferred by our system may prove particularly useful for studying aGPCRs that are
288 poorly expressed or less biochemically tractable in other assays. Of the 33 human aGPCRs,
289 several receptors do not possess canonical autoproteolytic cleavage sequences, raising the
290 question of whether they signal through intramolecular ligation of the TA.[53] Use of our fusion
291 protein platform to interrogate signaling of the CTF can determine whether aGPCRs that lack a
292 canonical autoproteolytic sequence are able to signal when the TA is unmasked.
293 Though our data show that TA unmasking is sufficient to activate aGPCRs, it is
294 important to recognize that our findings do not distinguish between existing models that favor
295 either complete dissociation of the TA from the rest of the GAIN domain or some form of de-
296 sequestration or unmasking of the TA as the key step in activation. Without a completely intact
297 GAIN domain, the default is a receptor molecule in an autonomously active state (Fig. 6 d). This
298 finding is also in line with previously reported clinical and *in vitro* observations that destabilizing
299 mutations of the GAIN domain lead to either autonomous activation or reduced force thresholds
300 of activation.[54,55] Additional data will be required to fully define how these disease-associated
301 mutations affect receptor function in molecular detail.
302

Methods

Materials

DMEM, PBS, and 0.25% trypsin were from Corning. Opti-MEM, 0.1mg/mL poly-D-lysine, and penicillin-streptomycin were from ThermoFisher Scientific. FBS was from GeminiBio. Lipofectamine 2000 was from Invitrogen. Carbachol was from Santa Cruz Biotechnology and (-)-isoproterenol hydrochloride was from Sigma-Aldrich. Rabbit anti-FLAG (D6W5B), Mouse anti-HA (6E2), Mouse anti-Beta Tubulin (D3U1W), Rabbit anti-GAPDH (14C10), and HRP-conjugated Rabbit anti-GAPDH (D16H11) XP antibodies were from Cell Signaling Technology. IRDye 800-conjugated Donkey Anti-Rabbit and IRDye 680-conjugated Goat Anti-Mouse secondary antibodies were from LI-COR Biosciences. To generate 647-conjugated Mouse anti-HA antibody, the antibody was purified from hybridoma supernatant and chemically conjugated to Alexa Fluor 647 NHS Ester from ThermoFisher Scientific. The Dual-Luciferase Reporter Assay System kit was from Promega.

Plasmid DNA Constructs

All proteins were assembled using human aGPCR and G protein sequences retrieved from the Uniprot database. Synthetic gene fragments (Integrated DNA Technologies) were inserted into vectors using Gibson assembly, and plasmid sequences were confirmed by DNA sequencing. Receptor constructs with a 5' IL2-signal sequence were subcloned into the pFuse-hIgG1-Fc2 vector using EcoRI and BglII restriction sites. M-CTF constructs (which do not contain an IL2-signal sequence) were inserted into the pFuse-hIgG1-Fc1 vector using EcoRI and BglII restriction sites. All receptor constructs included a 3' stop codon preceding the hIgG1-Fc coding sequence. An "empty vector" construct (referred to as "EV"; see Table S1) was generated by placement of a stop codon immediately after the IL2-signal sequence of the pFuse-hIgG1-Fc2 vector. A single glycine-serine linker was inserted between any epitope tag and the coding sequence of the protein of interest. Detailed construct features and calculated molecular weights for proteins used in this study are listed in Table S1.

Gα proteins were encoded using synthetic DNA sequences and subcloned into pcDNA3.1 using HindIII and NotI restriction sites. The secTEV construct was subcloned into pcDNA3.1 with HindIII and XhoI restriction sites. Point mutants were generated using a Q5 Site-Directed Mutagenesis Kit (New England Biolabs).

We generated luciferase reporter plasmids containing individual response elements, a minimal promoter (minP), and *luc2* firefly luciferase. These plasmids are analogous to the pGL4.29[CRE/minP/*luc2P*], pGL4.30[NFAT-RE/minP/*luc2P*], pGL4.33[SRE/*luc2P*], pGL4.34[SRF-RE/*luc2P*] plasmids with the exceptions that *luc2* is used in place of *luc2P* in our system and that minP is included in our SRE and SRF-RE plasmids. To assemble these plasmids, synthetic gene fragments containing response elements were inserted into a pGL4.23[*luc2*/minP] vector obtained from Promega using KpnI and HindIII (for SRE, SRF-RE, and CRE) or KpnI and NcoI (for NFAT-RE, which included the minP in the synthetic gene fragment) restriction sites. The pRL-TK plasmid (Promega) was used to express *Renilla* luciferase.

Cell Culture

347
348 Parental HEK293T and HEKΔ6[46] cells were cultured in DMEM supplemented with 10% FBS
349 and 1% penicillin-streptomycin at 37°C in a 5% CO_2, humidified incubator.
350
351 *Reporter Gene Assays*

352 Cells were seeded onto 96-well clear bottom, white walled, tissue culture-treated culture plates at
353 a density of 3 x 10^5 per well in 100 µL complete media (DMEM/10% FBS/1% penicillin-
354 streptomycin) and incubated overnight. Prior to seeding, plates were treated with 40 µL of 10
355 µg/mL aqueous poly-D-lysine solution for 30 minutes at room temperature. 16-20 hours after
356 seeding, media was replaced with 50 µL complete media and cells were transfected by adding a
357 solution volume of 20 µL containing plasmid DNA and Lipofectamine 2000 at a ratio of 3 µL
358 per 1 µg DNA in Opti-MEM.
359
360 For standard single-DNA dose assays in the CRE and NFAT-RE systems, each well was co-
361 transfected with 50 ng firefly reporter plasmid, 1 ng pRL-TK plasmid, and 10 ng receptor
362 plasmid. For standard single-DNA dose assays in the SRE and SRF-RE systems, each well was
363 co-transfected with 30 ng firefly reporter plasmid, 0.6 ng pRL-TK plasmid, and 30 ng receptor
364 plasmid. For receptor DNA titration experiments (Fig. 3 a, Extended Data Fig. 1 g,h, Extended
365 Data Fig. 3 a), total DNA was balanced with EV (see Table S1). For secTEV co-transfection
366 experiments (Fig. 3 c), a constant receptor DNA load of 30 ng was used and total DNA was
367 balanced with pcDNA3.1.
368
369 After 4-6 hours, the transfection media was replaced with 100 µL complete media and incubated
370 overnight. After 16-20 hours, media was replaced with 100 µL complete media (CRE and
371 NFAT-RE) or serum-free DMEM supplemented with 1% penicillin-streptomycin (SRE and
372 SRF-RE) and cells were incubated for 6-8 hours. Cells were then washed with 100 µL PBS and
373 lysed by addition of 20 µL 1 x Passive Lysis Buffer (Promega) to each well and nutated for 15
374 minutes at room temperature. Lysates were pipetted up and down, and the assay plate was spun
375 at 400 x g for 60 seconds to remove air bubbles. The ratio of firefly to *renilla* luciferase activity
376 was measured using automated reagent injection with a GloMax luminometer system (Promega),
377 in which 50 µL LAR II and 25 µL Stop & Glo reagent (Promega) were added to each well.
378
379 For Gα rescue optimization experiments in HEKΔ6 cells (Extended Data Fig. 5 a,d,e), varying
380 amounts of Gα subunits were transfected alongside 50 ng firefly reporter plasmid and 1 ng pRL-
381 TK plasmid (CRE, NFAT-RE) or 30 ng firefly reporter plasmid and 0.6 ng pRL-TK plasmid
382 (SRE), and total DNA was balanced with pcDNA3.1. For Gα rescue panels in the SRE assay
383 (Fig. 4 a, Extended Data Fig. 5 c), 30 ng receptor DNA was co-transfected with respective Gα
384 DNA loads (20 ng of $G\alpha_s$, $G\alpha_q$, $G\alpha_{11}$, and $G\alpha_{16}$; 0.3 ng of $G\alpha_{12}$ and $G\alpha_{13}$), 30 ng firefly reporter
385 plasmid, and 0.6 ng pRL-TK plasmid, and total DNA was balanced with pcDNA3.1. For $G\alpha_{12/13}$
386 titration experiments alongside CD97 receptor constructs (Fig. 4 b,c), total DNA was balanced
387 with pcDNA3.1.
388
389 *Bioinformatics and Homology Model Generation*

390 Sequence conservation of tethered agonist residues (Fig. 5 b) was derived from BLOSUM62
391 scores based upon a multiple sequence alignment beginning at the predicted autoproteolytic

392 cleavage sites of the 33 human aGPCR CTFs. For receptors without canonical cleavage sites, the
393 P1'-equivalent residue was determined using a multiple sequence alignment of the 32 human
394 GAIN domains. As ADGRA1 possesses neither a predicted cleavage motif nor GAIN domain,
395 the P1' residue was determined by alignment of its full-length sequence with the CTFs of the
396 other 32 human aGPCRs.
397
398 To generate a homology model of the CD97 CTF in the active signaling state (Fig. 5 a,d), the
399 sequences of human CD97 residues 531-835 and human $G\alpha_{13}$ residues 1-377 were submitted to
400 the ColabFold software server,[56] which predicted a structure of the TA-agonized CTF in complex
401 with $G\alpha_{13}$. The model of the CD97 GAIN domain (Extended Data Fig. 7 b) was extracted from
402 the alphafold2 structure prediction for full-length human CD97.[57,58]
403
404 *Western Blot Analysis*

405 Cells for western blots were prepared in poly-D-lysine-treated 24-well plates in parallel with
406 cells for reporter assays, using the same transfection mixtures (which included both reporter
407 plasmids and receptor and/or secTEV plasmids), media change timelines, and time of harvest.
408 Cell density, media volumes, and DNA loads were increased five-fold relative to the amounts
409 used in 96-well format reporter assays to account for the larger well size. At the time of harvest,
410 protein samples for blots were reduced and denatured in Laemmli buffer, loaded into 4 –20%
411 Tris-Glycine gels (Bio-Rad) for SDS-PAGE, and transferred onto nitrocellulose membranes
412 (Cytiva). Blots were blocked with 5% milk in TBST and incubated with primary antibodies
413 overnight at 4°C. Simultaneous analysis of proteins of interest and loading controls were carried
414 out for each blot (with the exception of those shown in Extended Data Fig. 2 e,f) through the use
415 of two imaging channels on a LI-COR Odyssey CLx Fluorescent Imaging System. Rabbit anti-
416 GAPDH antibody was used as a loading control for blots using Mouse anti-HA antibody, and
417 Mouse anti-Beta Tubulin antibody was used as a loading control for blots using Rabbit anti-
418 FLAG antibody. Blots were incubated with secondary antibodies for 30-60 minutes prior to
419 imaging. For blots simultaneously detecting FLAG, HA, and GAPDH (Extended Data Fig. 7 c,
420 Extended Data Fig. 8 b), FLAG and HA were detected using the LI-COR system and HRP-
421 conjugated GAPDH primary antibody was detected with Western Lightning Plus-ECL
422 (PerkinElmer) using a Bio-Rad Chemidoc System.
423
424 *Cell Surface Staining and Flow Cytometry*

425 Quantification of cell surface staining by flow cytometry was performed using cells from non-
426 coated 24-well plates prepared in parallel with cells for reporter assays, using the same
427 transfection mixtures (which included both reporter plasmids and receptor), media change
428 timelines, and time of harvest. Cell density, media volumes, and DNA loads were increased five-
429 fold relative to the amounts used in 96-well format reporter assays to account for the larger well
430 size. At the time of harvest, cells were resuspended in ice-cold 20 mM HEPES buffer, pH 7.4,
431 containing 150 mM NaCl and 0.1% bovine serum albumin (HBS/BSA), transferred to a 96-well
432 V-bottom plate (Corning), and washed once with HBS/BSA. Cells were incubated at 4°C with
433 647-conjugated mouse anti-HA antibody at 1 μg/mL in HBS/BSA for 20 minutes, washed twice
434 with HBS/BSA, and analyzed using a Beckman Coulter Cytoflex flow cytometer. Data were
435 analyzed with CytExpert software. Single cells were gated using forward and side scattering and
436 staining was quantified using the mean fluorescence intensity in the APC channel.

437
438 **References**

439 1. Schiöth HB, Lagerström MC. Structural diversity of G protein-coupled receptors and
440 significance for drug discovery. *Nature Reviews Drug Discovery*. 2008;7(4):339-357.
441 doi:10.1038/nrd2518

442 2. Ushkaryov YA, Lelianova V, Vysokov N v. Catching Latrophilin With Lasso: A
443 Universal Mechanism for Axonal Attraction and Synapse Formation. *Frontiers in*
444 *Neuroscience*. 2019;13. doi:10.3389/fnins.2019.00257

445 3. Arcos-Burgos M, Jain M, Acosta MT, et al. A common variant of the latrophilin 3 gene,
446 LPHN3, confers susceptibility to ADHD and predicts effectiveness of stimulant
447 medication. *Molecular Psychiatry*. 2010;15(11):1053-1066. doi:10.1038/mp.2010.6

448 4. Yona S, Lin H, Dri P, et al. Ligation of the adhesion-GPCR EMR2 regulates human
449 neutrophil function. *The FASEB Journal*. 2008;22(3):741-751. doi:10.1096/fj.07-9435com

450 5. Bolliger MF, Martinelli DC, Südhof TC. The cell-adhesion G protein-coupled receptor
451 BAI3 is a high-affinity receptor for C1q-like proteins. *Proc Natl Acad Sci U S A*.
452 2011;108(6):2534-2539. doi:10.1073/pnas.1019577108

453 6. Langenhan T, Piao X, Monk KR. Adhesion G protein-coupled receptors in nervous
454 system development and disease. *Nature Reviews Neuroscience*. 2016;17(9):550-561.
455 doi:10.1038/nrn.2016.86

456 7. McGee JA, Goodyear RJ, McMillan DR, et al. The very large G-protein-coupled receptor
457 VLGR1: A component of the ankle link complex required for the normal development of
458 auditory hair bundles. *Journal of Neuroscience*. 2006;26(24):6543-6553.
459 doi:10.1523/JNEUROSCI.0693-06.2006

460 8. Monk KR, Oshima K, Jörs S, Heller S, Talbot WS. Gpr126 is essential for peripheral
461 nerve development and myelination in mammals. *Development*. 2011;138(13):2673-2680.
462 doi:10.1242/dev.062224

463 9. Piao X, Hill SS, Bodell A, et al. G protein-coupled receptor-dependent development of
464 human frontal cortex. *Science*. 2004;303(5666):2033-2036.
465 doi:10.1126/SCIENCE.1092780

466 10. Fredriksson R, Lagerström MC, Lundin LG, Schiöth HB. The G-protein-coupled receptors
467 in the human genome form five main families. Phylogenetic analysis, paralogon groups,
468 and fingerprints. *Mol Pharmacol*. 2003;63(6):1256-1272. doi:10.1124/MOL.63.6.1256

469 11. Purcell RH, Hall RA. Adhesion G Protein-Coupled Receptors as Drug Targets. *Annu Rev*
470 *Pharmacol Toxicol*. 2018;58:429-449. doi:10.1146/ANNUREV-PHARMTOX-010617-
471 052933

472 12. Araç D, Boucard AA, Bolliger MF, et al. A novel evolutionarily conserved domain of
473 cell-adhesion GPCRs mediates autoproteolysis. *EMBO Journal*. 2012;31(6):1364-1378.
474 doi:10.1038/emboj.2012.26

475 13. Prömel S, Langenhan T, Araç D. Matching structure with function: The GAIN domain of
476 Adhesion-GPCR and PKD1-like proteins. *Trends in Pharmacological Sciences*.
477 2013;34(8):470-478. doi:10.1016/j.tips.2013.06.002

478 14. Stoveken HM, Hajduczok AG, Xu L, Tall GG. Adhesion G protein-coupled receptors are
479 activated by exposure of a cryptic tethered agonist. *Proc Natl Acad Sci U S A*.
480 2015;112(19):6194-6199. doi:10.1073/pnas.1421785112

481 15. Liebscher I, Schön J, Petersen SC, et al. A Tethered Agonist within the Ectodomain
482 Activates the Adhesion G Protein-Coupled Receptors GPR126 and GPR133. *Cell Reports*.
483 2014;9(6):2018-2026. doi:10.1016/j.celrep.2014.11.036

484 16. Mathiasen S, Palmisano T, Perry NA, et al. G12/13 is activated by acute tethered agonist
485 exposure in the adhesion GPCR ADGRL3. *Nature Chemical Biology*. 2020;16(12):1343-
486 1350. doi:10.1038/s41589-020-0617-7

487 17. Wilde C, Fischer L, Lede V, et al. The constitutive activity of the adhesion GPCR
488 GPR114/ADGRG5 is mediated by its tethered agonist. *FASEB Journal*. 2016;30(2):666-
489 673. doi:10.1096/fj.15-276220

490 18. Demberg LM, Winkler J, Wilde C, et al. Activation of adhesion G protein-coupled
491 receptors: Agonist specificity of Stachel sequence-derived peptides. *Journal of Biological*
492 *Chemistry*. 2017;292(11):4383-4394. doi:10.1074/jbc.M116.763656

493 19. Peeters MC, Fokkelman M, Boogaard B, et al. The adhesion G protein-coupled receptor
494 G2 (ADGRG2/GPR64) constitutively activates SRE and NFκB and is involved in cell
495 adhesion and migration. *Cellular Signalling*. 2015;27(12):2579-2588.
496 doi:10.1016/j.cellsig.2015.08.015

497 20. Kishore A, Purcell RH, Nassiri-Toosi Z, Hall RA. Stalk-dependent and stalk-independent
498 signaling by the adhesion G protein-coupled receptors GPR56 (ADGRG1) and BAI1
499 (ADGRB1). *Journal of Biological Chemistry*. 2016;291(7):3385-3394.
500 doi:10.1074/jbc.M115.689349

501 21. Vizurraga A, Adhikari R, Yeung J, Yu M, Tall GG. Mechanisms of adhesion G protein–
502 coupled receptor activation. *Journal of Biological Chemistry*. 2020;295(41):14065-14083.
503 doi:10.1074/jbc.REV120.007423

504 22. Demberg LM, Rothemund S, Schöneberg T, Liebscher I. Identification of the tethered
505 peptide agonist of the adhesion G protein-coupled receptor GPR64/ADGRG2. *Biochem*
506 *Biophys Res Commun*. 2015;464(3):743-747. doi:10.1016/J.BBRC.2015.07.020

507 23. Barros-Álvarez X, Nwokonko RM, Vizurraga A, et al. The tethered peptide activation
508 mechanism of adhesion GPCRs. *Nature 2022 604:7907*. 2022;604(7907):757-762.
509 doi:10.1038/s41586-022-04575-7

510 24. Xiao P, Guo S, Wen X, et al. Tethered peptide activation mechanism of the adhesion
511 GPCRs ADGRG2 and ADGRG4. *Nature 2022 604:7907*. 2022;604(7907):771-778.
512 doi:10.1038/s41586-022-04590-8

513 25. Ping YQ, Xiao P, Yang F, et al. Structural basis for the tethered peptide activation of
514 adhesion GPCRs. *Nature 2022 604:7907*. 2022;604(7907):763-770. doi:10.1038/s41586-
515 022-04619-y

516 26. Qu X, Qiu N, Wang M, et al. Structural basis of tethered agonism of the adhesion GPCRs
517 ADGRD1 and ADGRF1. *Nature 2022 604:7907*. 2022;604(7907):779-785.
518 doi:10.1038/s41586-022-04580-w

519 27. Frenster JD, Stephan G, Ravn-Boess N, et al. Functional impact of intramolecular
520 cleavage and dissociation of adhesion G protein-coupled receptor GPR133 (ADGRD1) on
521 canonical signaling. *Journal of Biological Chemistry*. 2021;296.
522 doi:10.1016/j.jbc.2021.100798

523 28. Lizano E, Hayes JL, Willard FS. A synthetic method to assay adhesion-family G-protein
524 coupled receptors. Determination of the G-protein coupling profile of
525 ADGRG6(GPR126). *Biochemical and Biophysical Research Communications*.
526 2021;534:317-322. doi:10.1016/j.bbrc.2020.11.086

527 29. Gray JX, Haino M, Roth MJ, et al. *CD97 Is a Processed, Seven-Transmembrane,*
528 *Heterodimeric Receptor Associated with Inflammation*. Vol 157.; 1996.
529 http://www.jimmunol.org/

530 30. EICHLER W, AUST G, HAMANN D. Characterization of an early activation-dependent
531 antigen on lymphocytes defined by the monoclonal antibody BL-Ac(F2). *Scand J*
532 *Immunol*. 1994;39(1):111-115. doi:10.1111/J.1365-3083.1994.TB03348.X

533 31. Liu D, Duan L, Rodda LB, et al. CD97 promotes spleen dendritic cell homeostasis
534 through the mechanosensing of red blood cells. *Science (1979)*. 2022;375(6581).
535 doi:10.1126/science.abi5965

536 32. Liu D, Duan L, Cyster JG. Chemo- and mechanosensing by dendritic cells facilitate
537 antigen surveillance in the spleen. *Immunol Rev*. 2022;306(1):25-42.
538 doi:10.1111/IMR.13055

539 33. Karpus ON, Veninga H, Hoek RM, et al. Shear stress-dependent downregulation of the
540 adhesion-G protein-coupled receptor CD97 on circulating leukocytes upon contact with its
541 ligand CD55. *J Immunol*. 2013;190(7):3740-3748. doi:10.4049/JIMMUNOL.1202192

542 34. Wang T, Tian L, Haino M, et al. Improved antibacterial host defense and altered
543 peripheral granulocyte homeostasis in mice lacking the adhesion class G protein receptor
544 CD97. *Infection and Immunity*. 2007;75(3):1144-1153. doi:10.1128/IAI.00869-06

545 35. Hamann J, Vogel B, van Schijndel GMW, van Lier RAW. The seven-span transmembrane
546 receptor CD97 has a cellular ligand (CD55, DAF). *J Exp Med*. 1996;184(3):1185-1189.
547 doi:10.1084/JEM.184.3.1185

548 36. Ward Y, Lake R, Faraji F, et al. Platelets Promote Metastasis via Binding Tumor CD97
549 Leading to Bidirectional Signaling that Coordinates Transendothelial Migration. *Cell*
550 *Reports*. 2018;23(3):808-822. doi:10.1016/j.celrep.2018.03.092

551 37. Safaee M, Clark AJ, Ivan ME, et al. CD97 is a multifunctional leukocyte receptor with
552 distinct roles in human cancers (Review). *International Journal of Oncology*.
553 2013;43(5):1343-1350. doi:10.3892/ijo.2013.2075

554 38. Tjong WY, Lin HH. The role of the RGD motif in CD97/ADGRE5-and EMR2/ADGRE2-
555 modulated tumor angiogenesis. *Biochemical and Biophysical Research Communications*.
556 2019;520(2):243-249. doi:10.1016/j.bbrc.2019.09.113

557 39. Aust G, Zheng L, Quaas M. To Detach, Migrate, Adhere, and Metastasize:
558 CD97/ADGRE5 in Cancer. *Cells 2022, Vol 11, Page 1538*. 2022;11(9):1538.
559 doi:10.3390/CELLS11091538

560 40. Cheng Z, Garvin D, Paguio A, Stecha P, Wood K, Fan F. Luciferase Reporter Assay
561 System for Deciphering GPCR Pathways. *Current Chemical Genomics*. 2010;4:84.
562 doi:10.2174/1875397301004010084

563 41. Ward Y, Lake R, Yin JJ, et al. LPA receptor heterodimerizes with CD97 to amplify LPA-
564 initiated RHO-dependent signaling and invasion in prostate cancer cells. *Cancer*
565 *Research*. 2011;71(23):7301-7311. doi:10.1158/0008-5472.CAN-11-2381

566 42. Hilbig D, Sittig D, Hoffmann F, et al. Mechano-Dependent Phosphorylation of the PDZ-
567 Binding Motif of CD97/ADGRE5 Modulates Cellular Detachment. *Cell Reports*.
568 2018;24(8):1986-1995. doi:10.1016/j.celrep.2018.07.071

569 43. Irannejad R, Tomshine JC, Tomshine JR, et al. Conformational biosensors reveal GPCR
570 signalling from endosomes. *Nature*. 2013;495(7442):534-538.
571 doi:10.1038/NATURE12000

572 44. Irannejad R, Pessino V, Mika D, et al. Functional selectivity of GPCR-directed drug
573 action through location bias. *Nat Chem Biol*. 2017;13(7):799-806.
574 doi:10.1038/NCHEMBIO.2389

575 45. Cesaratto F, López-Requena A, Burrone OR, Petris G. Engineered tobacco etch virus
576 (TEV) protease active in the secretory pathway of mammalian cells. *J Biotechnol*.
577 2015;212:159-166. doi:10.1016/J.JBIOTEC.2015.08.026

578 46. Grundmann M, Merten N, Malfacini D, et al. Lack of beta-arrestin signaling in the
579 absence of active G proteins. *Nature Communications 2018 9:1*. 2018;9(1):1-16.
580 doi:10.1038/s41467-017-02661-3

581 47. Inoue A, Raimondi F, Kadji FMN, et al. Illuminating G-Protein-Coupling Selectivity of
582 GPCRs. *Cell*. 2019;177(7):1933-1947.e25. doi:10.1016/J.CELL.2019.04.044

583 48. Avet C, Mancini A, Breton B, et al. Effector membrane translocation biosensors reveal G
584 protein and βarrestin coupling profiles of 100 therapeutically relevant GPCRs. *Elife*.
585 2022;11. doi:10.7554/ELIFE.74101

586 49. Hu QX, Dong JH, Du HB, et al. Constitutive Gαi Coupling Activity of Very Large G
587 Protein-coupled Receptor 1 (VLGR1) and Its Regulation by PDZD7 Protein. *The Journal*
588 *of Biological Chemistry*. 2014;289(35):24215. doi:10.1074/JBC.M114.549816

589 50. Nazarko O, Kibrom A, Winkler J, et al. A Comprehensive Mutagenesis Screen of the
590 Adhesion GPCR Latrophilin-1/ADGRL1. *iScience*. 2018;3:264-278.
591 doi:10.1016/J.ISCI.2018.04.019

592 51. Gupte J, Swaminath G, Danao J, Tian H, Li Y, Wu X. Signaling property study of
593 adhesion G-protein-coupled receptors. *FEBS Letters*. 2012;586(8):1214-1219.
594 doi:10.1016/j.febslet.2012.03.014

595 52. Shoichet BK. Screening in a spirit haunted world. *Drug Discovery Today*. 2006;11(13-
596 14):607-615. doi:10.1016/J.DRUDIS.2006.05.014

597 53. Langenhan T. Adhesion G protein–coupled receptors—Candidate metabotropic
598 mechanosensors and novel drug targets. *Basic & Clinical Pharmacology & Toxicology*.
599 2020;126(S6):5-16. doi:10.1111/BCPT.13223

600 54. Boyden SE, Desai A, Cruse G, et al. Vibratory Urticaria Associated with a Missense
601 Variant in ADGRE2. *N Engl J Med*. 2016;374(7):656-663.
602 doi:10.1056/NEJMOA1500611

603 55. Piao X, Hill SS, Bodell A, et al. G protein-coupled receptor-dependent development of
604 human frontal cortex. *Science*. 2004;303(5666):2033-2036.
605 doi:10.1126/SCIENCE.1092780

606 56. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold:
607 making protein folding accessible to all. *Nature Methods 2022 19:6*. 2022;19(6):679-682.
608 doi:10.1038/s41592-022-01488-1

609 57. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with
610 AlphaFold. *Nature 2021 596:7873*. 2021;596(7873):583-589. doi:10.1038/s41586-021-
611 03819-2

612 58. Varadi M, Anyango S, Deshpande M, et al. AlphaFold Protein Structure Database:
613 massively expanding the structural coverage of protein-sequence space with high-
614 accuracy models. *Nucleic Acids Research*. 2022;50(D1):D439-D444.
615 doi:10.1093/NAR/GKAB1061

616 59. Treisman R, Alberts AS, Sahai E. Regulation of SRF activity by Rho family GTPases.
617 *Cold Spring Harb Symp Quant Biol*. 1998;63:643-651. doi:10.1101/SQB.1998.63.643

618 60. Hill CS, Wynne J, Treisman R. The Rho family GTPases RhoA, Rac1, and CDC42Hs
619 regulate transcriptional activation by SRF. *Cell*. 1995;81(7):1159-1170.
620 doi:10.1016/S0092-8674(05)80020-0

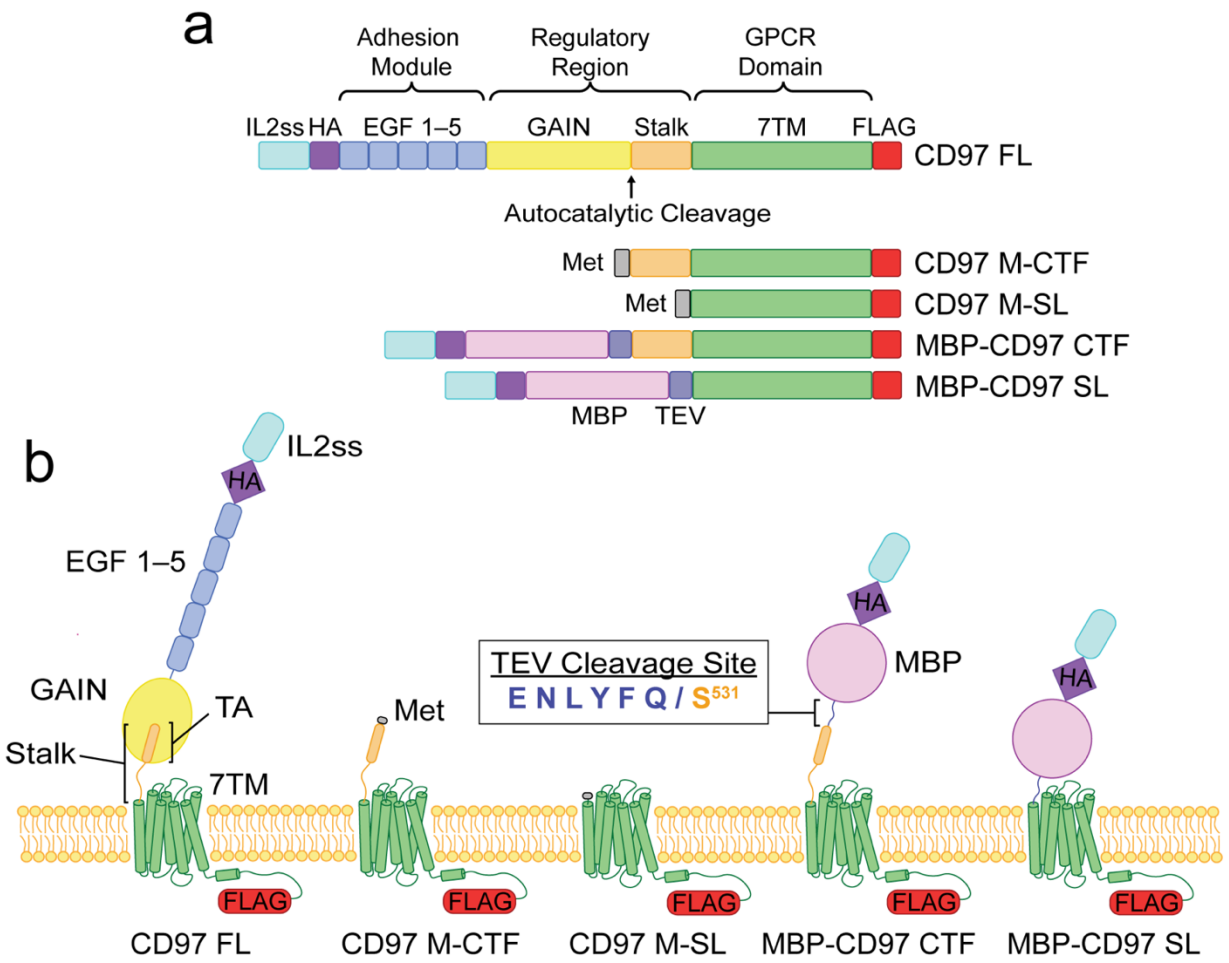
621

Acknowledgments
This work was supported by a grant from the Warren Alpert foundation (to S.C.B.) and award 1 R35-CA220340 (to S.C.B). A.N.D. was supported by a fellowship from the Fujifilm Corporation. M.A.S. was supported by a Merck Postdoctoral Fellowship from the Helen Hay Whitney Foundation. We thank Asuka Inoue (Tohoku University) for generously providing the HEKΔ6 knockout cell line. We also thank members of the Blacklow and Kruse labs for helpful discussions.

Author Contributions
A.N.D. and S.C.B. conceived of the research. A.N.D., D.T.D.J., M.A.S, A.C.K, and S.C.B designed experiments. A.N.D. and M.M.B. performed experiments. A.N.D., D.T.D.J., M.A.S, A.C.K., and S.C.B. interpreted data. A.N.D. and S.C.B. wrote the manuscript; D.T.D.J., M.A.S, A.C.K., and S.C.B. edited the manuscript.

Competing Interests
S.C.B. is on the scientific advisory board for and receives funding from Erasca, Inc., is an advisor to MPM Capital, and is a consultant for Scorpion Therapeutics, Odyssey Therapeutics, Droia Ventures, and Ayala Pharmaceuticals for unrelated projects. A.C.K. is a co-founder and consultant for Tectonic Therapeutic and Seismic Therapeutic and for the Institute for Protein Innovation, a non-profit research institute.

643 **Data Figures**
644



645 **Figure 1 – Overview of CD97 constructs used in this study. (a)** Schematic diagrams of
646 constructs. Full-length CD97 (CD97 FL) protein and Maltose-Binding Protein (MBP)-fused C-
647 terminal fragment (CTF) and "Stalkless" (SL) molecules include the interleukin 2 signal
648 sequence (IL2ss) in place of the native signal peptide. Methionine-capped C-terminal fragment
649 (M-CTF) and stalkless (M-SL) molecules have no signal sequence. In the MBP-fusion proteins,
650 a tobacco etch virus protease (TEV) cleavage sequence is located between the C-terminus of
651 MBP and the N-terminus of the CTF and SL molecules, utilizing serine 531 as its P1' residue.
652 **(b)** Schematic depicting the domain architecture of constructs in **(a)**. CD97 FL is a heterodimer
653 at the cell surface comprised of N-terminal and C-terminal fragments (NTF, CTF). CD97
654 contains five tandem EGF repeat domains followed by a GPCR autoproteolysis-inducing
655 (GAIN) domain, a 7TM domain, and a cytoplasmic tail. Autocatalytic cleavage occurs within the
656 GAIN domain to create the NTF and CTF. The stalk peptide is at the N-terminus of the CTF and
657 is sequestered within the GAIN domain.
658
659
660

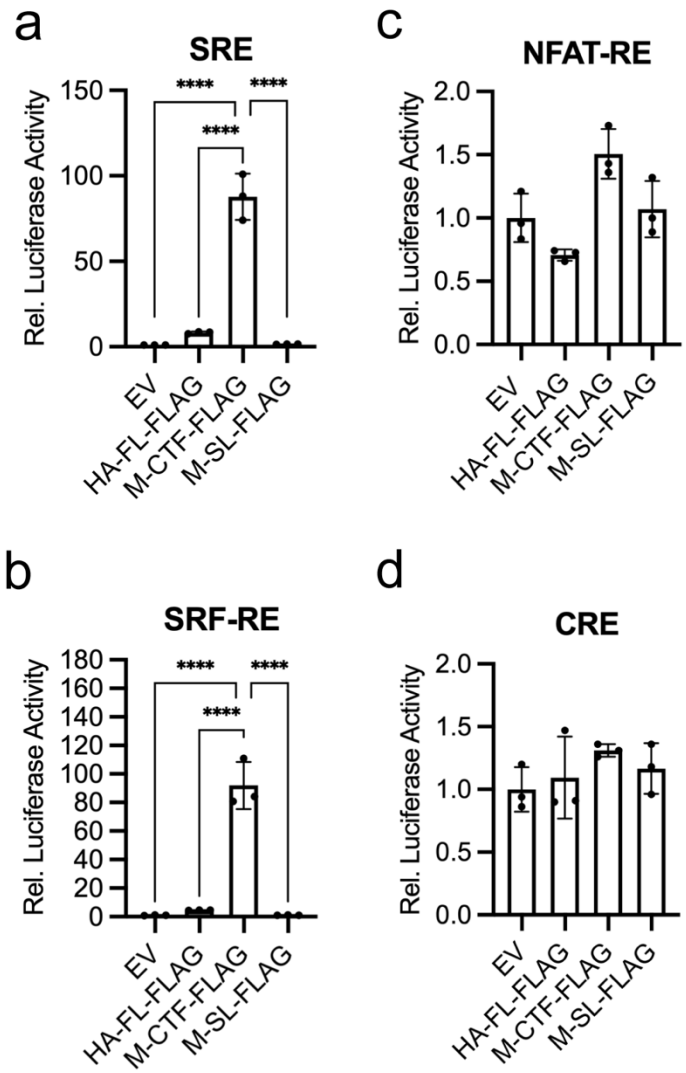


Figure 2 – CD97 signaling assays using SRE, SRF-RE, NFAT-RE, and CRE reporter genes. Signaling response of the indicated CD97-derived molecules in SRE **(a)**, SRF-RE **(b)**, NFAT-RE **(c)**, and CRE **(d)** reporter assays. Cells were transfected with 30 ng receptor DNA/well **(a,b)** or 10 ng receptor DNA/well **(c,d)**. Data are normalized to the mean value for empty vector. In all panels, data are presented as mean ± s.d. of three technical replicates and representative of three independent biological replicates. One-way analysis of variance (ANOVA) was used with Tukey's multiple-comparison post-hoc test to compare the difference between samples (****P<0.0001).

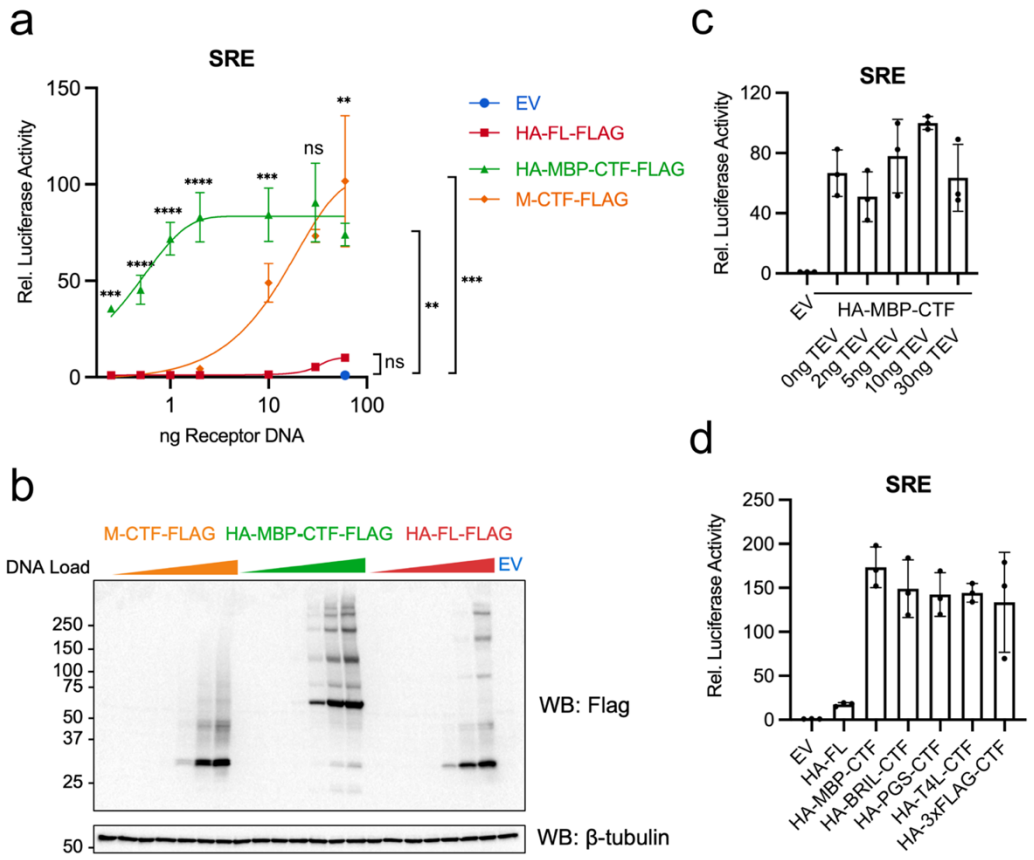


Figure 3 – Dose-response analysis of signaling by the MBP-CD97 fusion and analysis of other N-terminal fusions. (a) Comparison of HA-FL-FLAG, M-CTF-FLAG, and HA-MBP-CTF-FLAG in the SRE assay across a range of DNA transfection amounts (0.25, 0.5, 1, 2, 10, 30, or 60 ng per well). All data points are normalized to the mean value for empty vector (60 ng). Sigmoidal four parameter logistic curves were fit to each titration curve with nonlinear regression. Data are presented as mean ± s.d. of three technical replicates and representative of three independent biological replicates. Two-way ANOVA with Tukey's multiple-comparison post-hoc test was used to compare MBP-CTF and M-CTF data points at each DNA load, and one-way ANOVA with Tukey's multiple-comparison post-hoc test was used to compare 60 ng DNA loads with empty vector (ns, P>0.05, **P<0.01, ***P<0.001, ****P<0.0001). **(b)** α-FLAG Western blot comparison of protein abundance for the molecules evaluated in **(a)**. Individual lanes correspond to DNA amounts used in **(a)**. Blotting for β-tubulin was used to control for protein loading. **(c)** Signal response of the CD97 MBP-CTF when co-expressed with varying amounts (0-30 ng) of secreted TEV protease. Cells were transfected with 30 ng receptor DNA/well. Data are normalized to the mean value for empty vector and are presented as mean ± s.d. of three technical replicates and representative of three independent biological replicates. **(d)** Signaling data for a panel of N-terminal sequences fused to the CD97 CTF. Cells were transfected with 30 ng DNA/well. Data are normalized to the mean value for empty vector and are presented as mean ± s.d. of three technical replicates and representative of three independent biological replicates.

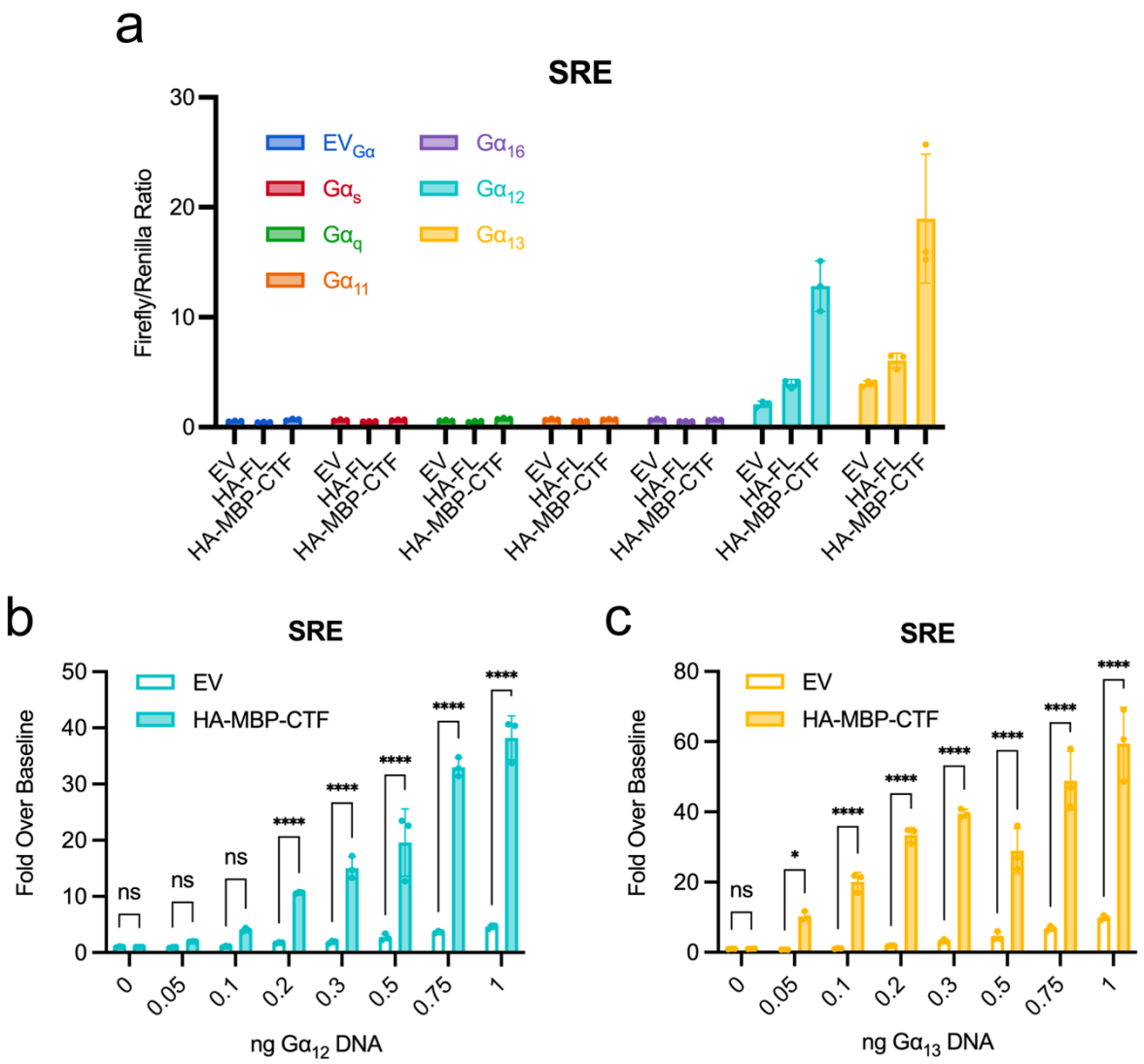


Figure 4 – Use of the MBP fusion platform to determine the G protein coupling profile of CD97. Engineered HEK293T cells lacking $G\alpha_s$, $G\alpha_{olf}$, $G\alpha_q$, $G\alpha_{11}$, $G\alpha_{12}$, and $G\alpha_{13}$ (HEKΔ6), as well as $G\alpha_{16}$ (which is not expressed in parental cells) were used to identify the G protein coupling profile of CD97. **(a)** Gα subunits (or empty vector for Gα-encoding plasmids, $EV_{G\alpha}$) were individually co-transfected with empty vector, CD97 FL, or CD97 MBP-CTF. Signals for each form of CD97 were evaluated in the SRE assay. Values are reported as the ratio of the raw luminescence values of firefly and renilla luciferases. **(b,c)** Signaling response of the CD97 MBP-CTF when co-transfected with increasing amounts of $G\alpha_{12}$ **(b)** or $G\alpha_{13}$ **(c)**. Data are normalized to the mean reporter signal generated by co-transfection of empty vector with 0 ng Gα plasmid (the leftmost bar on each graph). For all panels, cells were transfected with 30 ng receptor DNA/well. Data are presented as mean ± s.d. of three technical replicates and representative of three independent biological replicates. One-way ANOVA with Tukey's multiple-comparison post-hoc test was used to compare EV and MBP-CTF samples in **(b,c)** (ns, P>0.05,*P<0.05, ****P<0.0001).

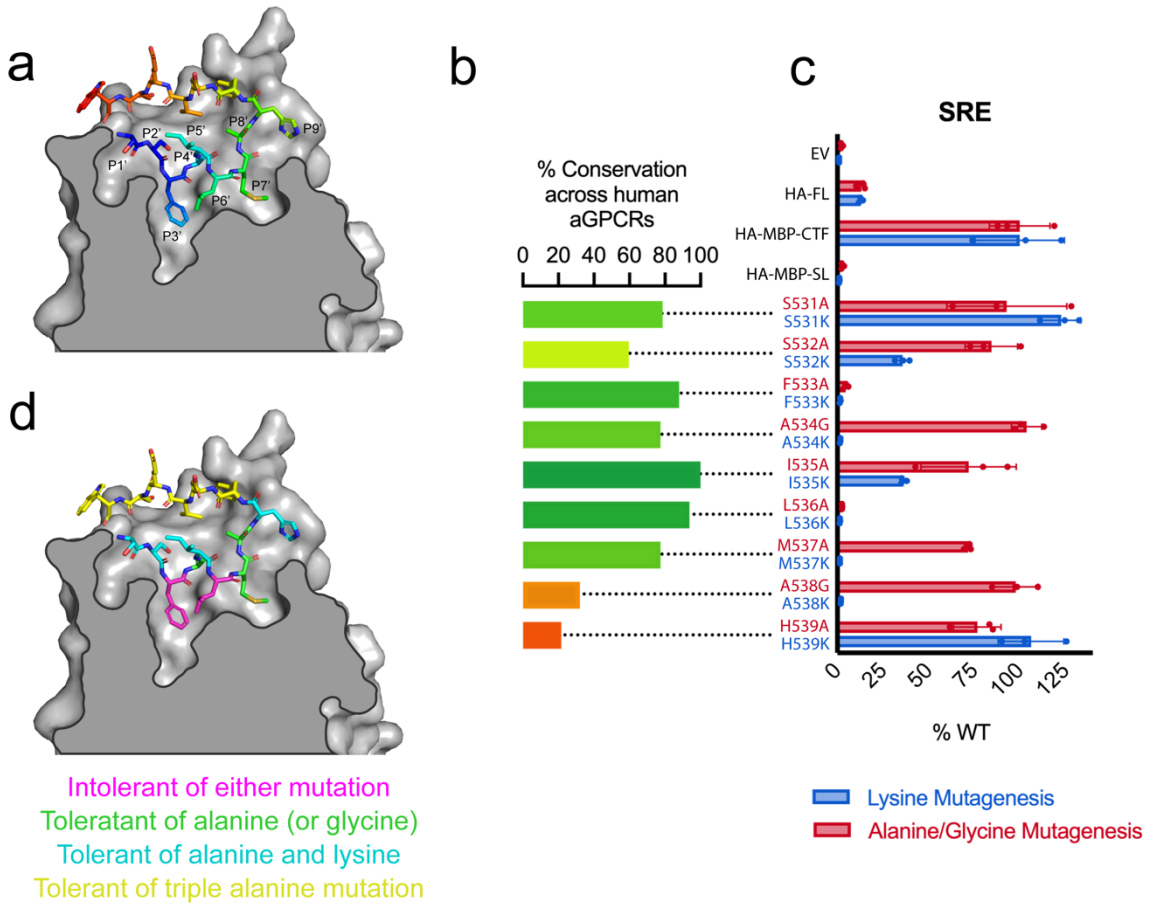


Figure 5 – Scanning mutagenesis of the CD97 stalk sequence using the N-terminal fusion platform. (a) ColabFold-generated structural prediction of the stalk-bound 7TM domain of CD97 in complex with $G\alpha_{13}$[56] A slice through the orthosteric pocket of the receptor is depicted, with the bound stalk sequence colored on a blue-to-red rainbow scale from N- to C-terminus. **(b)** Sequence conservation of the first nine stalk residues of CD97 based on alignment of the 33 human aGPCRs. **(c)** SRE signaling data for scanning alanine (or glycine for A534 and A538) and lysine mutagenesis of the stalk sequence in CD97. All mutations were made in the context of parental IL2ss-HA-MBP-TEV-CD97 CTF. Cells were transfected with 30 ng DNA/well. Data are normalized as a percentage of the mean value for WT and presented as mean ± s.d. of three technical replicates and representative of three independent biological replicates. **(d)** Mapping of mutagenesis results onto the structural model of the CD97 CTF.

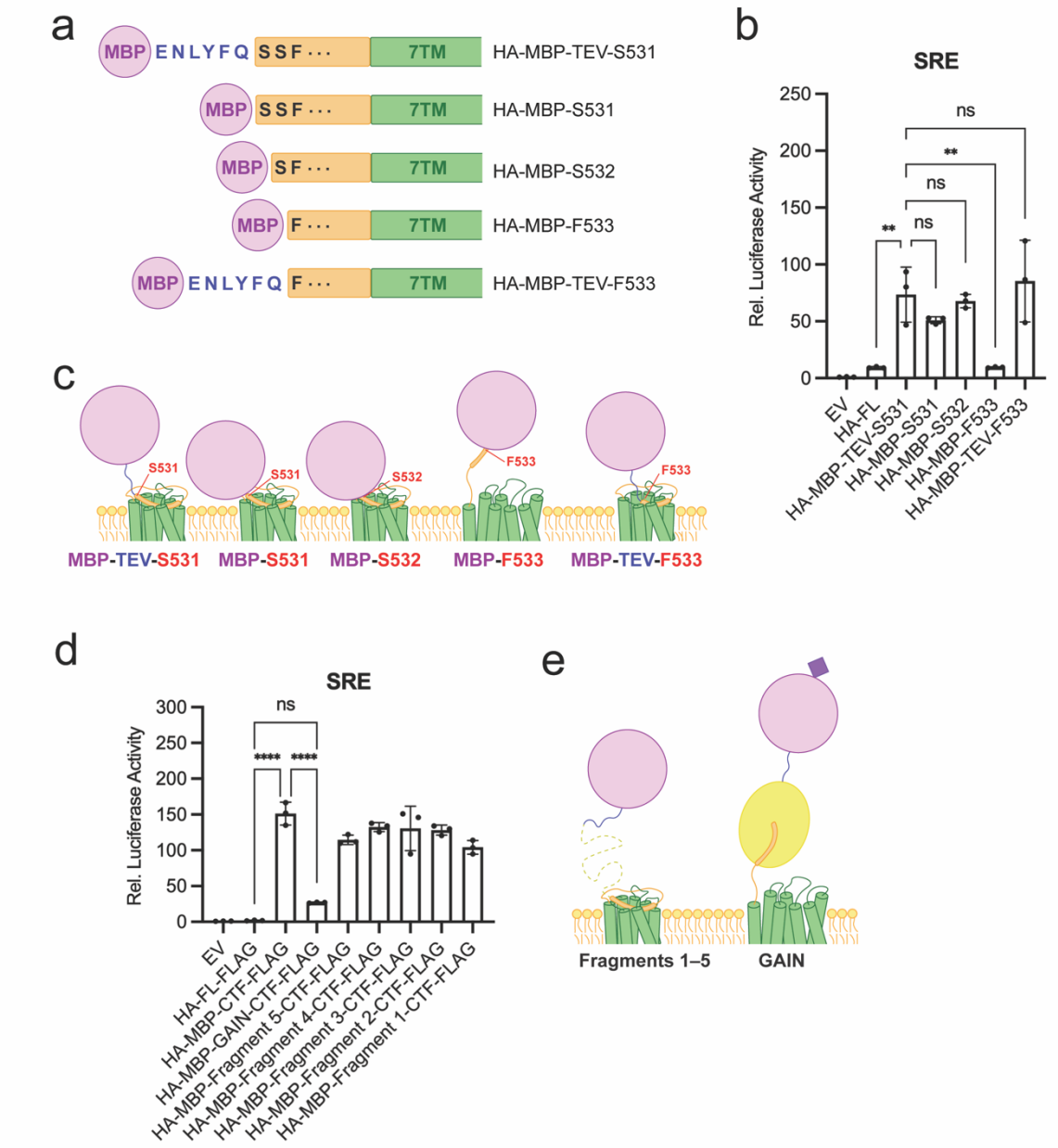


Figure 6 – Probing mechanisms of synthetic and natural autoinhibition in CD97. (a) Schematic diagrams depicting constructs in which MBP is fused at decreasing distances from the TA. **(b)** Signaling response of MBP fusion constructs shown in **(a)**. **(c)** Model illustrating the variation in steric interference depending on the fusion position of MBP. **(d)** Signaling response of proteins with insertion segments of varying lengths from the CD97 GAIN domain between MBP and the CD97 CTF (see Extended Data Fig. 7 a,b for construct design). **(e)** Schematic depiction of TA sequestration or exposure depending on whether part (left) or all of the GAIN domain is included (right). Data in **(b)** and **(d)** are normalized to the mean value for empty vector and are presented as mean ± s.d. of three technical replicates and representative of three independent biological replicates. One-way ANOVA with Tukey's multiple-comparison post-hoc test was used to compare MBP-CTF with all other receptor constructs in **(b)** and with the FL and MBP-GAIN-CTF constructs in **(d)** (ns, P>0.05,**P<0.01, ****P<0.0001).

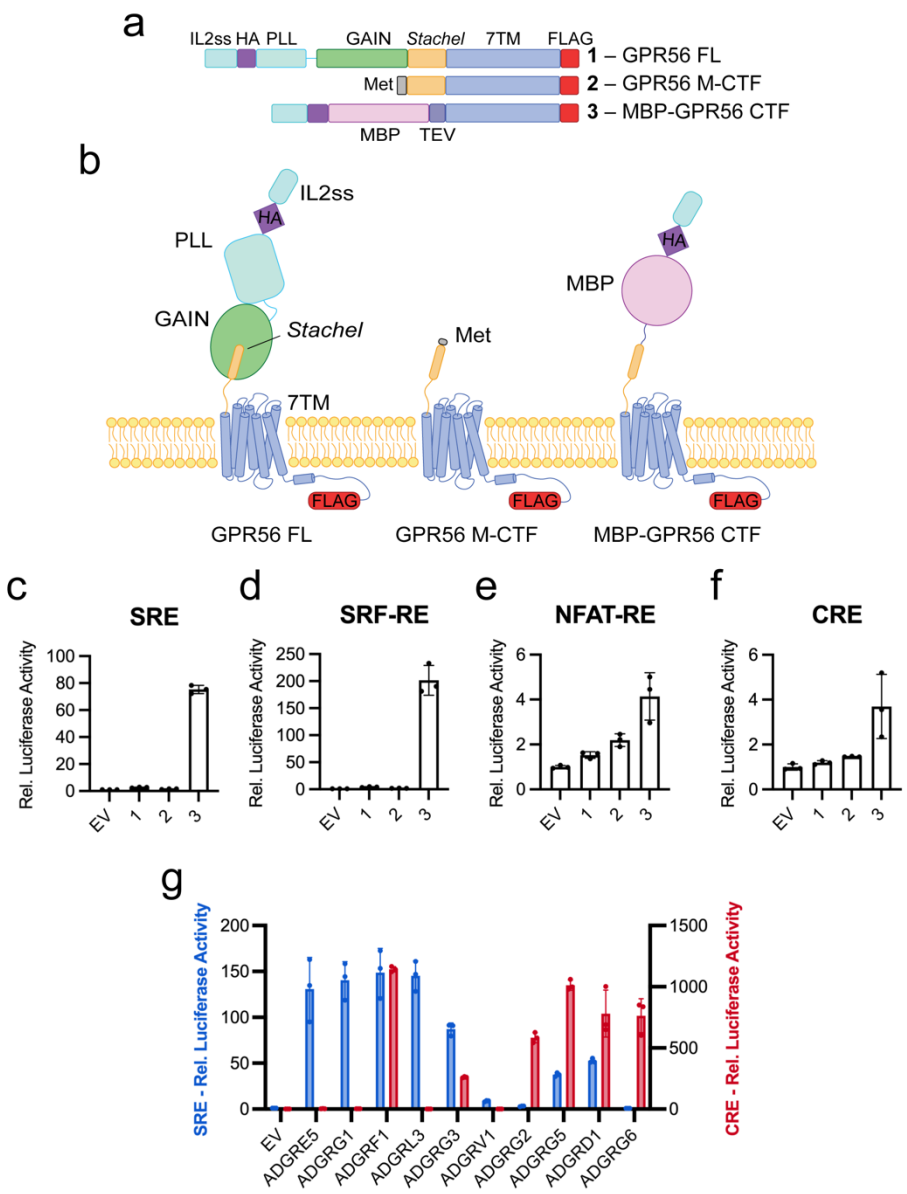
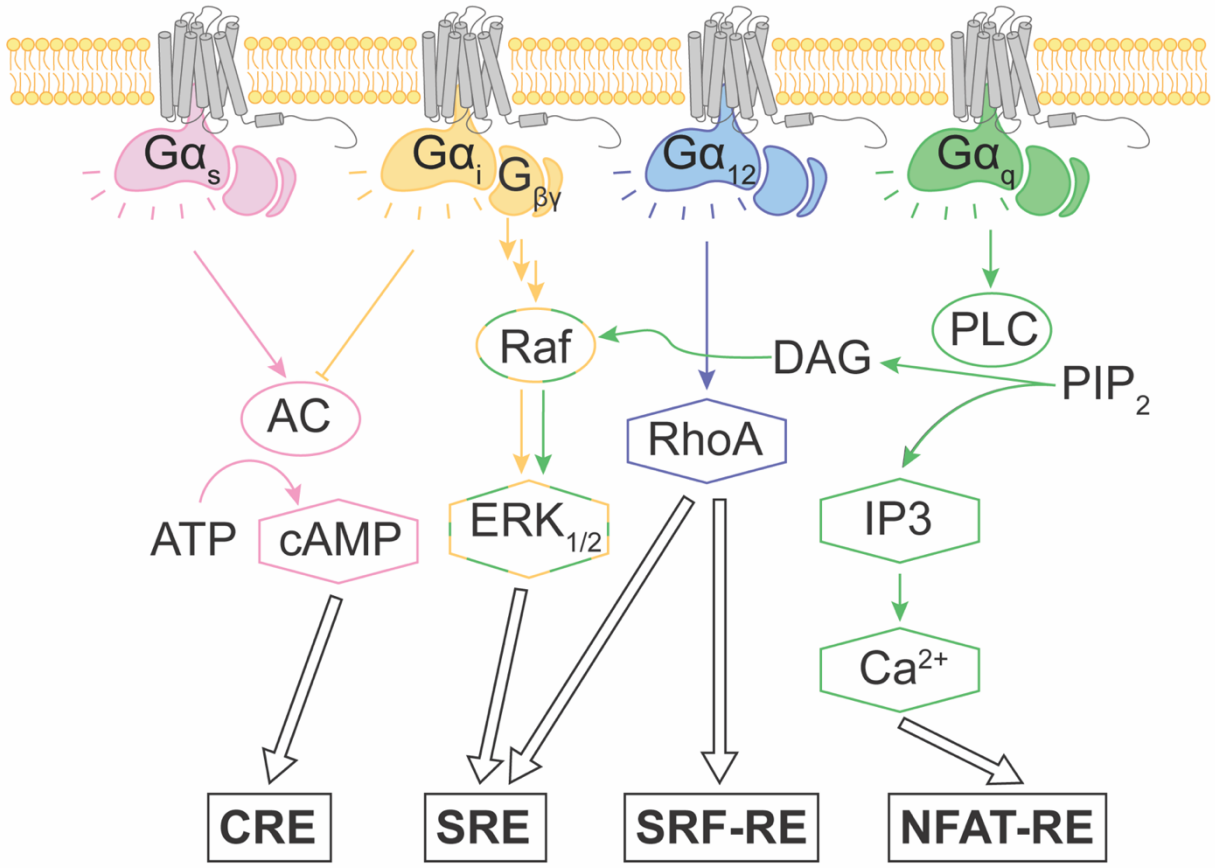
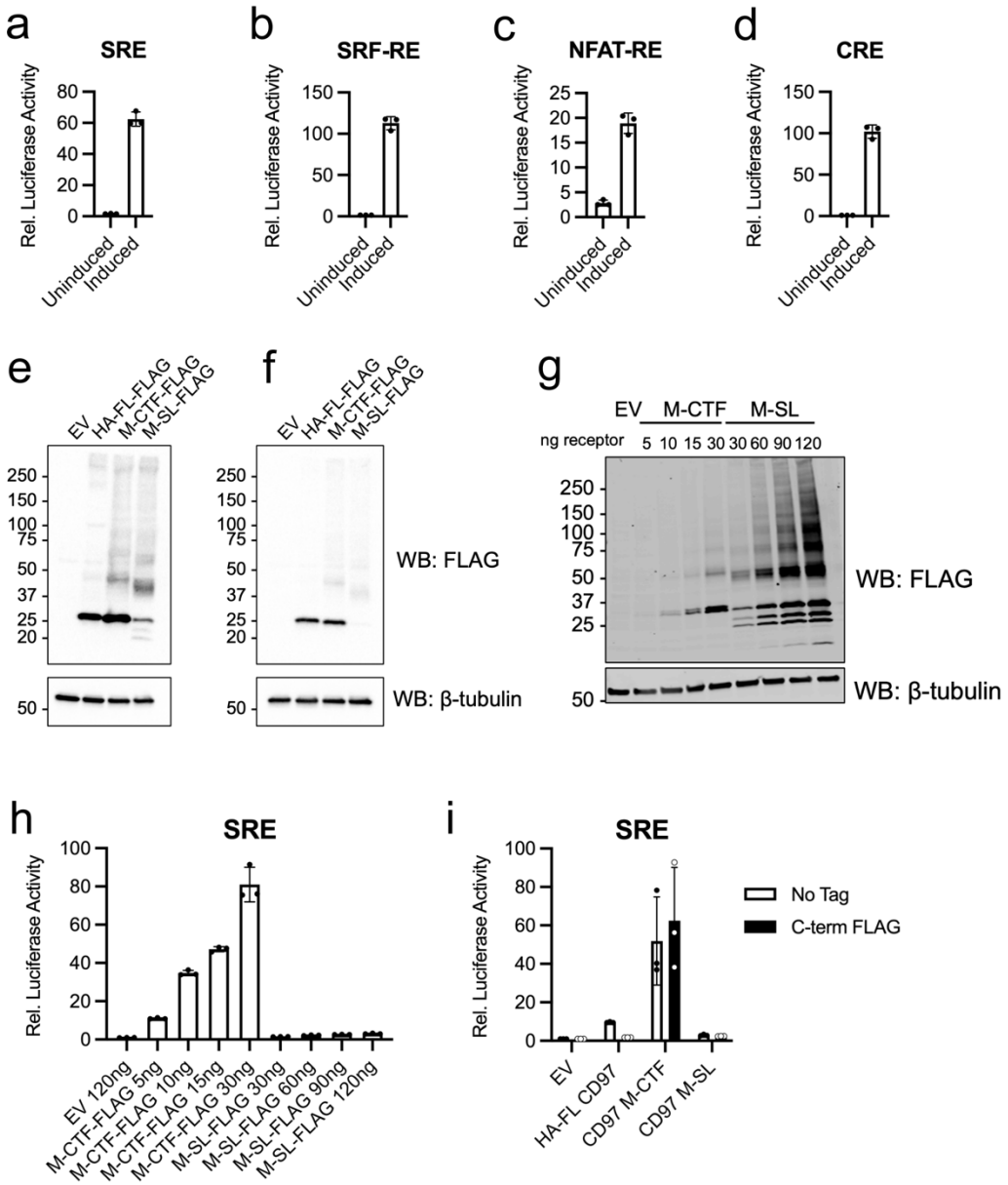


Figure 7 – Signaling assays for MBP-fused CTFs from a panel of aGPCRs. (a) Schematic diagrams of GPR56 constructs evaluated. **(b)** Schematic depicting the domain architecture of GPR56 constructs. The GPR56 headpiece is comprised of a Pentraxin/Laminin/neurexin/sex-hormone-binding-globulin-Like (PLL) domain followed by a conserved GAIN domain. **(c–f)** Signaling response of the indicated GPR56-derived molecules in SRE **(c)**, SRF-RE **(d)**, NFAT-RE **(e)**, and CRE **(f)** reporter assays. Cells were transfected with 30 ng DNA/well **(c,d)** or 10 ng DNA/well **(e,f)**. **(g)** MBP-CTF constructs for a panel of aGPCRs were evaluated in the SRE and CRE reporter assays. Cells were transfected with 30 ng DNA/well (SRE) or 10 ng DNA/well (CRE). Each receptor is expressed using the HA-MBP-TEV-CTF-FLAG protein design. The long C-terminal tail regions of ADGRL3 and ADGRG2 were truncated for ease of cloning and expression (see Table S1). Data in **(c–g)** are normalized to the mean value for empty vector (EV). Data are reported as mean ± s.d. of three technical replicates and representative of two **(c–f)** or three **(g)** independent biological replicates.

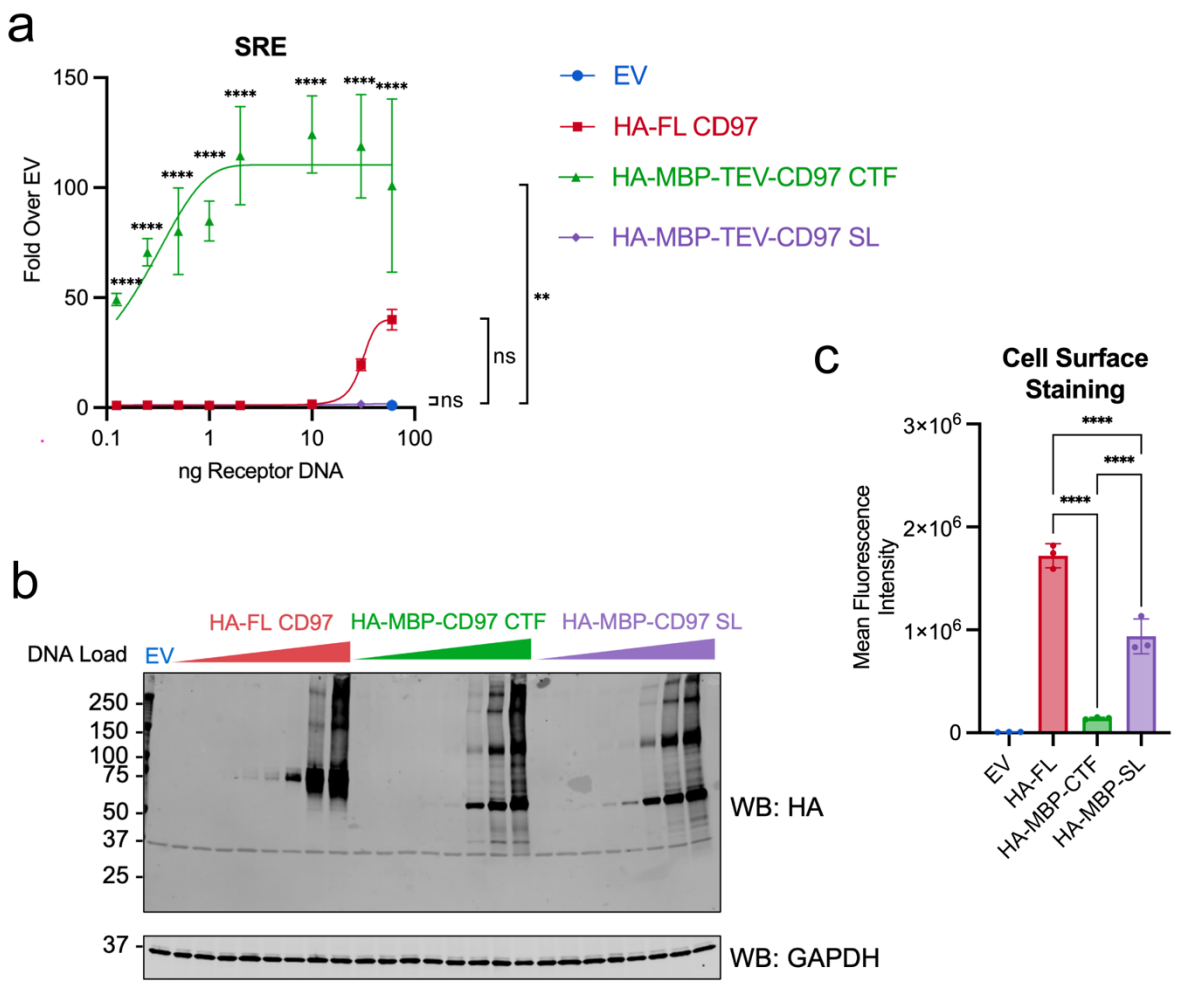
786 **Extended Data Figures**



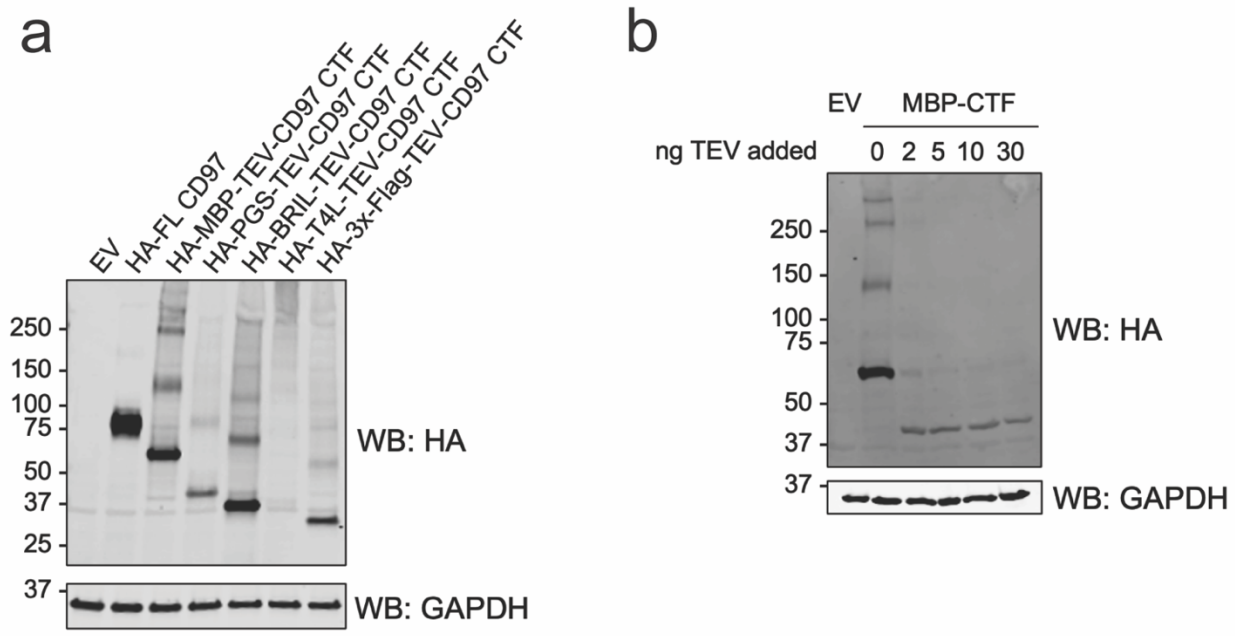
787 **Extended Data Figure 1 – Effector dependence of genetic reporter assays utilized in this**
788 **study.** G_s and G_i heterotrimers modulate cAMP production, which is reported by the cAMP
789 response element (CRE). The Serum Response Element (SRE) is dependent on signals from both
790 $ERK_{1/2}$ and RhoA proteins.[59] $ERK_{1/2}$ is stimulated by G_i-derived Gβγ subunits as well as a
791 signaling cascade downstream of phospholipase C (PLC) activity, which is controlled by G_q
792 heterotrimer stimulation. The small GTPase RhoA is activated by GTP-bound $G\alpha_{12}$. The Serum
793 Response Factor Response Element (SRF-RE) is an engineered form of the SRE reporter to
794 respond exclusively to RhoA activity.[60] Finally, the Nuclear Factor of Activated T Cells
795 Response Element (NFAT-RE) is downstream of the canonically G_q-modulated intracellular
796 calcium signaling cascade. These pathways are not meant to be exhaustive, but reflect the
797 generally-accepted signaling cascades downstream of the four primary G protein subfamilies.
798
799
800



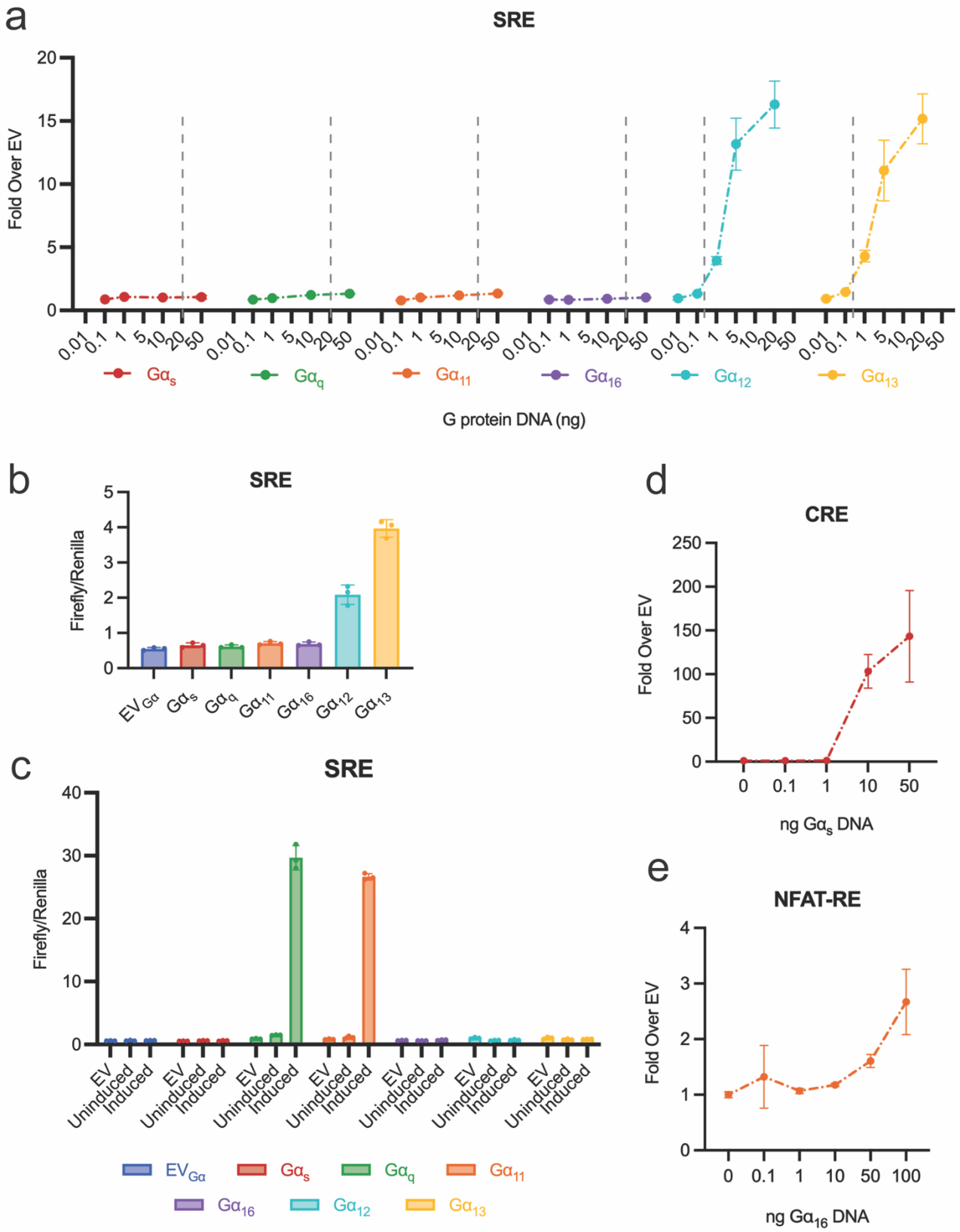
Extended Data Figure 2 – Validation of genetic reporter assays and aGPCR construct design. (a–d) Validation of reporters with muscarinic M3 receptors (M3R) stimulated with 10μM carbachol **(a–c)** or endogenous β-adrenergic receptors (β-ARs) stimulated with 100nM isoproterenol **(d)**. Cells were treated for 6 hours with drug (induced) or DMSO (uninduced) prior to readout. **(e,f)** Western blot detection of C-terminal FLAG epitope tags to analyze total protein amounts for proteins evaluated in Fig. 2 a,b **(e)** and Fig. 2 c,d **(f)**. Blotting for β-tubulin was used to control for protein loading. **(g,h)** Titration of M-CTF and M-SL DNA in Western blot against C-terminal FLAG epitope tags **(g)** and the SRE assay **(h)**. **(i)** Effects of the addition of C-terminal FLAG epitope tags on signaling constructs used in Fig. 2. Data in **(a–d,h,i)** are normalized to the mean value for empty vector. Cells were transfected with 30 ng DNA/well. Data are presented as mean ± s.d. of three technical replicates and representative of two **(h)** or three **(a–d,i)** independent biological replicates.



Extended Data Figure 3 – Effects of removing the stalk sequence from the MBP-CTF framework. (a) Comparison of the indicated CD97-derived proteins in the SRE assay across a range of transfected DNA amounts (0.25, 0.5, 1, 2, 10, 30, or 60 ng per well). All data points are normalized to the mean value for empty vector (60 ng). Sigmoidal four parameter logistic curves were fit to each titration curve with nonlinear regression. Data are presented as mean ± s.d. of three technical replicates and representative of three independent biological replicates. Two-way ANOVA with Tukey's multiple-comparison post-hoc test was used to compare MBP-CTF and M-CTF data points at each DNA load, and one-way ANOVA with Tukey's multiple-comparison post-hoc test was used to compare 60 ng DNA loads with empty vector (ns, P>0.05,**P<0.01, ****P<0.0001). **(b)** Western blot detection of N-terminal HA tags on the proteins evaluated in **(a)**. Individual lanes correspond to the amounts of transfected DNA used in **(a).** Note that CD97 FL is predicted to be heavily glycosylated. Blotting for GAPDH was used to control for protein loading. **(c)** Flow cytometry to assess cell surface protein abundance. Proteins were detected using an antibody against the N-terminal HA epitope tag. Data are presented as mean fluorescence intensity values. Data are presented as mean ± s.d. of three technical replicates and representative of three independent biological replicates. One-way ANOVA with Tukey's multiple-comparison post-hoc test was used to compare samples (****P<0.0001).

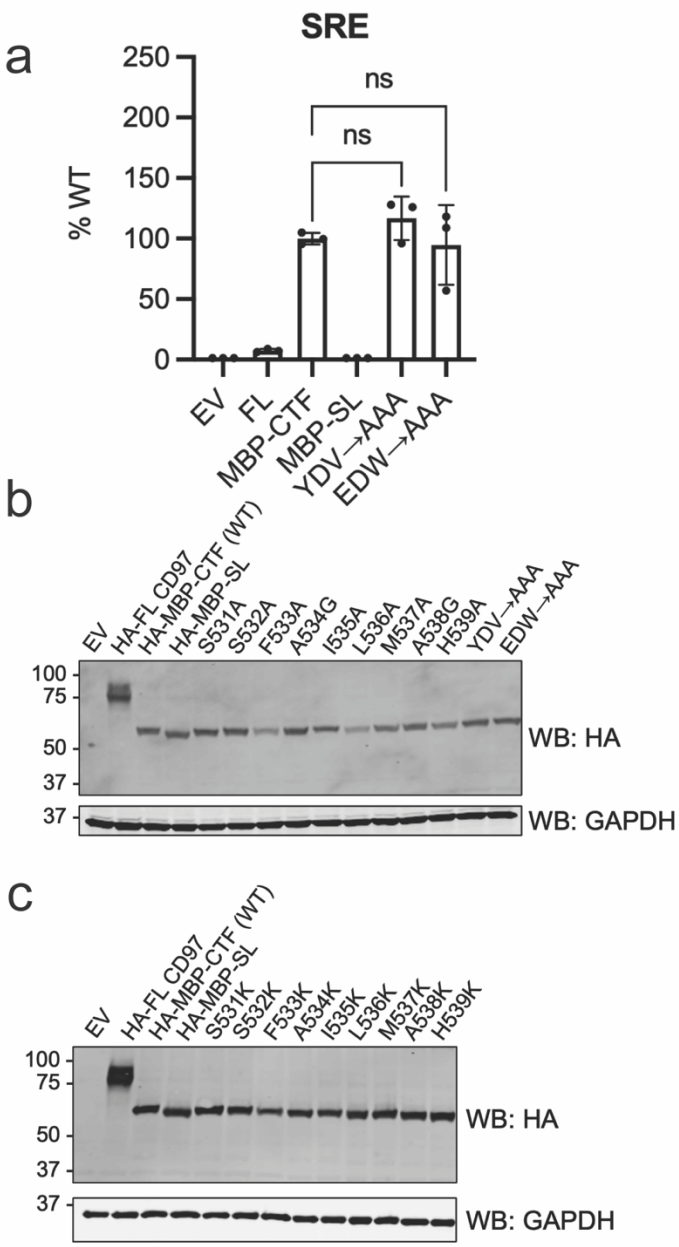


Extended Data Figure 4 – Western blots probing a panel of N-terminal CTF fusions TEV protease-catalyzed release of the MBP tag from MBP-CTF. (a) Western blot detection of N-terminal HA epitope tags for fusion proteins evaluated in Fig. 3 c. Note that CD97 FL is predicted to be heavily glycosylated. Blotting for GAPDH was used to control for protein loading. **(b)** Western blot detection of N-terminal HA epitope tags of the CD97 MBP-CTF when co-expressed with varying amounts (2-30 ng) of secreted TEV protease. Expected weights of and pre- and post-processed protein are 76.6 kDa and 42.3 kDa, respectively. Blotting for GAPDH was used to control for protein loading.


a
SRE
Fold Over EV
G protein DNA (ng)
Gαs
Gαq
Gα11
Gα16
Gα12
Gα13
b
SRE
Firefly/Renilla
EVGα
c
SRE
Firefly/Renilla
EV
Uninduced
Induced
d
CRE
Fold Over EV
ng Gαs DNA
e
NFAT-RE
Fold Over EV
ng Gα16 DNA

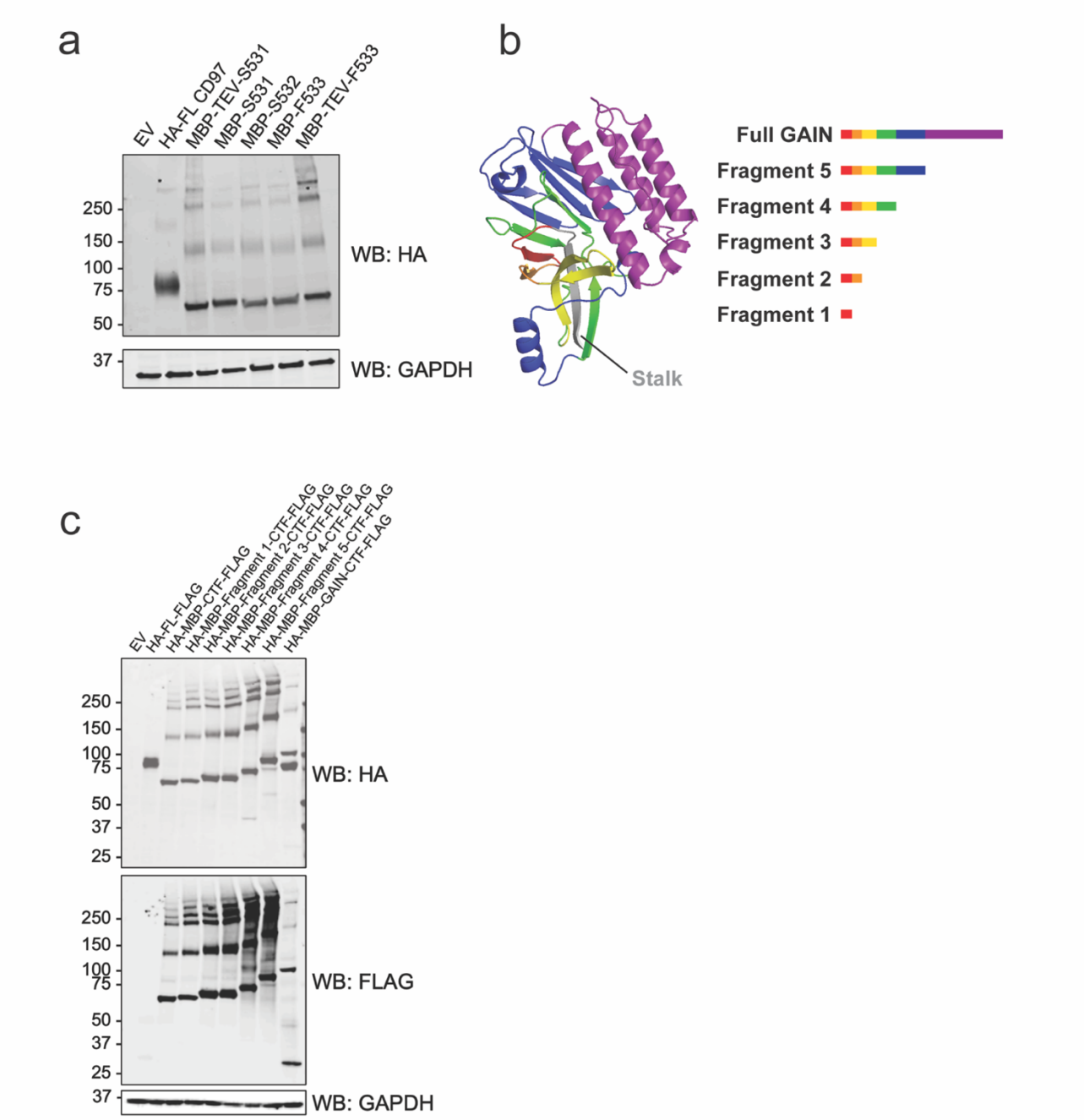
843

Extended Data Figure 5 – Calibration of G protein knockout cells and Gα rescue system to deconvolute G protein subtype specificity. (a) Reporter plasmids were co-transfected into HEKΔ6 cells with DNA encoding each Gα subtype (titrated up to 50 ng) and the relative luciferase activity was measured 30 hours after transfection. Gα DNA transfection amounts selected for rescue experiments are shown by dashed lines intersecting the x-axis: 20 ng for $G\alpha_s$, $G\alpha_q$, $G\alpha_{11}$, and $G\alpha_{16}$; 0.3 ng for $G\alpha_{12}$ and $G\alpha_{13}$. Data are normalized to the mean value for empty vector (50 ng). Data are presented as mean ± s.d. of three technical replicates and representative of two independent biological replicates. **(b)** Basal signaling activity induced by the amounts of Gα DNA transfected in rescue experiments. Values are reported as the ratio of the raw luminescence values of firefly and renilla luciferases. Data are presented as mean ± s.d. of three technical replicates and representative of three independent biological replicates. **(c)** Rescue experiment with 10 μM carbachol-stimulated M3R to confirm its known coupling profile of $G\alpha_q$ and $G\alpha_{11}$. Values are reported as the ratio of the raw luminescence values of firefly and renilla luciferases. Cells were transfected with 30 ng receptor DNA/well. Data are presented as mean ± s.d. of three technical replicates and representative of three independent biological replicates. **(d,e)** DNA titration with $G\alpha_s$ **(d)** and $G\alpha_{16}$ **(e)** to confirm their ability to induce signals in HEKΔ6 cells. Data are normalized to the mean value for empty vector (50 ng for **(d)** and 100 ng for **(e)**). Data are presented as mean ± s.d. of three technical replicates and representative of three independent biological replicates.

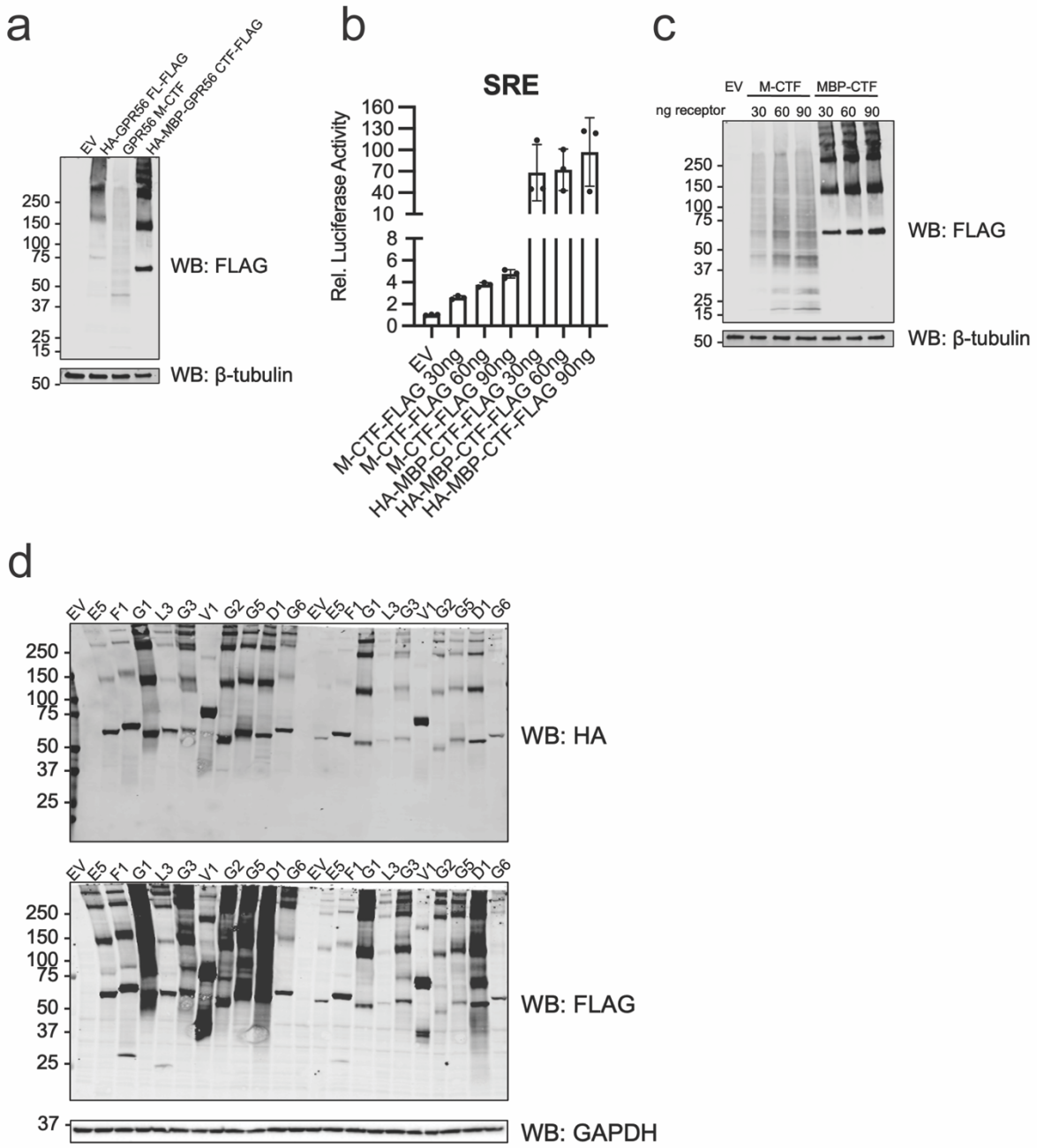


Extended Data Figure 6 – Effect of triple alanine mutations in the stalk linker and Western blot evaluation of the expression of all mutant proteins tested. (a) Signaling responses of MBP-CTF constructs harboring triple alanine mutations of residues in the P10-15 positions of the CD97 stalk sequence. Data are normalized to the mean value for empty vector. Data are presented as mean ± s.d. of three technical replicates and representative of three independent biological replicates. One-way ANOVA was used with Tukey's multiple-comparison post-hoc test to compare the difference between samples (ns, P>0.05). **(b,c)**. Western blot detection of N-terminal HA epitope tags to evaluate total protein expression for constructs evaluated in Fig. 5 c and Extended Data Fig. 6 a. Blotting for GAPDH was used to control for protein loading.

934



935 **Extended Data Figure 7 – Construct design and validation of molecules used to test**
936 **synthetic and natural autoinhibition in CD97. (a)** Western blot detection of N-terminal HA
937 epitope tags to evaluate total protein expression for GAIN domain insertions evaluated in Fig. 6
938 b. Blotting for GAPDH was used to control for protein loading. **(b)** Alphafold2 prediction of the
939 CD97 GAIN domain.[57,58] GAIN fragments tested in reporter assays are colored on the model as
940 indicated on the schematics at right. **(c)** Western blot detection of N-terminal HA and C-terminal
941 FLAG epitope tags on proteins tested in Fig. 6 d. Blotting for GAPDH was used to control for
942 protein loading



Extended Data Figure 8 – Evaluation of expression and processing of a panel of aGPCR MBP-CTF proteins. (a) Western blot detection of C-terminal FLAG epitope tags on GPR56-derived proteins tested in Fig. 7 c,d. Blotting for β-tubulin was used to control for protein loading. **(b,c)** Titration of M-CTF and MBP-CTF DNA in the SRE assay **(b)** and Western blot against C-terminal FLAG epitope tags **(c)**. Data in **(b)** are normalized to the mean value for empty vector. Data are presented as mean ± s.d. of three technical replicates and representative of two independent biological replicates. **(d)** Western blot detection of N-terminal HA and C-terminal FLAG epitope tags for proteins evaluated in Fig. 7 g. Blotting for GAPDH was used to control for protein loading.

Table S1 – Description of constructs utilized in this study

Name	Mutations/ Modifications	Vector	Description	Calculated MW (if blotted)	Ref
SRE/MinP/*luc2*		pGL4.23	Luciferase reporter vector using the serum response element, a minimal promoter, and luc2 firefly luciferase		
SRF-RE/MinP/*luc2*		pGL4.23	Luciferase reporter vector using the serum response factor response element, a minimal promoter, and luc2 firefly luciferase		
NFAT-RE/MinP/*luc2*		pGL4.23	Luciferase reporter vector using the nuclear factor of activated T cells response element, a minimal promoter, and luc2 firefly luciferase		
CRE/MinP/*luc2*		pGL4.23	Luciferase reporter vector using the cAMP response element, a minimal promoter, and luc2 firefly luciferase		
EV		pFuse-hIgG1-Fc2	Empty pFuse-hIgG1-Fc2 plasmid with a stop codon inserted between the IL2 signal peptide and the multiple cloning site		
HA-CD97 FL-FLAG		pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, full-length CD97 isoform 1, and a FLAG tag	56.7 kDa (processed NTF); 35.4 kDa (processed CTF)	
CD97 M-CTF-FLAG	Residues 1-530 of CD97 isoform 1 removed	pFuse-hIgG1-Fc1	Methionine residue followed by a CD97 construct N-terminally truncated at the autoproteolytic cleavage site (CD97 CTF) and a FLAG tag	35.5 kDa	
CD97 M-SL-FLAG	Residues 1-545 of CD97 isoform 1 removed	pFuse-hIgG1-Fc1	Methionine residue followed by a CD97 construct N-terminally truncated at the start of TM1 (CD97 SL) and a FLAG tag	33.8 kDa	
HA-MBP-CD97 CTF-FLAG	Residues 1-530 of CD97 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, the CD97 CTF, and a FLAG tag	77.7 kDa	
HA-MBP-CD97 CTF		pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
secTEV	N23Q, C130S, and T173G substitutions	pcDNA3.1	IL2 signal peptide followed by the coding sequence of secTEV protease		45
HA-CD97 FL		pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, full-length CD97 isoform 1	90.9 kDa (unprocessed); 56.7 kDa (processed NTF)	
HA-BRIL-CD97 CTF		pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, bone restricted ifitm-like protein (BRIL), TEV cleavage sequence residues P6-P1, and the CD97 CTF	48.0 kDa	
HA-PGS-CD97 CTF		pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, a glycogen synthase domain from Pyrococcus abyssi (PGS), TEV cleavage sequence residues P6-P1, and the CD97 CTF	58.1 kDa	
HA-T4L-CD97 CTF		pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, T4-lysozyme, TEV cleavage sequence residues P6-P1, and the CD97 CTF	54.4 kDa	
HA-3xFLAG-CD97 CTF		pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, a 3xFLAG tag, TEV cleavage sequence residues P6-P1, and the CD97 CTF	39.0 kDa	
HA-MBP-CD97 CTF S531A	Alanine substitution S531A	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF S532A	Alanine substitution S532A	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF F533A	Alanine substitution F533A	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF A534G	Glycine substitution A534G	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	

HA-MBP-CD97 CTF I535A	Alanine substitution I535A	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF L536A	Alanine substitution L536A	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF M537A	Alanine substitution M537A	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF A538G	Glycine substitution A538G	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF H539A	Alanine substitution H539A	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF S531K	Lysine substitution S531K	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF S532K	Lysine substitution S532K	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF F533K	Lysine substitution F533K	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF A534K	Lysine substitution A534K	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF I535K	Lysine substitution I535K	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF L536K	Lysine substitution L536K	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF M537K	Lysine substitution M537K	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF A538K	Lysine substitution A538K	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF H539K	Lysine substitution H539K	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-S531		pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, the CD97 CTF, and a FLAG tag	75.8 kDa	
HA-MBP-S532	Residues 1-531 of CD97 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, an N-terminally-truncated CD97 CTF, and a FLAG tag	75.7 kDa	
HA-MBP-F533	Residues 1-532 of CD97 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, an N-terminally-truncated CD97 CTF, and a FLAG tag	75.6 kDa	
HA-MBP-TEV-F533	Residues 1-532 of CD97 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, an N-terminally truncated CD97 CTF, and a FLAG tag	76.4 kDa	
HA-MBP-GAIN-CTF-FLAG	Residues 1-256 of CD97 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, the CD97 GAIN domain, the CD97 CTF, and a FLAG tag	107.5 kDa (unprocessed); 72.2 kDa (processed NTF);	
HA-MBP-Fragment 5-CTF-FLA	Residues 1-347 of CD97 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, a fragment of the CD97 GAIN domain, the CD97 CTF, and a FLAG tag	97.5 kDa	
HA-MBP-Fragment 4-CTF-FLA	Residues 1-434 of CD97 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, a fragment of the CD97 GAIN domain, the CD97 CTF, and a FLAG tag	88.2 kDa	

HA-MBP-Fragment 3-CTF-FLA	Residues 1-489 of CD97 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, a fragment of the CD97 GAIN domain, the CD97 CTF, and a FLAG tag	82.2 kDa	
HA-MBP-Fragment 2-CTF-FLA	Residues 1-510 of CD97 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, a fragment of the CD97 GAIN domain, the CD97 CTF, and a FLAG tag	79.8 kDa	
HA-MBP-Fragment 1-CTF-FLA	Residues 1-520 of CD97 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, a fragment of the CD97 GAIN domain, the CD97 CTF, and a FLAG tag	78.8 kDa	
HA-GPR56 FL-FLAG		pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, full-length GPR56 isoform 1, and a FLAG tag	77.4 kDa	
GPR56 M-CTF-FLAG	Residues 1-382 of GPR56 isoform 1 removed	pFuse-hIgG1-Fc1	Methionine residue followed by a GPR56 construct N-terminally truncated at the autoproteolytic cleavage site (GPR56 CTF) and a FLAG tag	35.8 kDa	
HA-MBP-GPR56 CTF-FLAG	Residues 1-382 of GPR56 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, the GPR56 CTF, and a FLAG tag	78.0 kDa	
ADGRF1	Residues 1-566 of ADGRF1 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, the ADGRF1 CTF, and a FLAG tag	81.8 kDa	
ADGRL3	Residues 1-841 and 1134-1447 of ADGRL3 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, the ADGRL3 CTF with a truncated C-terminal tail, and a FLAG tag	76.6 kDa	
ADGRG3	Residues 1-249 of ADGRG3 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, the ADGRG3 CTF, and a FLAG tag	76.9 kDa	
ADGRV1	Residues 1-5883 of ADGRV1 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, the ADGRV1 CTF, and a FLAG tag	89.5 kDa	
ADGRG2	Residues 1-606 and 912-1017 of ADGRG2 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, the ADGRG2 CTF with a truncated C-terminal tail, and a FLAG tag	78.3 kDa	
ADGRG5	Residues 1-226 of ADGRG5 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, the ADGRG5 CTF, and a FLAG tag	76.8 kDa	
ADGRD1	Residues 1-544 of ADGRD1 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, the ADGRD1 CTF, and a FLAG tag	79.5 kDa	
ADGRG6	Residues 1-840 of ADGRG6 isoform 1 removed	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, the ADGRG6 CTF, and a FLAG tag	86.7 kDa	
HA-CHRM3-FLAG		pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, the CHRM3 coding sequence, and a FLAG tag		
CD97 M-CTF		pFuse-hIgG1-Fc1	Methionine residue followed by the CD97 CTF		
CD97 M-SL		pFuse-hIgG1-Fc1	Methionine residue followed by CD97 SL		
HA-MBP-CD97 SL		pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and CD97 SL	74.8 kDa	

HA-MBP-CD97 CTF YDV->A	Alanine substitutions Y540A, D541A, and V542A	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	
HA-MBP-CD97 CTF EDW->A	Alanine substitutions E543A, D544A, and W545A	pFuse-hIgG1-Fc2	IL2 signal peptide followed by an HA tag, maltose binding protein, TEV cleavage sequence residues P6-P1, and the CD97 CTF	76.6 kDa	