Heterogeneity of Tethered Agonist Signaling in Adhesion G Protein-Coupled Receptors

Andrew N. Dates[1], Daniel T.D. Jones[1], Jeffrey S. Smith[1,2], Meredith A. Skiba[1], Maria F. Rich[3], Maggie M. Burruss[1], Andrew C. Kruse[1], and Stephen C. Blacklow[1,4,*]

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

[2]Department of Dermatology, Brigham and Women's Hospital, 221 Longwood Avenue, Boston, MA 02115

[3]University of Cincinnati School of Medicine, Department of Molecular Genetics, Biochemistry, and Microbiology, Cincinnati, OH 45267, USA

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

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

Abstract

Adhesion G Protein-Coupled Receptor (aGPCR) signaling influences development and homeostasis in a wide range of tissues. In the current model for aGPCR signaling, ligand binding liberates or unmasks a highly conserved tethered agonist (TA) that acts as an intramolecular ligand to stimulate G protein coupling. However, a systematic, comprehensive examination of signaling has not been performed for all aGPCR family members. Here, we report a platform for profiling TA-dependent activities of aGPCRs in several different assays and apply it to all 33 human family members. Activity profiling identified a heterogeneity of responses among these aGPCRs, with only ~50% showing robust, TA-dependent signals. AlphaFold2 predictions assessing TA engagement in the predicted intramolecular binding pocket aligned with the TA-dependence of the cellular responses. The signaling information in this dataset is a comprehensive resource relevant for the investigation of all human aGPCRs and for targeting aGPCRs therapeutically.

36 **Introduction**

37 As the largest class of membrane proteins in the human genome, G protein-coupled
38 receptors (GPCRs) are central modulators of physiology[1]. The 33 Adhesion G Protein-Coupled
39 Receptors (aGPCRs, GPCR family B2) comprise the second-largest family of GPCRs, and they
40 influence development, cell maturation, and tissue maintenance through roles in synapse
41 formation, myelination, immune cell migration, auditory perception, and other critical processes[2–
42 9].

43 Despite their abundance in the human genome, aGPCRs are relatively understudied
44 compared to other GPCRs[10–14]. Unlike GPCRs from other families, which have been
45 systematically screened for coupling to G proteins and β-arrestins[15–19], aGPCRs have not been
46 comprehensively profiled for their signaling properties, limiting efforts to link molecular
47 mechanism to physiological role.

48 The aGPCRs are distinguished from other GPCRs by their modular domain organization,
49 which has three distinct elements (Fig. 1A). At the N-terminus, there is an adhesion module that
50 typically consists of a series of protein interaction domains responsible for ligand binding[20].
51 Between the adhesion module and the membrane lies a regulatory domain, called the GPCR
52 autoproteolysis-inducing (GAIN) domain. The third module consists of the seven-transmembrane
53 (7TM) GPCR domain, which couples to intracellular effectors to transduce signals across the
54 membrane. The aGPCRs also have cytoplasmic C-terminal tails of variable length.

55 A distinctive feature of some GAIN domains is that they can undergo autocatalytic
56 cleavage during maturation[21,22]. This self-processing step divides these aGPCRs into two non-
57 covalently associated subunits: an N-terminal fragment (NTF) containing the adhesion module and
58 most of the GAIN domain and a C-terminal fragment (CTF) comprised of the residual "stalk"
59 sequence from the GAIN domain, the 7TM domain, and the cytoplasmic tail (Fig. 1A). The stalk
60 is roughly 15-25 residues long, containing a conserved N-terminal segment of seven amino acids
61 that is largely hydrophobic. The first seven stalk residues are sequestered as a β-strand in the GAIN
62 domain under resting conditions[21,23–25]. For several aGPCRs, this sequence can be liberated from
63 the GAIN domain, putatively in response to force delivered to the NTF by bound ligand, to
64 function as a tethered agonist (TA) for signal activation[14,26–33].

65 An atomic-level view of tethered agonism in aGPCRs has emerged from single particle
66 cryo-electron microscopy (cryo-EM) studies of seven CTFs from different aGPCRs in G protein-
67 bound states[34–38]. For these aGPCRs, the structural studies define a conserved binding pose of the
68 TA in the orthosteric site, which facilitate transition of the GPCR to an active, G protein-bound
69 state.

70 Although the TA-induced signaling model clearly applies for some family members,
71 including those for which active state structures have been reported, whether TA ligation elicits a
72 signaling response for all aGPCRs remains unclear. For about half of aGPCRs, TA-dependent
73 signaling has not been established, in part because of the difficulty in developing an assay generally
74 applicable to all aGPCRs. Of the aGPCRs shown to exhibit TA-dependent signaling, few have
75 been evaluated side-by-side and even fewer across a comprehensive set of functional readouts such
76 as signal-induced transcription, G protein recruitment, and receptor internalization, making it
77 difficult to compare signaling properties among them[11–13].

78 Here, we report a platform for profiling TA-dependent activities of aGPCRs in several
79 different assays and apply it to all 33 human family members. Activity profiling identified a
80 heterogeneity of responses, with only ~50% of family members showing robust, TA-dependent
81 signals, largely consistent with AlphaFold2 predictions of TA engagement[39,40]. The signaling
82 information in this dataset should be broadly relevant for tool compound discovery, therapeutic
83 development, and future *in vivo* studies.

84

85 **Results**

86 ***Development of an assay platform for assessment of aGPCR signaling.*** We selected ADGRE5
87 (also known as CD97) as a prototype for development of a general approach to evaluate aGPCR
88 signaling[41–43]. We expressed recombinant full-length protein (FL), a C-terminal fragment (CTF)
89 with an N-terminal methionine residue (methionine-capped, M-CTF), and two truncated forms of
90 the CTF, one lacking the seven-residue tethered agonist sequence (M-CTFΔTA), and another
91 lacking all fifteen residues preceding TM1 (a "stalkless" construct called M-CTFΔstalk), in
92 HEK293 cells (Fig. 1A). The steady state amounts of these methionine-capped proteins varied
93 substantially (Fig. 1B), however, making it difficult to compare signaling activities of the variants
94 produced using this approach.

95 We used protein engineering to achieve consistency in expression among ADGRE5
96 variants. Active state structures of CTFs of aGPCRs show that the TA is positioned in a
97 conformation with an exit vector compatible with an N-terminal fusion extending into the
98 extracellular space (Fig. S1A)[34–38]. We postulated that insertion of a stable N-terminal fusion
99 protein would permit TA-dependent signaling and reduce variation in steady-state protein
100 abundance among different aGPCRs and their variants, enabling the measurement of TA-
101 dependent and TA-independent signaling activity for each aGPCR using a single expression
102 cassette. We designed the same series of ADGRE5 variants with a secretion signal sequence at the
103 N-terminus, followed by an HA epitope tag, the thermostable maltose binding protein (MBP), and
104 a TEV protease cleavage site preceding CTF (MBP-CTF), CTFΔTA (MBP-CTFΔTA), and
105 CTFΔstalk (MBP-CTFΔstalk) sequences (Fig. 1A,C). We compared the steady state amounts of
106 these proteins in HEK293 cells to one another and to the methionine-capped CTFs (Fig. 1B).
107 Unlike the M-CTF proteins, the MBP fusion proteins were present at similar – and substantially
108 greater – total protein abundance when transfected in equivalent amounts, enabling their direct
109 comparison in downstream functional assays.

110 We compared the signaling properties of these ADGRE5 variants in a transcriptional
111 reporter assay, an established and sensitive metric of GPCR activity[44]. Transfection of full-length
112 ADGRE5 into HEK293 cells resulted in low basal activity of the Serum Response Element (SRE)
113 reporter, whereas transfection of either MBP-CTF or M-CTF induced a strong, saturable SRE
114 reporter signal (Fig. 1D). Strikingly, the MBP-CTF fusion was roughly 30-fold more potent than
115 M-CTF as a function of plasmid dose without the need for addition of TEV to catalyze release of
116 the N-terminal fusion domain. Although the MBP-CTF fusion was produced in greater amounts
117 than the M-CTF protein when equivalent quantities of DNA were transfected (Fig. S1B), this
118 difference in abundance does not fully account for its increased potency. In contrast, the MBP-
119 CTFΔTA and MBP-CTFΔstalk proteins show negligible signaling activity, confirming that the
120 stalk sequence is necessary to induce a transcriptional reporter signal (Fig. S1C). The identity of
121 the fusion domain also had little effect on the signaling activity of ADGRE5, because a range of
122 different N-terminal fusion sequences also led to constitutive signaling by the CTF (Fig. S1D,E).

123 We serially deleted linker residues to determine the minimal spacer length between the
124 fusion protein and the TA that permitted signaling. Deletion of the entire TEV site and even the

125 first serine residue (S531$^{P1'}$) of the tethered agonist from the ADGRE5 MBP-CTF protein did not
126 silence its constitutive signaling activity. Additional removal of S532$^{P2'}$, however, reduced the
127 signaling activity to that of the quiescent FL receptor (Fig. S2). This decrease in activity resulted
128 from steric interference by the MBP fusion, because reinsertion of the TEV cleavage site between
129 MBP and F533$^{P3'}$ restored activity to that of the parental MBP-TEV-CTF protein (Fig. S2).

130 We analyzed the effects of systematically mutating individual ADGRE5 TA residues on
131 reporter gene activity to interrogate structure-function relationships in the ADGRE5 stalk. Our
132 scanning mutagenesis showed that signaling from ADGRE5 was intolerant of either alanine or
133 lysine substitution at both F533 (P3') and L536 (P6') (Fig. S3A), consistent with the known
134 importance of these positions for signaling by other aGPCRs and the deep burial of these TA
135 positions in active state aGPCR structures[26,27,34–38]. The observed effects of other stalk sequence
136 mutations aligned with a homology model of the ADGRE5 CTF (Fig. 3B)[40], with solvent exposed
137 and predicted linker residues tolerant of both alanine and lysine substitutions, and other partially
138 buried positions tolerant only of alanine (or glycine, for P4' residue A534 and P8' residue A538).
139 The MBP-CTF platform resulted in similar amounts of expressed protein for all mutants tested
140 except F533A and L536A, both of which were signaling incompetent in the reporter assay and
141 present in lower abundance than wild-type protein at steady state as judged by western blot (Fig.
142 S3C,D). The reduced abundance of these variants is likely a consequence, rather than a cause, of
143 their signaling incompetence; the analogous signaling-deficient F569A mutant in the TA of
144 ADGRF1 was also nearly undetectable when expressed as a CTF protein lacking a stabilizing
145 domain such as MBP[45].

146 ***Assessment of TA-dependent signaling across the complete human aGPCR family.*** We evaluated
147 the ability of each human aGPCR to be activated by its TA peptide in the context of an HA-MBP-
148 TEV-CTF-FLAG fusion analogous to the ADGRE5 prototype. We aligned the GAIN domain
149 sequences of all human aGPCRs to identify the P1' position of each and assembled a complete set
150 of CTF fusion proteins with the P1' residue positioned immediately after the TEV site (designated
151 MBP-CTF) (Fig. S4). We also made a matched set of fusion proteins lacking the tethered agonist
152 in which the conserved P1'-P7' residues that comprise the TA were deleted, designating these
153 variants MBP-CTFΔTA (Fig. S4). For ADGRA1, which lacks a GAIN domain, the amino acid
154 sequence of the full-length receptor was present in the MBP-CTF construct, and we deleted all

155 residues prior to the start of TM1 in the MBP-CTFΔTA construct. All MBP-CTF and MBP-
156 CTFΔTA proteins were readily detectable by western blot when expressed in HEK293T cells
157 except for ADGRE4, which is a predicted pseudogene (Fig. S5). Comparable amounts of protein
158 were expressed by almost all matched MBP-CTF and MBP-CTFΔTA protein pairs, showing that
159 the MBP fusion strategy suppressed variability in expression not merely for ADGRE5, but for the
160 full family of aGPCRs.

161 We assessed TA-dependent and TA-independent signaling activity by measuring the ability
162 of each MBP-CTF and MBP-CTFΔTA protein to induce the expression of a luciferase reporter
163 gene under the control of four different response elements[44]. Assays were performed using a serum
164 response element (SRE) reporter, a cAMP Response Element (CRE) reporter, a Nuclear Factor of
165 Activated T Cells Response Element (NFAT-RE) reporter, and a Serum Response Factor Response
166 Element (SRF-RE) reporter (Fig. 2A). Signaling patterns varied substantially among the different
167 subfamilies, and there was even heterogeneity of response patterns within the different subfamilies
168 (Fig. 2B). We note that the SRE and SRF-RE reporters showed near identical response patterns
169 across aGPCR family members, suggesting that they are integrating very similar signaling inputs.
170 Of the 32 family members, 16 showed robust TA-dependent transcriptional responses with at least
171 one reporter, including proteins from each of the D, E, F, G, and L subfamilies. However, no
172 proteins of the A, B, C, or V subfamilies showed TA-dependent expression of the luciferase
173 reporter, nor did ADGRE2, ADGRE3, ADGRF2, ADGRF5, ADGRL1, or ADGRL4 show a TA-
174 dependent reporter response. Members of the B, C, and V subfamilies, as well as ADGRG3 did
175 produce activity with some reporters, but this transcriptional response was independent of the
176 presence of the TA. The TA-independent activity was also typically substantially weaker than the
177 TA-dependent signal observed in D, E, F, G, and L family members. Together, these transcriptional
178 reporter data define a functional signaling landscape for the complete family of human aGPCRs,
179 identifying a spectrum of TA responsiveness among them.

180 Transcriptional activity is a distal signaling readout downstream of G proteins, which are
181 the primary transducers of GPCR-initiated signaling cascades. We also adapted a series of
182 TRUPATH G protein biosensors to measure the TA-dependent ability of each aGPCR to stimulate
183 the dissociation of Gα subunits from Gβγ complexes, a direct consequence of coupling between
184 the aGPCR and the G protein[46]. These assays detect changes in Bioluminescence Resonance

185 Energy Transfer (BRET) associated with the dissociation of RLucII-fused Gα subunits from GFP-
186 fused Gβγ complexes upon receptor activation (Fig. 3A). For GPCRs activated by *trans* addition
187 of a soluble agonist, biosensor activation is typically reported as a ligand-dependent change in
188 BRET efficiency (ΔBRET). Because the TA is an intramolecular aGPCR ligand, here we measured
189 the BRET efficiencies of the MBP-CTF and MBP-CTFΔTA (Fig. S6) proteins and calculated the
190 difference between them to determine the TA-dependent change in BRET efficiency (ΔBRET).
191 The aGPCRs showed distinct patterns of TA-dependent G protein activation with the G_s, G_q, G_{i1},
192 and G_{13} biosensors (Fig. 3B–E). The TA-dependence of aGPCR signaling activity in the BRET
193 assay closely tracked the aGPCR activity in the transcriptional reporter assay, with proteins of the
194 A, B, C, and V subfamilies again quiescent in the BRET assay. Strikingly, the G_{13} BRET biosensor
195 was the most frequently activated, indicating a G protein preference of aGPCRs that differs from
196 family A and family B1 GPCRs (Fig. 3F). Together, the combination of reporter and BRET assays
197 revealed that the aGPCRs fall into two distinct groups: those that transduce TA-dependent signals
198 and those that do not.

199 ***Assessment of aGPCR localization.*** We also measured the steady state cell-surface abundance of
200 all MBP-CTF and MBP-CTFΔTA matched pairs by flow cytometry, using an anti-HA antibody
201 recognizing the extracellular part of the protein. Whereas western blotting showed that the amount
202 of expressed protein was roughly equivalent for each MBP-CTF and MBP-CTFΔTA matched pair
203 (Fig. S5), many of the MBP-CTF proteins from the D, E, F, G and L families were depleted from
204 the cell surface when compared with their MBP-CTFΔTA counterparts (Fig. 4). The cohort of
205 proteins that showed reduced MBP-CTF staining relative to their matched MBP-CTFΔTA partners
206 closely correlated with the group of proteins that showed TA-dependent activity in the reporter
207 gene and G protein biosensor assays. The tendency of autonomously active, MBP-CTF forms of
208 TA-dependent aGPCRs to be depleted from the cell surface relative to their inactive, MBP-
209 CTFΔTA forms is consistent with examples from other families of GPCRs in which receptor
210 activation is followed by receptor internalization[47].

211 ***Comparison of structural predictions with signaling and internalization data.*** We modeled the
212 structure of each CTF using Colabfold[40] to assess whether the TA sequence was predicted to be
213 bound intramolecularly for each CTF (*i.e.*, without the MBP fusion domain). The predicted
214 structures consistently adopted one of two conformations (Fig. 5A, Fig. S7), either with the TA

215 occupying the orthosteric site of the 7TM domain and a kink in the TM6 helix at residue $G^{6.50}$, or
216 with the TA unbound and an unkinked TM6 helix. Strikingly, the structural predictions closely
217 correlated with the TA-mediated signaling responses from our transcriptional readouts, G protein
218 coupling assays, and surface staining analyses, providing a functional landscape of aGPCR
219 activities extending from agonist binding to receptor desensitization (Fig. 5B-F).

220 Despite the strong conservation of the TA sequence and autoproteolysis motif (Fig. S8),
221 the signaling properties of the human aGPCRs are heterogeneous with respect to their TA
222 dependence. The invariance of the TA sequence thus suggests that structural motifs within the 7TM
223 itself accounts for the activity differences. Multiple sequence alignment revealed that residue $G^{6.50}$
224 within the 7TM domain is completely conserved in TA-activated aGPCRs but largely absent in the
225 quiescent receptors (Fig. 6A). This residue lies in a PXXG motif also found in family B1 (Secretin)
226 GPCRs[48], which is required to facilitate the kinking of the TM6 helix characteristic of the fully
227 active state in family B GPCRs. Mutation of this position to alanine in four TA-activated aGPCRs
228 (ADGE5, ADGRF1, ADGRG1, and ADGRL3) impaired their ability to induce genetic reporters,
229 confirming the importance of this residue in TA-dependent aGPCR activation (Fig. 6B, Fig. S9),
230 and offering a molecular rationale for the observed functional heterogeneity of the human aGPCRs.

231 It is possible that, for certain receptors, the N-terminal MBP fusion interfered with access
232 of the TA to its putative binding site on the 7TM domain, preventing TA-dependent signaling. We
233 evaluated this possibility by testing whether TEV protease-catalyzed release of the MBP fusion
234 domain would restore TA-dependent signaling activity for inactive aGPCRs. Co-expression of a
235 secreted, ER-retained form of TEV (secTEV)[49] with three TA-activated receptors (ADGRE5,
236 ADGRF1, and ADGRL3) led to processing of each MBP-CTF protein at the TEV cleavage site
237 (Fig. S10) and did not substantially change the signaling activity of these proteins in the SRF-RE
238 reporter assay. Co-expression of secTEV also led to MBP release for all receptors that were silent
239 in the TA-dependent signaling assays, except ADGRG3, and yet none of these receptors exhibited
240 TA-dependent signaling in the CRE, SRF-RE, and NFAT-RE reporters. These data confirm that
241 the quiescent subset of aGPCRs do not exhibit TA-dependent signals, even when the TA is
242 completely exposed at the N-terminus after MBP removal (Fig. S10).

243

Discussion

Numerous GPCRs belonging to families A, B1, and C have been profiled to assess G protein coupling activity and downstream transcriptional responses, but platform-based profiling of aGPCRs has been elusive because the agonist is an intramolecular ligand that is masked in the mature full-length protein. By constructing an MBP fusion platform that was modular and generalizable, we were able to evaluate the coupling and signaling response of each aGPCR to its intramolecular TA.

We compared the signaling activities of two variants for each receptor, one with and one without the TA (referred to as CTF and CTFΔTA, respectively), in reporter gene assays and with G protein biosensors, and used ColabFold to predict whether or not the TA would be bound in the orthosteric site. Our findings uncovered a striking and unexpected dichotomy among the aGPCRs, in which some exhibited TA-dependent signals and some did not.

Our studies revealed tethered agonism for three aGPCRs that had not previously been evaluated: ADGRD2, ADGRE1, and ADGRE3. We also confirmed G protein coupling responses for fourteen other receptors and identified secondary signaling responses for ADGRD1, ADGRG2, ADGRG6, ADGRF4, and ADGRL2. The aGPCRs of the A, B, C, and V subfamilies, however, as well as ADGRE2, ADGRF2, ADGRF5, ADGRL1, and ADGRL4, did not show evidence of TA-dependent signaling in transcriptional reporter assays or with G protein biosensors. ColabFold predictions of CTF structures were also consistent with the experimental signaling data with few exceptions.

It is possible that some of the receptors failing to exhibit TA-dependent signaling in our system may signal through the TA *in vivo*. ADGRE2, for example, can signal through G_{16} heterotrimers[50], which are not present in HEK293T cells, and the ADGRF2 proteins used in this study were poorly expressed. AlphaFold2 predicted these two proteins, as well as ADGRF5 and ADGRL4, to be TA binding-competent, and these receptors contain an intact PXXG motif, but these proteins showed no TA-dependent signaling in our assays. Nevertheless, release of the MBP to reveal a free TA sequence at the N-terminal end of the CTF for all of the receptors that did not exhibit TA-dependent signaling did not produce any TA-dependent signals in the reporter gene assays.

272 Some of the receptors that did not produce TA-dependent signals did have increased basal
273 amounts of TA-independent activity in some of the reporter gene assays. Whether this activity
274 represents functionally important basal signaling that is biologically relevant independent of ligand
275 induction remains unclear.

276 We used flow cytometry to assess the cell-surface amounts of each protein because receptor
277 internalization can be a mechanism for desensitizing or compartmentalizing GPCR
278 signaling[47,51,52]. The CTFΔTA form of the protein tended to be present in higher amounts at the
279 cell surface than the CTF form for proteins that exhibited TA-dependent signals, suggesting an
280 agonist-dependent effect on aGPCR internalization. These data also suggest that the constitutive
281 signaling properties of CTF proteins may complicate the use of surface abundance to normalize
282 for the relative potency of different proteins in activity assays.

283 Heterogeneity of TA-dependent signaling among the aGPCRs existed despite the high
284 conservation of the TA sequence. Its high conservation, however, may derive from its role as the
285 terminal β-strand within the core of the GAIN domain in each receptor[21]. The segregation of the
286 aGPCRs into two groups may instead be due to the presence or absence of the TM6 helix PXXG
287 motif, which is critical for the activation of family B1 GPCRs[48].

288 Although autocatalytic cleavage in the GAIN domain is necessary for dissociative presentation
289 of the TA sequence at the N-terminus of the CTF, self-cleavage within the GAIN domain does not
290 occur in all aGPCRs. Indeed, receptors lacking canonical cleavage motifs, including ADGRE1 and
291 ADGRF4, exhibited TA-dependent signaling. In ADGRF4, a polybasic consensus sequence for
292 furin cleavage lies ~30 residues toward the N-terminus from the TA and may substitute for the
293 autocatalytic cleavage event to enable subunit dissociation. For ADGRE1, however, no such
294 alternative cleavage site exists, consistent with the possibility that cleavage-independent access of
295 the TA sequence can occur, as has been proposed[53,54].

296 Finally, the aGPCRs showed a general coupling preference for $G_{12/13}$ heterotrimers, the least
297 frequently activated effectors for other GPCR families. $G_{12/13}$ activation leads to actin cytoskeletal
298 rearrangement[55,56], which correlates with the biological processes governed by many aGPCRs. Our
299 catalog of TA-dependent signals and their effector dependencies should be a valuable resource to
300 assist in analyzing the molecular bases and effector dependencies of aGPCR-driven biological

phenotypes, and ultimately should facilitate the development of tool compounds for therapeutic development.

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. J.S.S. is supported by National Institutes of Health Dermatology Training Grant 5T32AR007098-49 (J.S.S.) and the Dermatology Foundation (J.S.S.). We thank Justin English (University of Utah) for generously providing the tricistronic TRUPATH plasmids. 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., J.S.S, M.A.S, A.C.K, and S.C.B designed experiments. A.N.D., M.F.R., and M.M.B. performed experiments. A.N.D., D.T.D.J., J.S.S., M.A.S, A.C.K., and S.C.B. interpreted data. A.N.D. and S.C.B. wrote the manuscript with input from all authors.

Competing Interests

S.C.B. is on the board of directors for the non-profit Revson Foundation, non-profit Institute for Protein Innovation, is on the scientific advisory board for and receives funding from Erasca, Inc. for an unrelated project, 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. J.S.S has received consulting fees from Biogen for an unrelated project.

331 **Methods**

332 *Materials*

333 DMEM, PBS, and 0.25% trypsin were from Corning. DMEM with no phenol red/glucose/
334 glutamine, Opti-MEM, GlutaMAX, 0.1mg/mL poly-D-lysine, and penicillin-streptomycin were
335 from ThermoFisher Scientific. FBS was from GeminiBio. Lipofectamine 2000 was from
336 Invitrogen. Rabbit anti-FLAG (D6W5B), Mouse anti-HA (6E2), Mouse anti-Beta Tubulin
337 (D3U1W), Rabbit anti-GAPDH (14C10), and HRP-conjugated Rabbit anti-GAPDH (D16H11) XP
338 antibodies were from Cell Signaling Technology. IRDye 800-conjugated Donkey Anti-Rabbit and
339 IRDye 680-conjugated Goat Anti-Mouse secondary antibodies were from LI-COR Biosciences.
340 To generate 647-conjugated Mouse anti-HA antibody, the antibody was purified from hybridoma
341 supernatant and chemically conjugated to Alexa Fluor 647 NHS Ester from ThermoFisher
342 Scientific. The Dual-Luciferase Reporter Assay System kit was from Promega. Coelenterazine
343 400a was from Nanolight Technology (Cat# 340).

344

345 *Plasmid DNA Constructs*

346 Detailed construct features and calculated molecular weights for proteins used in this study are
347 listed in Table S1.

348 All proteins were assembled using human aGPCR and G protein sequences retrieved from the
349 Uniprot database. Synthetic gene fragments (Integrated DNA Technologies) were inserted into
350 vectors using Gibson assembly, and plasmid sequences were confirmed by DNA sequencing.
351 Receptor constructs were subcloned into the pFuse-hIgG1-Fc2 vector (Invivogen) using EcoRI
352 and BglII restriction sites. All receptor constructs included a 3' stop codon preceding the hIgG1-
353 Fc coding sequence. An "empty vector" construct (referred to as "EV"; see Table S1) was
354 generated by placement of a stop codon immediately after the IL2-signal sequence of the pFuse-
355 hIgG1-Fc2 vector. A single glycine-serine linker was inserted between any epitope tag and the
356 coding sequence of the protein of interest. Point mutants were generated using a Q5 Site-Directed
357 Mutagenesis Kit (New England Biolabs).

358 We generated luciferase reporter plasmids containing individual response elements, a minimal
359 promoter (minP), and *luc2* firefly luciferase. These plasmids are identical to the

360 pGL4.29[CRE/minP/*luc2P*], pGL4.30[NFAT-RE/minP/*luc2P*], pGL4.33[SRE/*luc2P*],
361 pGL4.34[SRF-RE/*luc2P*] plasmids (Promega) except that *luc2* is used in place of *luc2P* in our
362 system and that minP is included in our SRE and SRF-RE plasmids. To assemble these plasmids,
363 synthetic gene fragments containing response elements were inserted into a pGL4.23[*luc2*/minP]
364 vector obtained from Promega using KpnI and HindIII (for SRE, SRF-RE, and CRE) or KpnI and
365 NcoI (for NFAT-RE, which included the minP in the synthetic gene fragment) restriction sites. The
366 pRL-TK plasmid (Promega) was used to express *Renilla* luciferase.

367 Tricistronic plasmids encoding TRUPATH G protein biosensors were a gift from Dr. Justin
368 English, University of Utah. These plasmids utilize identical protein sequences (for various Gα
369 subunits, Gβ1, and Gγ2) to those initially reported.

370

371 *Cell Culture*

372 Parental HEK293T and HEKΔ6 cells were cultured in DMEM supplemented with 10% FBS and
373 1% penicillin-streptomycin at 37°C in a 5% CO_2, humidified incubator.

374

375 *Reporter Gene Assays*

376 96-well clear bottom, white walled, tissue culture-treated plates were coated with 40 µL of 10
377 µg/mL aqueous poly-D-lysine solution for 15 minutes at room temperature followed by aspiration
378 to remove excess solution. Cells were seeded at a density of 3 x 10^5 per well in 100 µL complete
379 media (DMEM/10% FBS/1% penicillin-streptomycin) and incubated overnight (16-20 hours).
380 Cells in 50 µL fresh complete media were transfected by adding 20 µL of a solution containing
381 plasmid DNA and Lipofectamine 2000 at a ratio of 3 µL per 1 µg DNA in Opti-MEM.

382 For the CRE and NFAT-RE assays, each well was transfected with 50 ng firefly reporter plasmid,
383 1 ng pRL-TK plasmid, and 10 ng receptor plasmid. For the SRE and SRF-RE assays, each well
384 was transfected with 30 ng firefly reporter plasmid, 0.6 ng pRL-TK plasmid, and 30 ng receptor
385 plasmid. For receptor DNA titration experiments, total DNA was balanced with empty vector (see
386 Table S1).

387 After 4-6 hours, the transfection media was replaced with 100 µL complete media and cells
388 incubated overnight (16-20 hours). Media was then replaced with 100 µL complete media (CRE

389 and NFAT-RE) or serum-free DMEM supplemented with 1% penicillin-streptomycin (SRE and
390 SRF-RE), and cells were incubated for 6-8 hours. Cells were then washed with 100 μL PBS and
391 lysed by addition of 20 μL 1 x Passive Lysis Buffer (Promega) to each well and nutated for 15
392 minutes at room temperature. Lysates were pipetted up and down, and the assay plate was spun at
393 400 x g for 60 seconds to remove air bubbles. The ratio of firefly to *Renilla* luciferase activity was
394 measured using automated reagent injection with a GloMax luminometer system (Promega), in
395 which 50 μL LAR II and 25 μL Stop & Glo reagent (Promega) were added to each well.

396

397 *G Protein Biosensor Assays*

398 Cells were seeded and transfected using identical methods to those used for reporter gene assays.
399 Each well was transfected with 80 ng of biosensor plasmid and 20 ng of receptor plasmid. After
400 4-6 hours, the transfection media was replaced with 100 μL complete media and cells incubated
401 overnight. After 24 hours, media was replaced with 100 μL DMEM with no phenol red, glucose,
402 or glutamine supplemented with 2% GlutaMAX and 1% penicillin-streptomycin. 14-20 hours
403 later, media was replaced with Hank's Balanced Salt Solution with no phenol red, calcium, or
404 magnesium and supplemented with 20 mM HEPES pH 7.4 and 5 μM coelenterazine 400a from
405 NanoLight Technology (Cat# 340-5). An opaque white vinyl sticker was placed on the bottom of
406 the plate and BRET ratios were measured 2-8 minutes later using a GloMax luminometer system.
407 RLucII luminescence was measured using ET405/40x and ET510lp filters (Chroma Technology
408 Corp), for donor (*Renilla*) and acceptor (GFP) readouts, respectively.

409

410 *Cell Surface Staining and Flow Cytometry*

411 Quantification of cell surface staining by flow cytometry was performed using cells plated in non-
412 coated 96-well plates at a density of 3 x 10^5 per well in 100 μL complete media and incubated
413 overnight (16-20 hours). Then 50 μL media was removed from each well, and cells were
414 transfected by adding 20 μL of a solution containing 60 ng plasmid DNA and Lipofectamine 2000
415 at a ratio of 3 μL per 1 μg DNA in Opti-MEM. After 4-6 hours, 30 μL of fresh complete media was
416 added to each well, and cells were incubated overnight (18-20 hours). 24 hours post-transfection,

417 cells were resuspended and transferred to a 96-well V-bottom plate (Corning). Cells were pelleted
418 at 500 x g for 3 minutes, washed once with HEPES buffered saline (HBS), pH 7.4, and then rocked
419 at 4° C in 100 μL of HBS supplemented with 0.1% bovine serum albumin (HBS/BSA) and 647-
420 conjugated mouse anti-HA antibody at 1 μg/mL. Cells were rocked at 4° C for 20 minutes, washed
421 twice with HBS/BSA, and analyzed using a Beckman Coulter Cytoflex flow cytometer. Data were
422 analyzed with CytExpert software. Single cells were gated using forward and side scattering, and
423 α-HA surface staining was quantified using the MFI in the APC channel.

424

425 *Western Blot Analysis*

426 Cells for western blots were plated on poly-D-lysine-treated 24-well plates and transfected in
427 parallel with cells for reporter assays, using the same transfection mixtures (which included both
428 reporter plasmids and receptor and/or secTEV plasmids), and treated identically. Cell density,
429 media volumes, and DNA loads were increased five-fold relative to the amounts used in 96-well
430 format reporter assays to account for the larger well size. At the time of harvest, protein samples
431 for blots were reduced and denatured in Laemmli buffer, loaded into 4 –20% Tris-Glycine gels
432 (Bio-Rad) for SDS-PAGE, and transferred onto nitrocellulose membranes (Cytiva). Blots were
433 blocked with 5% milk in TBST and incubated with primary antibodies overnight at 4°C.
434 Simultaneous analysis of proteins of interest and loading controls were carried out for each blot
435 through the use of two imaging channels on a LI-COR Odyssey CLx Fluorescent Imaging System.
436 Blots were incubated with secondary antibodies for 30-60 minutes prior to imaging. For blots
437 simultaneously detecting FLAG, HA, and GAPDH (Fig. S4), FLAG and HA were detected using
438 the LI-COR system and HRP-conjugated GAPDH primary antibody was incubated with blots for
439 2 hours at room temperature and detected with Western Lightning Plus-ECL (PerkinElmer) using
440 a Bio-Rad Chemidoc System.

441

442 *Bioinformatics and Structural Model Generation*

443 To generate a homology model of the ADGRE5 CTF in the active signaling state (Fig. S3C), the
444 sequences of human ADGRE5 residues 531-835 and human $G\alpha_{13}$ residues 1-377 were submitted
445 to the ColabFold software server[40], which predicted a structure of the TA-agonized CTF in

complex with $G\alpha_{13}$. ColabFold implementation of AlphaFold installed on SBGrid was used to make structural predictions of all 32 human aGPCR CTF proteins[39,40,57]. All 5 models were run for 5 recycles with version 1.3.0, without templates and without amber relaxation. Multiple sequence alignments were generated using MMseqs2[58]. Models were ranked by the Predicted Template Modeling Score (pTM) and the top ranked model for each human adhesion GPCR was used for further analysis.

Heat Map Analysis and Scaling of Measured Responses

TA-dependent genetic reporter scores (for heat map and scatter plots) were quantified as Log1P-transformed differences in signal given by MBP-CTF and MBP-CTFΔTA proteins. Values were min-max scaled from zero to one within each assay. G protein biosensor scores were quantified as the ΔBRET value, linearly scaled from zero to the maximum value given in each respective biosensor for heat maps or from zero to one for scatter plots. Negative ΔBRET values were represented as zero in both metrics. Signaling score (in scatter plots) is the largest score amongst all reporter gene scores and G protein biosensor scores. TA-dependent surface depletion was quantified as the percentage decrease in surface staining signal given by MBP-CTF proteins when compared to the cognate MBP-CTFΔTA proteins. For scatter plots, this was scaled from zero (0% depletion) to one (100% depletion).

468 **Bibliography**

469 1. Schiöth HB, Lagerström MC. Structural diversity of G protein-coupled receptors and
470 significance for drug discovery. *Nat Rev Drug Discov*. 2008;7(4):339-357.
471 doi:10.1038/nrd2518
472 2. Ushkaryov YA, Lelianova V, Vysokov N v. Catching Latrophilin With Lasso: A Universal
473 Mechanism for Axonal Attraction and Synapse Formation. *Front Neurosci*. 2019;13.
474 doi:10.3389/fnins.2019.00257
475 3. Arcos-Burgos M, Jain M, Acosta MT, et al. A common variant of the latrophilin 3 gene,
476 LPHN3, confers susceptibility to ADHD and predicts effectiveness of stimulant
477 medication. *Mol Psychiatry*. 2010;15(11):1053-1066. doi:10.1038/mp.2010.6
478 4. Yona S, Lin H, Dri P, et al. Ligation of the adhesion-GPCR EMR2 regulates human
479 neutrophil function. *The FASEB Journal*. 2008;22(3):741-751. doi:10.1096/fj.07-9435com
480 5. Bolliger MF, Martinelli DC, Südhof TC. The cell-adhesion G protein-coupled receptor
481 BAI3 is a high-affinity receptor for C1q-like proteins. *Proc Natl Acad Sci U S A*.
482 2011;108(6):2534-2539. doi:10.1073/pnas.1019577108
483 6. Langenhan T, Piao X, Monk KR. Adhesion G protein-coupled receptors in nervous system
484 development and disease. *Nat Rev Neurosci*. 2016;17(9):550-561.
485 doi:10.1038/nrn.2016.86
486 7. McGee JA, Goodyear RJ, McMillan DR, et al. The very large G-protein-coupled receptor
487 VLGR1: A component of the ankle link complex required for the normal development of
488 auditory hair bundles. *Journal of Neuroscience*. 2006;26(24):6543-6553.
489 doi:10.1523/JNEUROSCI.0693-06.2006
490 8. Monk KR, Oshima K, Jörs S, Heller S, Talbot WS. Gpr126 is essential for peripheral
491 nerve development and myelination in mammals. *Development*. 2011;138(13):2673-2680.
492 doi:10.1242/dev.062224
493 9. Piao X, Hill SS, Bodell A, et al. G protein-coupled receptor-dependent development of
494 human frontal cortex. *Science*. 2004;303(5666):2033-2036.
495 doi:10.1126/SCIENCE.1092780
496 10. Fredriksson R, Lagerström MC, Lundin LG, Schiöth HB. The G-protein-coupled receptors
497 in the human genome form five main families. Phylogenetic analysis, paralogon groups,
498 and fingerprints. *Mol Pharmacol*. 2003;63(6):1256-1272. doi:10.1124/MOL.63.6.1256
499 11. Purcell RH, Hall RA. Adhesion G Protein-Coupled Receptors as Drug Targets. *Annu Rev*
500 *Pharmacol Toxicol*. 2018;58:429-449. doi:10.1146/ANNUREV-PHARMTOX-010617-
501 052933
502 12. Langenhan T. Adhesion G protein–coupled receptors—Candidate metabotropic
503 mechanosensors and novel drug targets. *Basic Clin Pharmacol Toxicol*. 2020;126(S6):5-
504 16. doi:10.1111/BCPT.13223
505 13. Lala T, Hall RA. Adhesion G Protein-Coupled Receptors: Structure, Signaling,
506 Physiology, and Pathophysiology. *Physiol Rev*. 2022;102(4):1587-1624. doi:
507 10.1152/physrev.00027.2021
508 14. Vizurraga A, Adhikari R, Yeung J, Yu M, Tall GG. Mechanisms of adhesion G protein–
509 coupled receptor activation. *Journal of Biological Chemistry*. 2020;295(41):14065-14083.
510 doi:10.1074/jbc.REV120.007423
511 15. Inoue A, Raimondi F, Kadji FMN, et al. Illuminating G-Protein-Coupling Selectivity of
512 GPCRs. *Cell*. 2019;177(7):1933-1947.e25. doi:10.1016/J.CELL.2019.04.044

513 16. Avet C, Mancini A, Breton B, et al. Effector membrane translocation biosensors reveal G
514 protein and βarrestin coupling profiles of 100 therapeutically relevant GPCRs. *Elife*.
515 2022;11. doi:10.7554/ELIFE.74101
516 17. Kroeze WK, Sassano MF, Huang XP, et al. PRESTO-Tango as an open-source resource
517 for interrogation of the druggable human GPCRome. *Nature Publishing Group*. Published
518 online 2015. doi:10.1038/nsmb.3014
519 18. Masuho I, Kise R, Gainza P, et al. Rules and mechanisms governing G protein coupling
520 selectivity of GPCRs. *Cell Rep*. 2023;42(10):113173.
521 doi:10.1016/J.CELREP.2023.113173
522 19. Isberg V, Vroling B, Van Der Kant R, Li K, Vriend G, Gloriam D. GPCRDB: An
523 information system for G protein-coupled receptors. *Nucleic Acids Res*. 2014;42(D1).
524 doi:10.1093/NAR/GKT1255
525 20. Araç D, Leon K. Structure, function and therapeutic potential of adhesion GPCRs.
526 Published online 2020. doi:10.1016/B978-0-12-816228-6.00002-7
527 21. Araç D, Boucard AA, Bolliger MF, et al. A novel evolutionarily conserved domain of cell-
528 adhesion GPCRs mediates autoproteolysis. *EMBO Journal*. 2012;31(6):1364-1378.
529 doi:10.1038/emboj.2012.26
530 22. Prömel S, Langenhan T, Araç D. Matching structure with function: The GAIN domain of
531 Adhesion-GPCR and PKD1-like proteins. *Trends Pharmacol Sci*. 2013;34(8):470-478.
532 doi:10.1016/j.tips.2013.06.002
533 23. Salzman GS, Ackerman SD, Ding C, et al. Structural basis for regulation of
534 GPR56/ADGRG1 by its alternatively spliced extracellular domains. *Neuron*.
535 2016;91(6):1292. doi:10.1016/J.NEURON.2016.08.022
536 24. Leon K, Cunningham RL, Riback JA, et al. Structural basis for adhesion G protein-
537 coupled receptor Gpr126 function. *Nature Communications 2020 11:1*. 2020;11(1):1-14.
538 doi:10.1038/s41467-019-14040-1
539 25. Chu TY, Zheng-Gérard C, Huang KY, et al. GPR97 triggers inflammatory processes in
540 human neutrophils via a macromolecular complex upstream of PAR2 activation. *Nat*
541 *Commun*. 2022;13(1). doi:10.1038/S41467-022-34083-1
542 26. Stoveken HM, Hajduczok AG, Xu L, Tall GG. Adhesion G protein-coupled receptors are
543 activated by exposure of a cryptic tethered agonist. *Proc Natl Acad Sci U S A*.
544 2015;112(19):6194-6199. doi:10.1073/pnas.1421785112
545 27. Liebscher I, Schön J, Petersen SC, et al. A Tethered Agonist within the Ectodomain
546 Activates the Adhesion G Protein-Coupled Receptors GPR126 and GPR133. *Cell Rep*.
547 2014;9(6):2018-2026. doi:10.1016/j.celrep.2014.11.036
548 28. Mathiasen S, Palmisano T, Perry NA, et al. G12/13 is activated by acute tethered agonist
549 exposure in the adhesion GPCR ADGRL3. *Nat Chem Biol*. 2020;16(12):1343-1350.
550 doi:10.1038/s41589-020-0617-7
551 29. Wilde C, Fischer L, Lede V, et al. The constitutive activity of the adhesion GPCR
552 GPR114/ADGRG5 is mediated by its tethered agonist. *FASEB Journal*. 2016;30(2):666-
553 673. doi:10.1096/fj.15-276220
554 30. Demberg LM, Winkler J, Wilde C, et al. Activation of adhesion G protein-coupled
555 receptors: Agonist specificity of Stachel sequence-derived peptides. *Journal of Biological*
556 *Chemistry*. 2017;292(11):4383-4394. doi:10.1074/jbc.M116.763656
557 31. Peeters MC, Fokkelman M, Boogaard B, et al. The adhesion G protein-coupled receptor
558 G2 (ADGRG2/GPR64) constitutively activates SRE and NFκB and is involved in cell

559 adhesion and migration. *Cell Signal*. 2015;27(12):2579-2588.
560 doi:10.1016/j.cellsig.2015.08.015
561 32. Kishore A, Purcell RH, Nassiri-Toosi Z, Hall RA. Stalk-dependent and stalk-independent
562 signaling by the adhesion G protein-coupled receptors GPR56 (ADGRG1) and BAI1
563 (ADGRB1). *Journal of Biological Chemistry*. 2016;291(7):3385-3394.
564 doi:10.1074/jbc.M115.689349
565 33. Demberg LM, Rothemund S, Schöneberg T, Liebscher I. Identification of the tethered
566 peptide agonist of the adhesion G protein-coupled receptor GPR64/ADGRG2. *Biochem*
567 *Biophys Res Commun*. 2015;464(3):743-747. doi:10.1016/J.BBRC.2015.07.020
568 34. Barros-Álvarez X, Nwokonko RM, Vizurraga A, et al. The tethered peptide activation
569 mechanism of adhesion GPCRs. *Nature 2022 604:7907*. 2022;604(7907):757-762.
570 doi:10.1038/s41586-022-04575-7
571 35. Xiao P, Guo S, Wen X, et al. Tethered peptide activation mechanism of the adhesion
572 GPCRs ADGRG2 and ADGRG4. *Nature 2022 604:7907*. 2022;604(7907):771-778.
573 doi:10.1038/s41586-022-04590-8
574 36. Ping YQ, Xiao P, Yang F, et al. Structural basis for the tethered peptide activation of
575 adhesion GPCRs. *Nature 2022 604:7907*. 2022;604(7907):763-770. doi:10.1038/s41586-
576 022-04619-y
577 37. Qu X, Qiu N, Wang M, et al. Structural basis of tethered agonism of the adhesion GPCRs
578 ADGRD1 and ADGRF1. *Nature 2022 604:7907*. 2022;604(7907):779-785.
579 doi:10.1038/s41586-022-04580-w
580 38. Wang N, Qian Y, Xia R, et al. Structural basis of CD97 activation and G-protein coupling.
581 *Cell Chem Biol*. Published online September 2023.
582 doi:10.1016/J.CHEMBIOL.2023.08.003
583 39. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with
584 AlphaFold. *Nature 2021 596:7873*. 2021;596(7873):583-589. doi:10.1038/s41586-021-
585 03819-2
586 40. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold:
587 making protein folding accessible to all. *Nature Methods 2022 19:6*. 2022;19(6):679-682.
588 doi:10.1038/s41592-022-01488-1
589 41. Hilbig D, Sittig D, Hoffmann F, et al. Mechano-Dependent Phosphorylation of the PDZ-
590 Binding Motif of CD97/ADGRE5 Modulates Cellular Detachment. *Cell Rep*.
591 2018;24(8):1986-1995. doi:10.1016/j.celrep.2018.07.071
592 42. Ward Y, Lake R, Yin JJ, et al. LPA receptor heterodimerizes with CD97 to amplify LPA-
593 initiated RHO-dependent signaling and invasion in prostate cancer cells. *Cancer Res*.
594 2011;71(23):7301-7311. doi:10.1158/0008-5472.CAN-11-2381
595 43. Liu D, Duan L, Rodda LB, et al. CD97 promotes spleen dendritic cell homeostasis
596 through the mechanosensing of red blood cells. *Science (1979)*. 2022;375(6581).
597 doi:10.1126/science.abi5965
598 44. Cheng Z, Garvin D, Paguio A, Stecha P, Wood K, Fan F. Luciferase Reporter Assay
599 System for Deciphering GPCR Pathways. *Curr Chem Genomics*. 2010;4:84.
600 doi:10.2174/1875397301004010084
601 45. Jones DTD, Dates AN, Rawson SD, Burruss MM, Lipper CH, Blacklow SC. Tethered
602 agonist activated ADGRF1 structure and signalling analysis reveal basis for G protein
603 coupling. *Nature Communications 2023 14:1*. 2023;14(1):1-13. doi:10.1038/s41467-023-
604 38083-7

605 46. Olsen RHJ, DiBerto JF, English JG, et al. TRUPATH, an open-source biosensor platform
606 for interrogating the GPCR transducerome. *Nature Chemical Biology 2020 16:8*.
607 2020;16(8):841-849. doi:10.1038/s41589-020-0535-8
608 47. Rajagopal S, Shenoy SK. GPCR Desensitization: Acute and Prolonged Phases. *Cell*
609 *Signal*. 2018;41:9. doi:10.1016/J.CELLSIG.2017.01.024
610 48. Cary BP, Zhang X, Cao J, et al. New Insights into the Structure and Function of Class B1
611 GPCRs. *Endocr Rev*. 2023;44(3):492-517. doi:10.1210/ENDREV/BNAC033
612 49. Cesaratto F, López-Requena A, Burrone OR, Petris G. Engineered tobacco etch virus
613 (TEV) protease active in the secretory pathway of mammalian cells. *J Biotechnol*.
614 2015;212:159-166. doi:10.1016/J.JBIOTEC.2015.08.026
615 50. I KY, Huang YS, Hu CH, et al. Activation of adhesion GPCR EMR2/ADGRE2 induces
616 macrophage differentiation and inflammatory responses via Gα16/Akt/MAPK/NF-κB
617 signaling pathways. *Front Immunol*. 2017;8(APR). doi:10.3389/fimmu.2017.00373
618 51. Smith JS, Lefkowitz RJ, Rajagopal S. Biased signalling: from simple switches to allosteric
619 microprocessors. *Nature Reviews Drug Discovery 2018 17:4*. 2018;17(4):243-260.
620 doi:10.1038/nrd.2017.229
621 52. Irannejad R, Tomshine JC, Tomshine JR, et al. Conformational biosensors reveal GPCR
622 signalling from endosomes. *Nature*. 2013;495(7442):534-538.
623 doi:10.1038/NATURE12000
624 53. Liebscher I, Schöneberg T, Thor D. Stachel-mediated activation of adhesion G protein-
625 coupled receptors: insights from cryo-EM studies. *Signal Transduction and Targeted*
626 *Therapy 2022 7:1*. 2022;7(1):1-2. doi:10.1038/s41392-022-01083-y
627 54. Prömel S, Frickenhaus M, Hughes S, et al. The GPS Motif Is a Molecular Switch for
628 Bimodal Activities of Adhesion Class G Protein-Coupled Receptors. *Cell Rep*.
629 2012;2(2):321-331. doi:10.1016/j.celrep.2012.06.015
630 55. Suzuki N, Hajicek N, Kozasa T. Regulation and Physiological Functions of G12/13-
631 Mediated Signaling Pathways. *Neurosignals*. 2009;17(1):55. doi:10.1159/000186690
632 56. Worzfeld T, Wettschureck N, Offermanns S. G 12 /G 13-mediated signalling in
633 mammalian physiology and disease. doi:10.1016/j.tips.2008.08.002
634 57. Morin A, Eisenbraun B, Key J, et al. Collaboration gets the most out of software. *Elife*.
635 2013;2013(2). doi:10.7554/ELIFE.01456
636 58. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the
637 analysis of massive data sets. *Nature Biotechnology 2017 35:11*. 2017;35(11):1026-1028.
638 doi:10.1038/nbt.3988
639
640

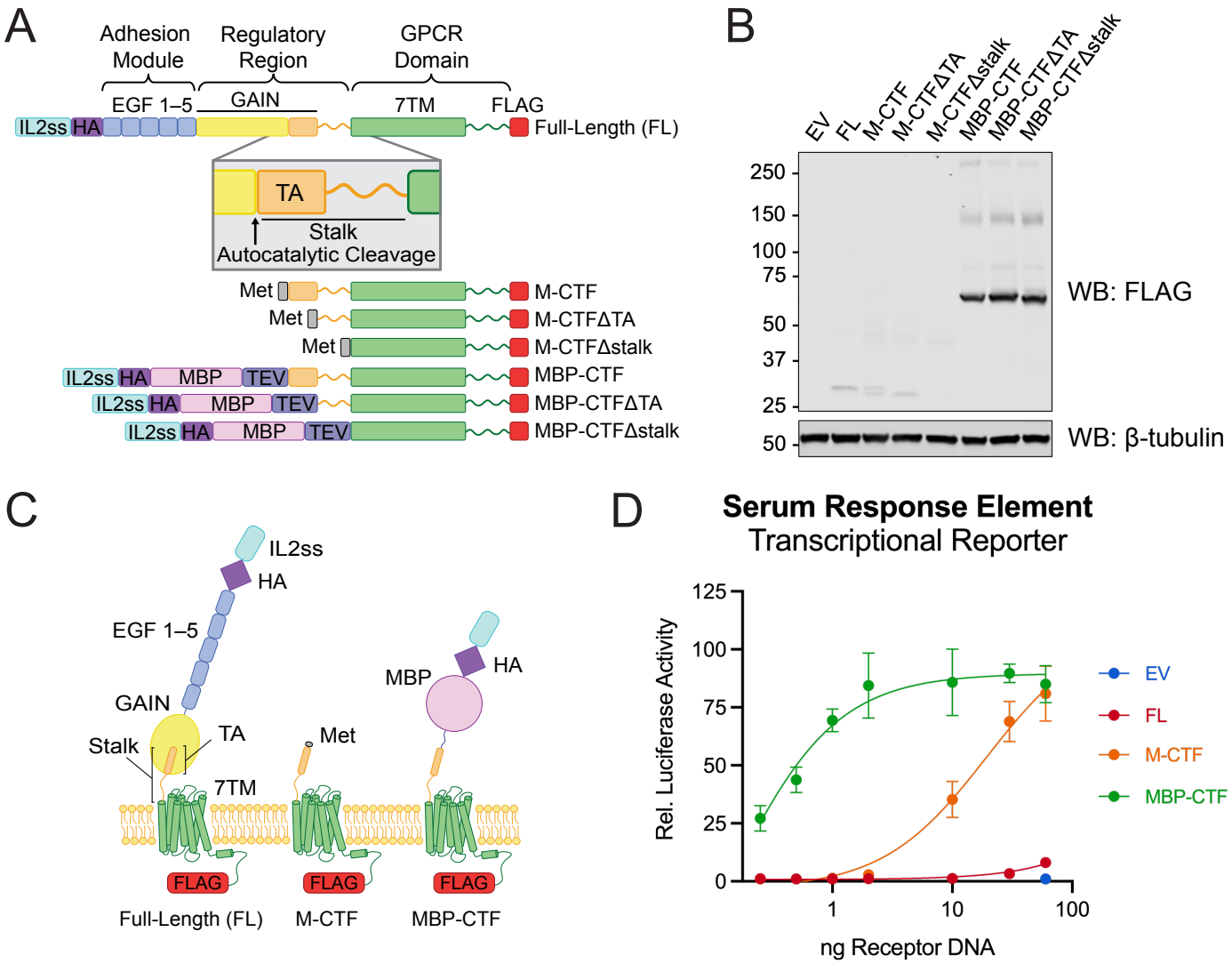


Fig. 1 – A fusion protein system to evaluate aGPCR signaling. (A) Domain architectures of constructs used to evaluate signaling in ADGRE5. In the full-length (FL) protein, autocatalytic cleavage occurs within the GAIN domain, immediately prior to the TA peptide, which is sequestered within the folded GAIN domain. C-terminal fragment proteins with exposed TA peptides are expressed with an N-terminal methionine (M-CTF) or after a signal sequence, HA epitope tag, and maltose-binding protein fusion domain (MBP-CTF). The design of the ΔTA and Δstalk variants is also shown. (B) Western blot detection of the FLAG C-terminal epitope tag to assess total protein abundance, with β-tubulin as a protein loading control. (C) Schematic diagrams of ADGRE5 signaling constructs. (D) Comparison of ADGRE5 variants in the SRE transcriptional reporter signaling assay as a function of the amount of transfected DNA. All data points are normalized to the mean value for empty vector (60 ng). Dose-response curves were fit for each titration curve using nonlinear regression. Data are presented as mean ± SEM of three independent biological replicates (n=3).

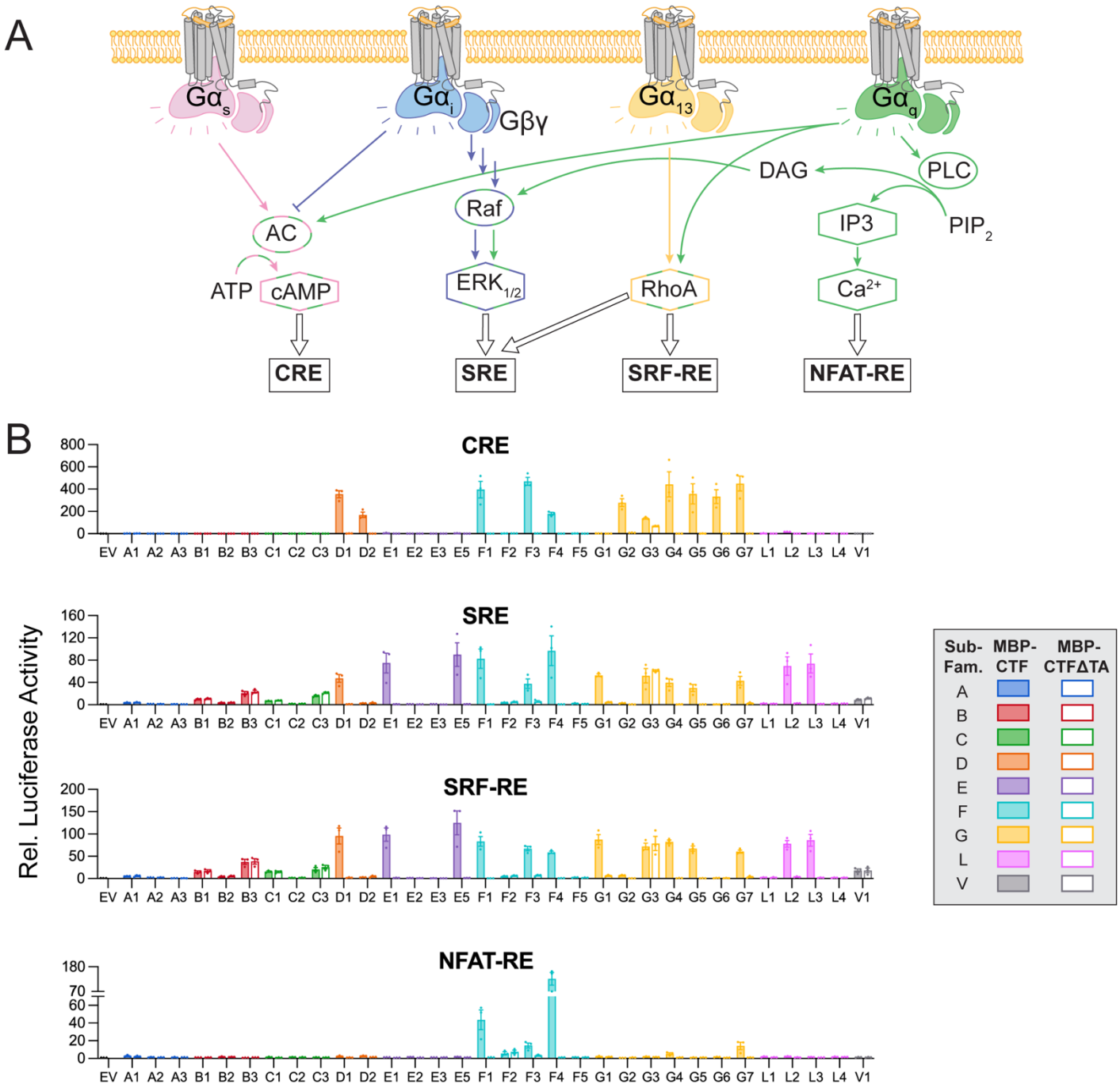


Fig. 2 – Transcriptional reporter responses for all human aGPCR family members. (A) CRE, SRE, SRF-RE, and NFAT-RE transcriptional reporter gene assays, illustrating potential sources of stimulation downstream of G protein signaling. (B) CRE, SRE, SRF-RE, and NFAT-RE reporter activity of cells expressing either MBP-CTF (filled bars) or MBP-CTFΔTA (open bars). Data are normalized to the mean value for empty vector and are presented as mean ± SEM of three independent biological replicates (n=3).

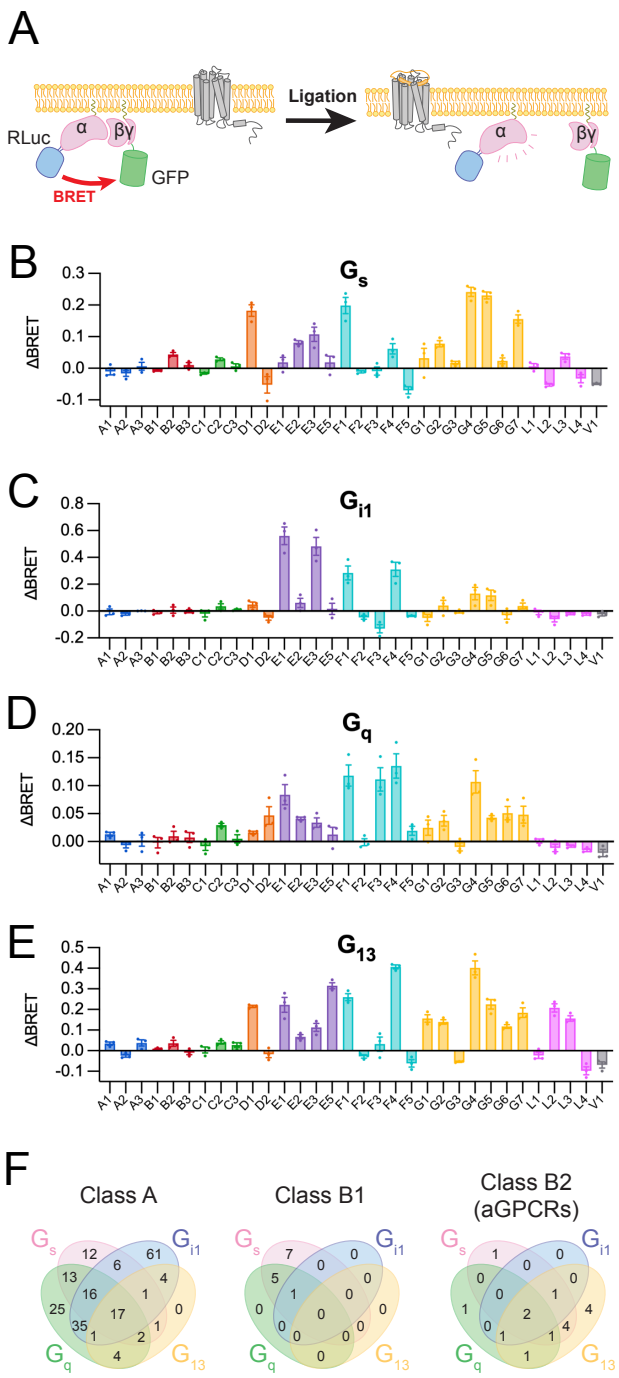
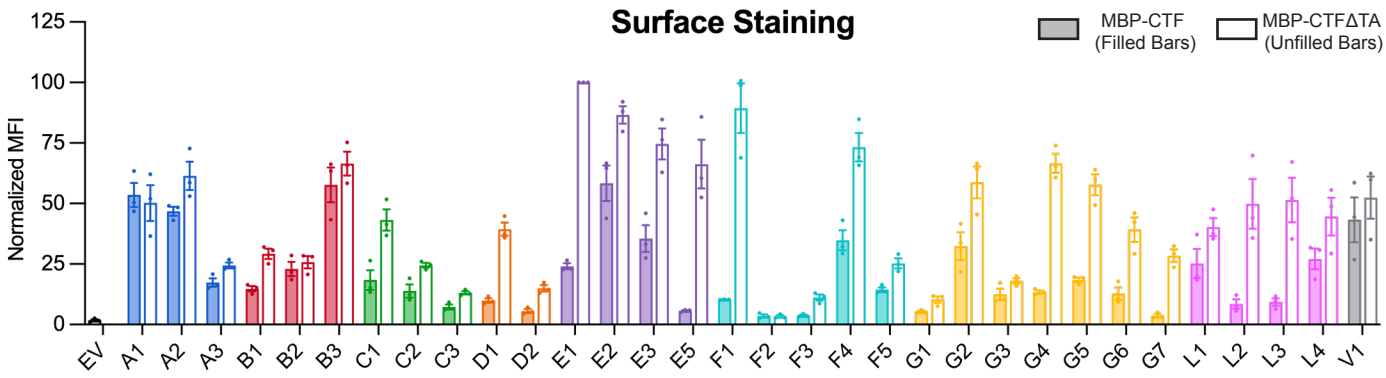


Fig. 3 – Tethered agonist-dependent G protein activation activity of all human aGPCRs (A) Schematic depicting G protein biosensor assays, which measure the dissociation of Gα subunits from Gβγ complexes. (B–E) TA-dependent activation of the G_s (B), G_{i1} (C), G_q (D), and G_{13} (E) biosensors for all aGPCRs. ΔBRET values are the difference in resting BRET ratios between MBP-CTF and MBP-CTFΔTA for each protein pair. Data are presented as mean ± SEM of three independent biological replicates (n=3). (F) Summaries of G protein biosensor activation for family A[19], B1[19], and B2 GPCRs. Family B2 signals were assigned using one-way ANOVA with Tukey's multiple-comparison post-hoc test to compare ΔBRET values for each receptor with that of EV (zero) (* P<0.05).



674
675 **Fig. 4 – Cell surface amounts of all aGPCR protein constructs.** α-HA surface staining of
676 non-permeabilized HEK293T cells expressing HA-MBP-CTF-FLAG (filled bars) and HA-MBP-
677 CTFΔTA-FLAG (open bars) proteins. Data from three biological replicates (n=3) are presented
678 as mean fluorescence intensity (MFI) ± SEM, normalized to the value for ADGRE1, which was
679 set to 100.
680

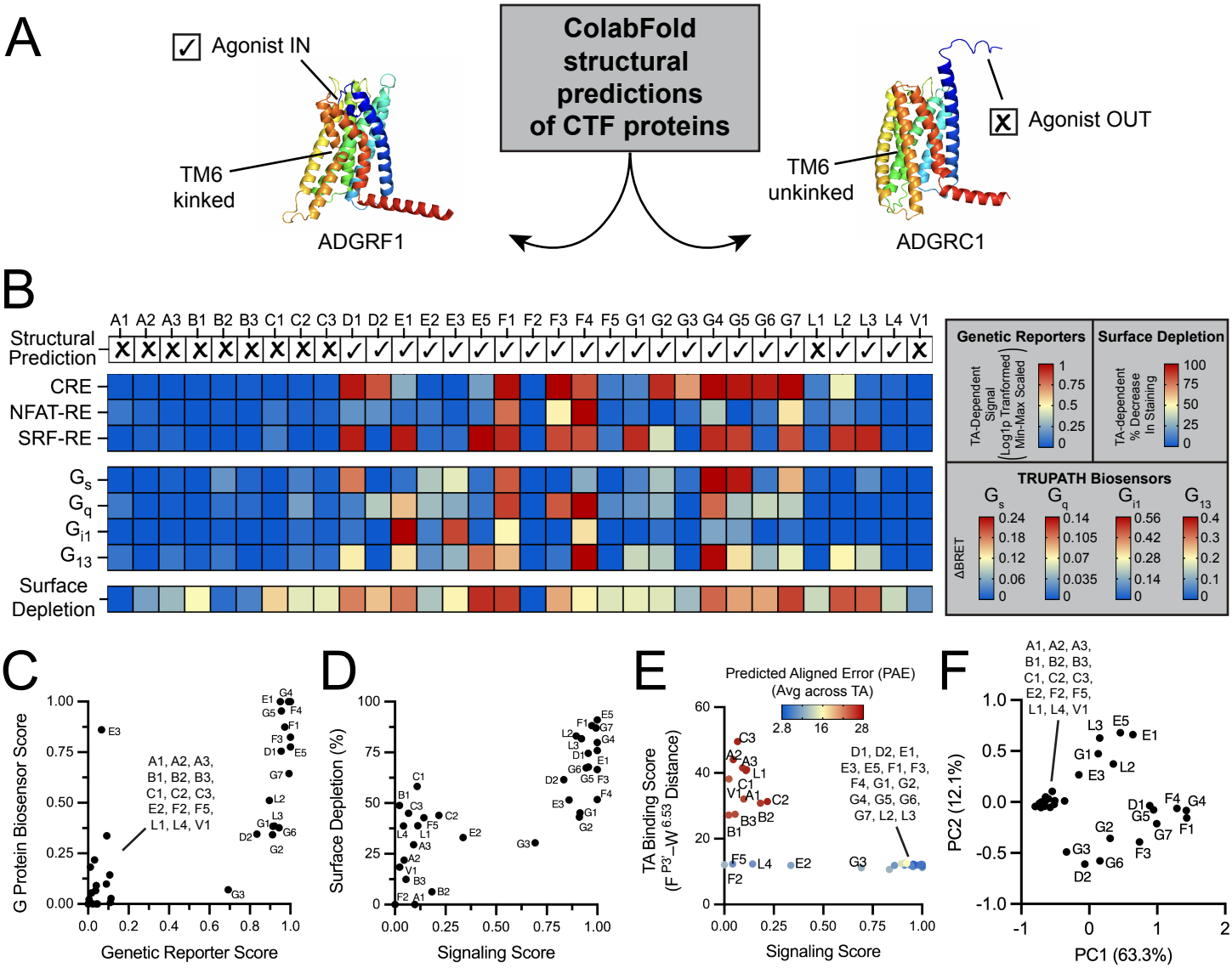


Fig. 5 – Structural predictions and heat map representations of responses for all human aGPCRs. (A) Representative examples of CTF "in" and "out" conformations predicted by ColabFold, showing the position of the TA peptide and degree of kinking by TM6, which is a hallmark of Family B GPCR activation. (B) Heat map summarizing TA-dependent transcriptional activation, G protein activation, surface localization, and ColabFold predictions for the human aGPCRs. (C–E) Scatter plots correlating G protein biosensor score with genetic reporter score (C), surface depletion with signaling score (including all reporters and G protein biosensors) (D), and ColabFold-generated TA binding scores with signaling score (E) for each aGPCR. Points in (E) are colored by the Predicted Aligned Error confidence metric averaged across the seven TA residues for each receptor. (F) Principal Component Analysis of all measured TA-dependent responses when each measured response is scaled zero to one across all aGPCRs evaluated in this study.

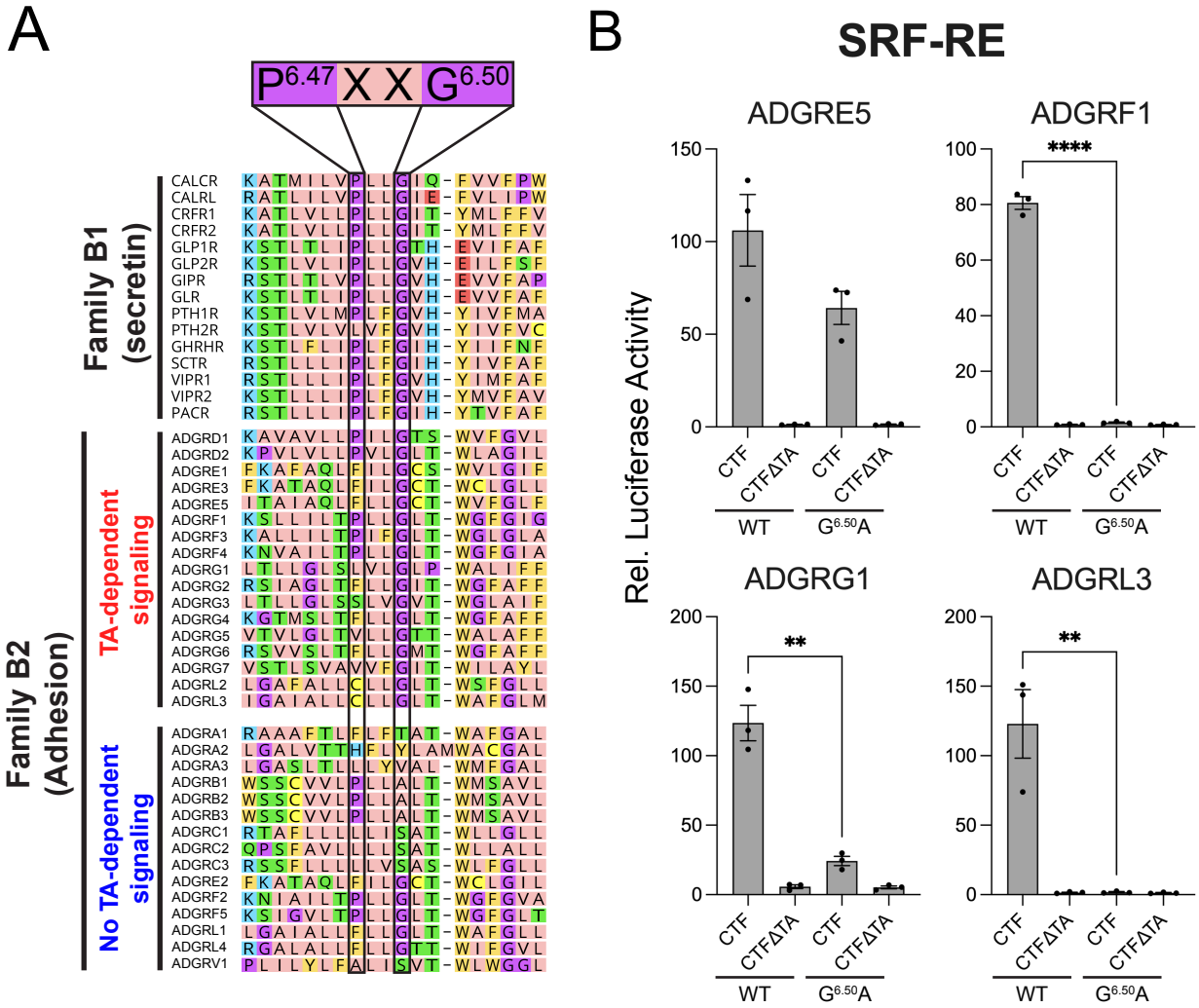


Fig. 6 – Conservation and functional relevance of the PXXG activation motif in aGPCRs. (A) Multiple sequence alignment of TM6 across human Family B GPCRs, with aGPCRs segregated by TA-signaling competence in our MBP fusion system. (B) Effects of alanine substitution at $G^{6.50}$ within the PXXG motif of four TA-activated aGPCRs in the SRF-RE reporter. Data are normalized to the mean value for empty vector and are presented as mean ± SEM of three independent biological replicates (n=3). Unpaired two-tailed T-tests were performed to compare WT and mutant CTF proteins (** P<0.01, **** P<0.0001).

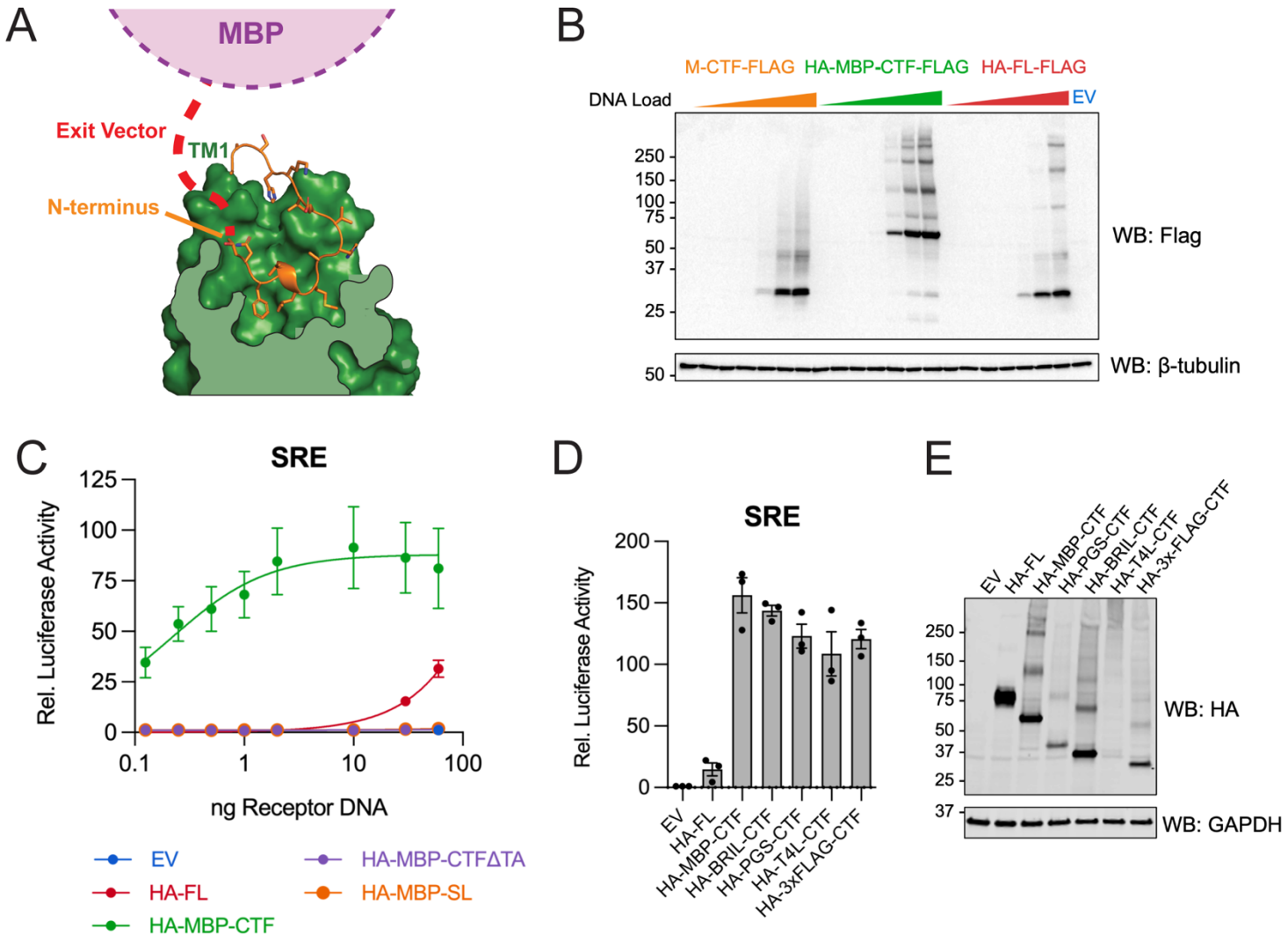


Fig. S1 –Signaling properties of CTF fusion proteins. (A) Binding pose of the TA peptide within the 7TM domain of ADGRG1 (PDB ID: 7SF8)[34]. The presence of a potential exit vector path provides a physical explanation for the signaling capacity of CTF molecules with N-terminal fusions. (B) α-FLAG western blot analysis of protein abundance for molecules evaluated in the dose-response analysis of signaling in the SRE assay (Fig. 1D). Individual lanes correspond to DNA amounts used, with β-tubulin as a protein loading control. (C) Comparison of the indicated ADGRE5-derived proteins in the SRE assay as a function of the amount of transfected DNA. (D) SRE signaling activity for a panel of different N-terminal sequences fused to the ADGRE5 CTF. (E) Western blot detection of N-terminal HA epitope tags for fusion proteins evaluated in (D), with GADPH as a protein loading control. For panels C and D, all data points are normalized to the mean value for empty vector and presented as mean ± SEM of three independent biological replicates (n=3). In panel C, dose-response curves were fit to each titration curve with nonlinear regression.

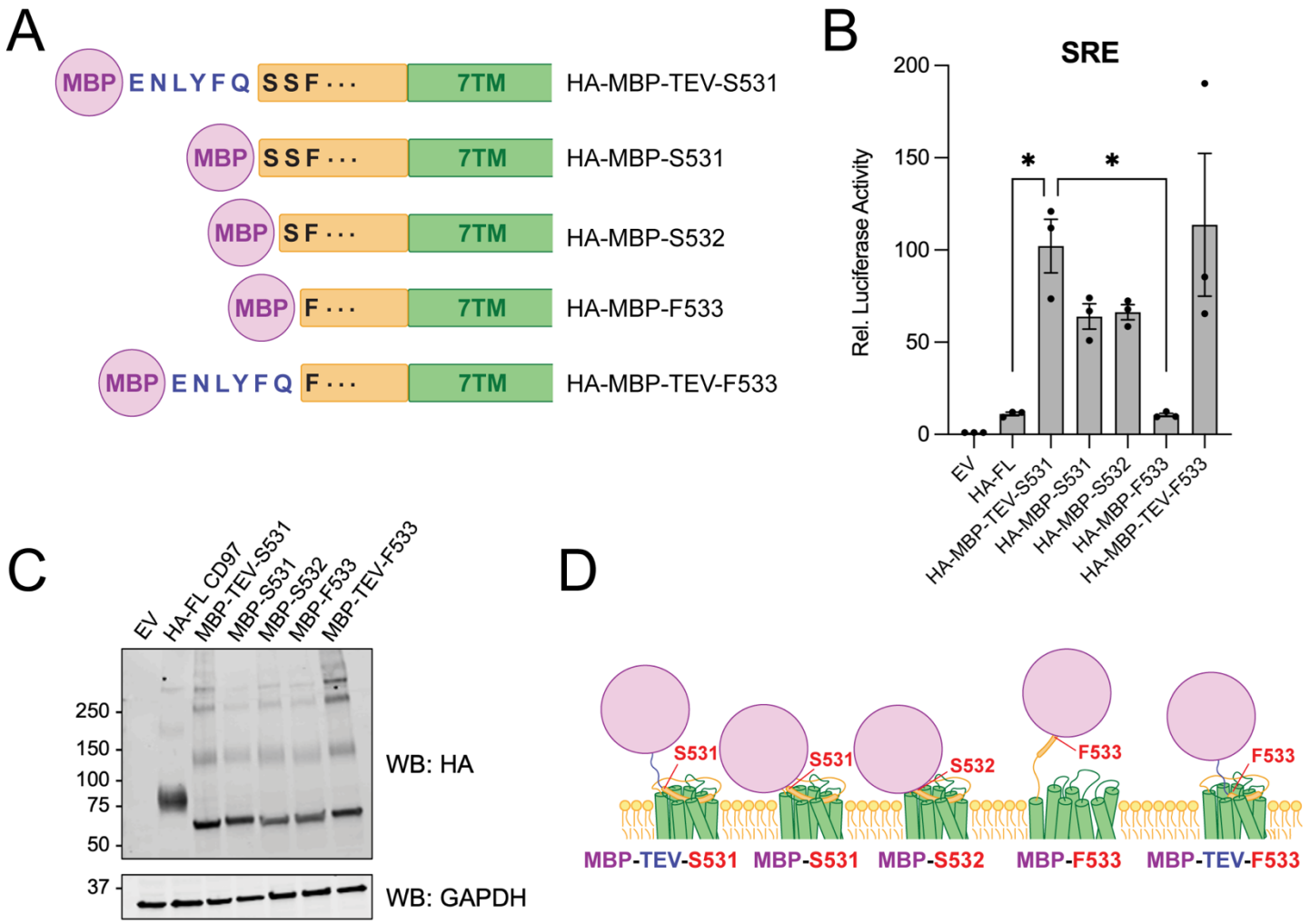


Fig. S2 – Identification of the closest MBP fusion tolerable for autonomous aGPCR signaling in ADGRE5. (A) Schematic diagrams depicting constructs in which MBP is fused at decreasing distances from the TA. (B) Signaling response of MBP fusion constructs shown in (A). Data are normalized to the mean value for empty vector and presented as mean ± SEM of three independent biological replicates (n=3). One-way ANOVA was used with Tukey's multiple-comparison post-hoc test to compare conditions (* $P<0.05$). (C) Western blot detection of N-terminal HA epitope tags for proteins evaluated in (B). Blotting for GAPDH was used to control for protein loading. (D) Model illustrating the variation in steric interference depending on the fusion position of MBP.

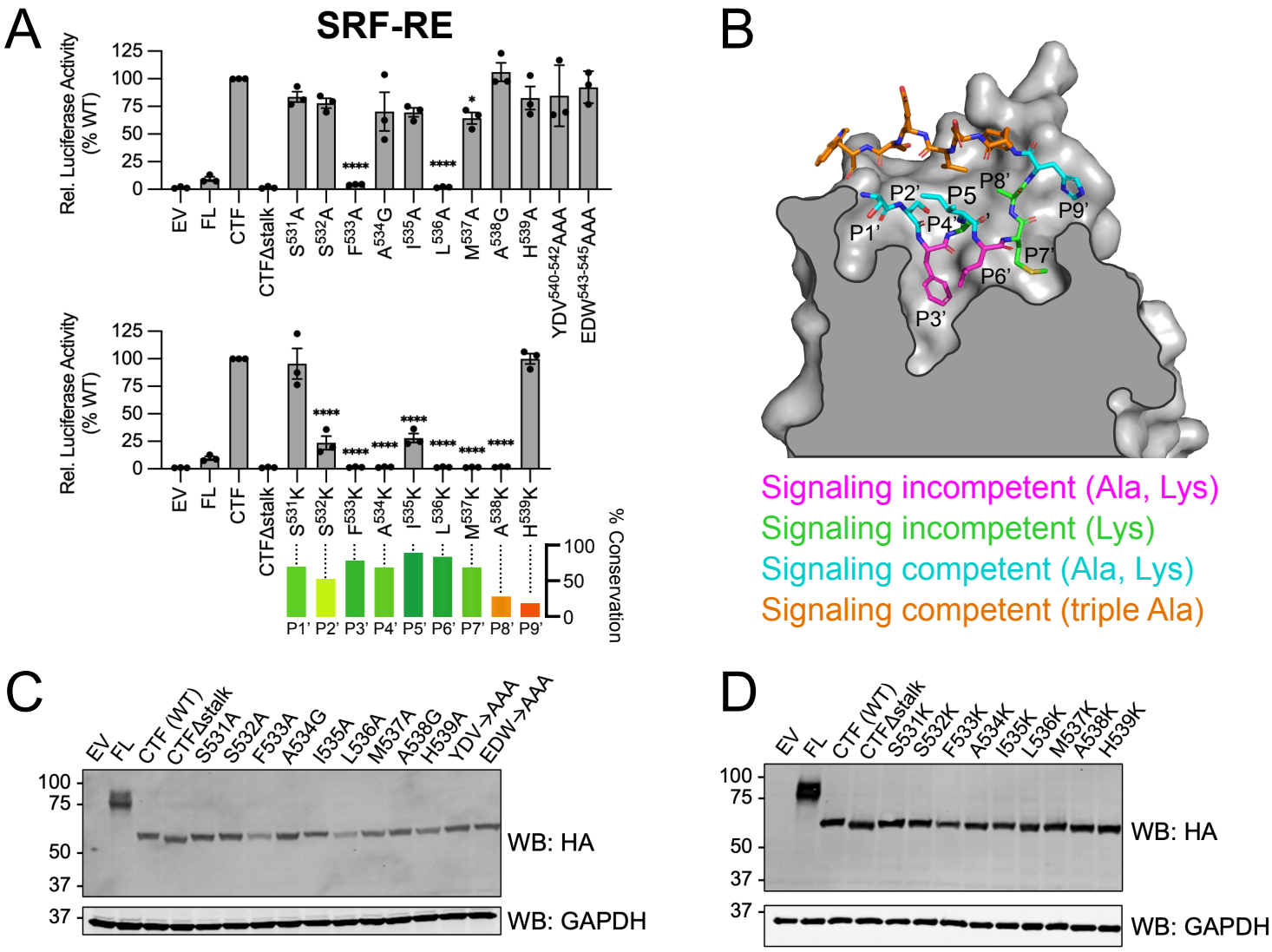


Fig. S3 – Scanning mutagenesis evaluation of the ADGRE5 TA sequence in the MBP-CTF scaffold. (A) SRE signaling data for scanning alanine (or glycine for A534 and A538) and lysine mutagenesis of the stalk sequence in ADGRE5. All mutations were made in the context of parental MBP-CTF protein. Data are normalized as a percentage of the mean value for WT and presented as mean ± SEM of three independent biological replicates (n=3). Sequence conservation is shown below and based on alignment of the 33 human aGPCRs. One-way ANOVA was used with Tukey's multiple-comparison post-hoc test to compare each mutant to wild-type protein (CTF) (* $P<0.05$, **** $P<0.0001$). (B) ColabFold-generated structural prediction of the TA-bound 7TM domain of ADGRE5 in complex with $G\alpha_{13}$[40]. The orthosteric pocket on the extracellular face of the receptor is depicted, with the bound stalk sequence colored according to mutagenesis results. (C,D) Western blot detection of N-terminal HA epitope tags for constructs evaluated in (A). Blotting for GAPDH was used to control for protein loading.

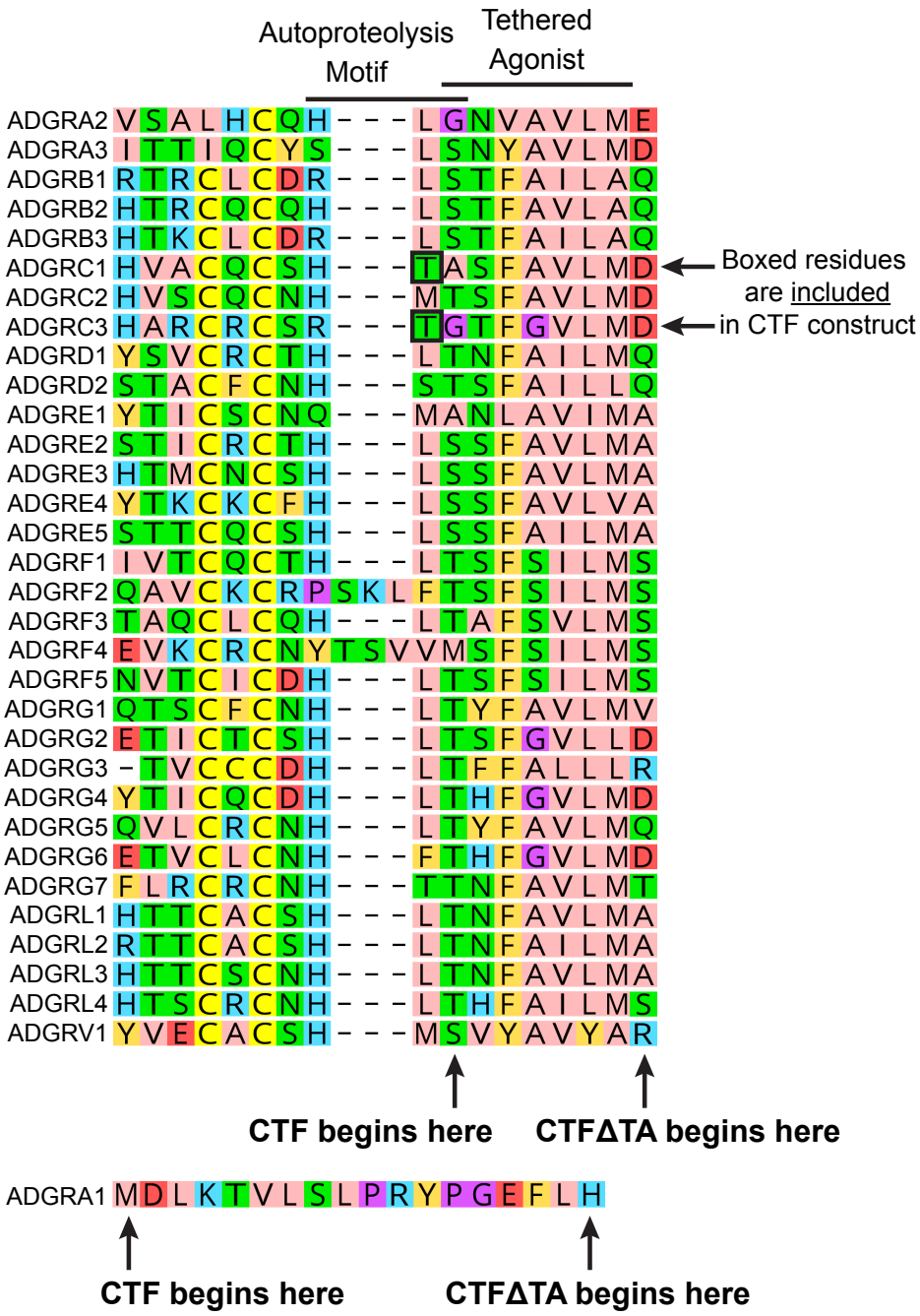
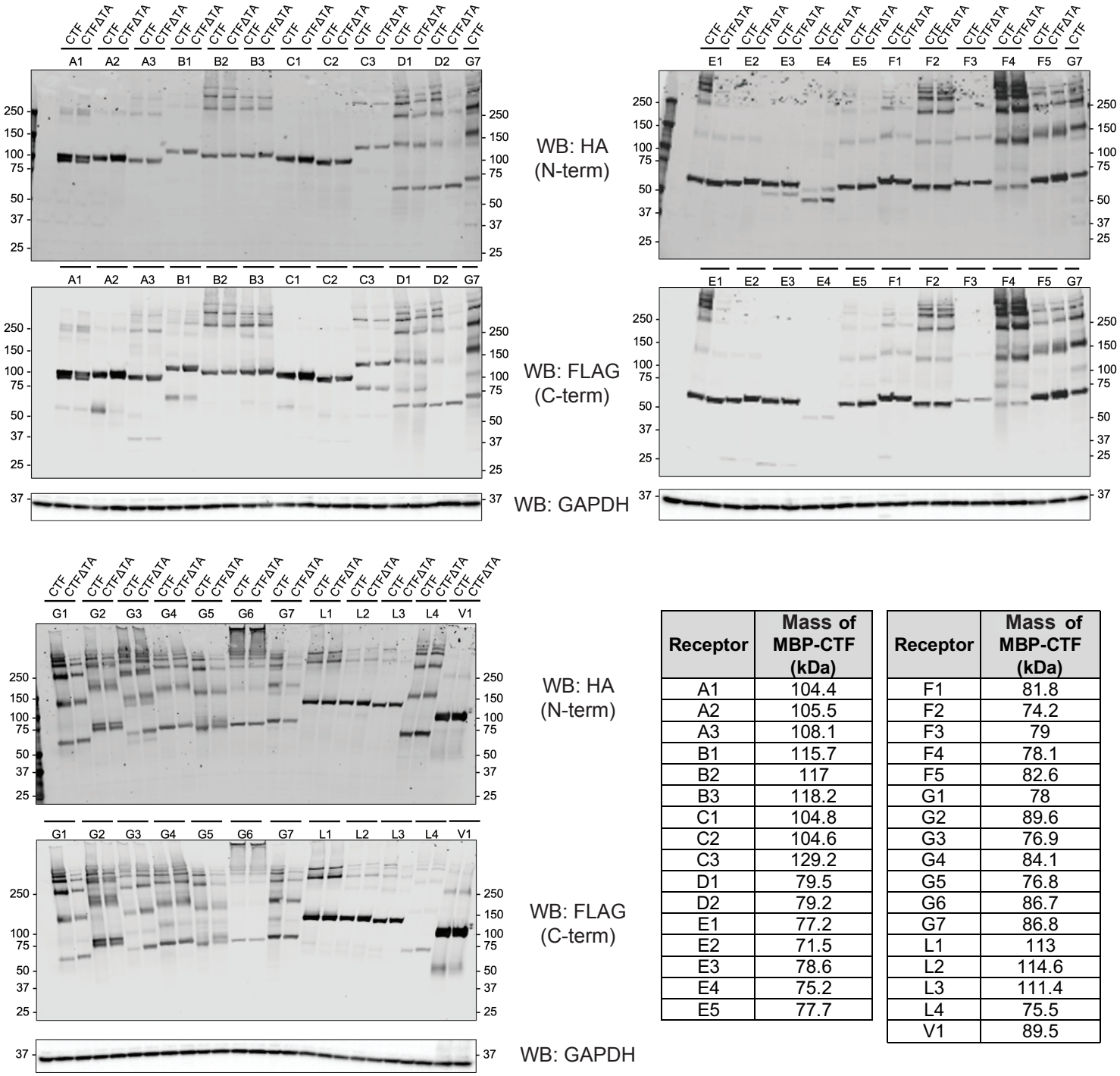


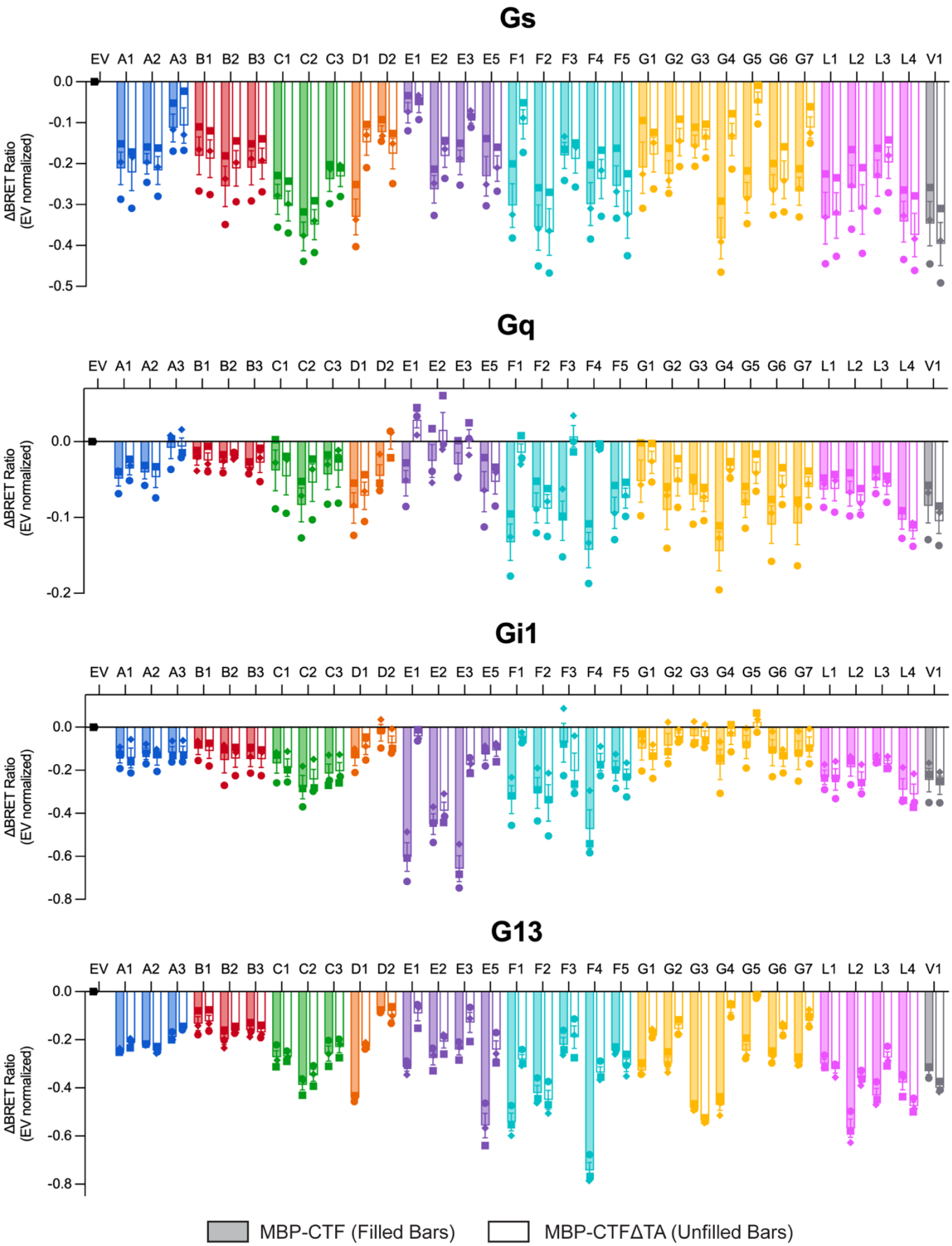
Fig. S4 – Design of CTF and CTFΔTA protein constructs. A multiple sequence alignment of 32 human GAIN domains was used to identify the first residue of each CTF and CTFΔTA protein. The P1' residue of each TA is the first residue of each CTF protein, except for ADGRC1 and ADGRC3, which start at the residue preceding P1' because each possesses a nucleophilic threonine residue immediately after the basic residue of the autoproteolysis motif. The CTFΔTA proteins begin with the residue immediately following the P7' TA position, as indicated. For ADGRA1, which lacks a GAIN domain, the CTF construct consists of the full-length receptor and the CTFΔTA construct lacks the 17 N-terminal residues prior to TM1.



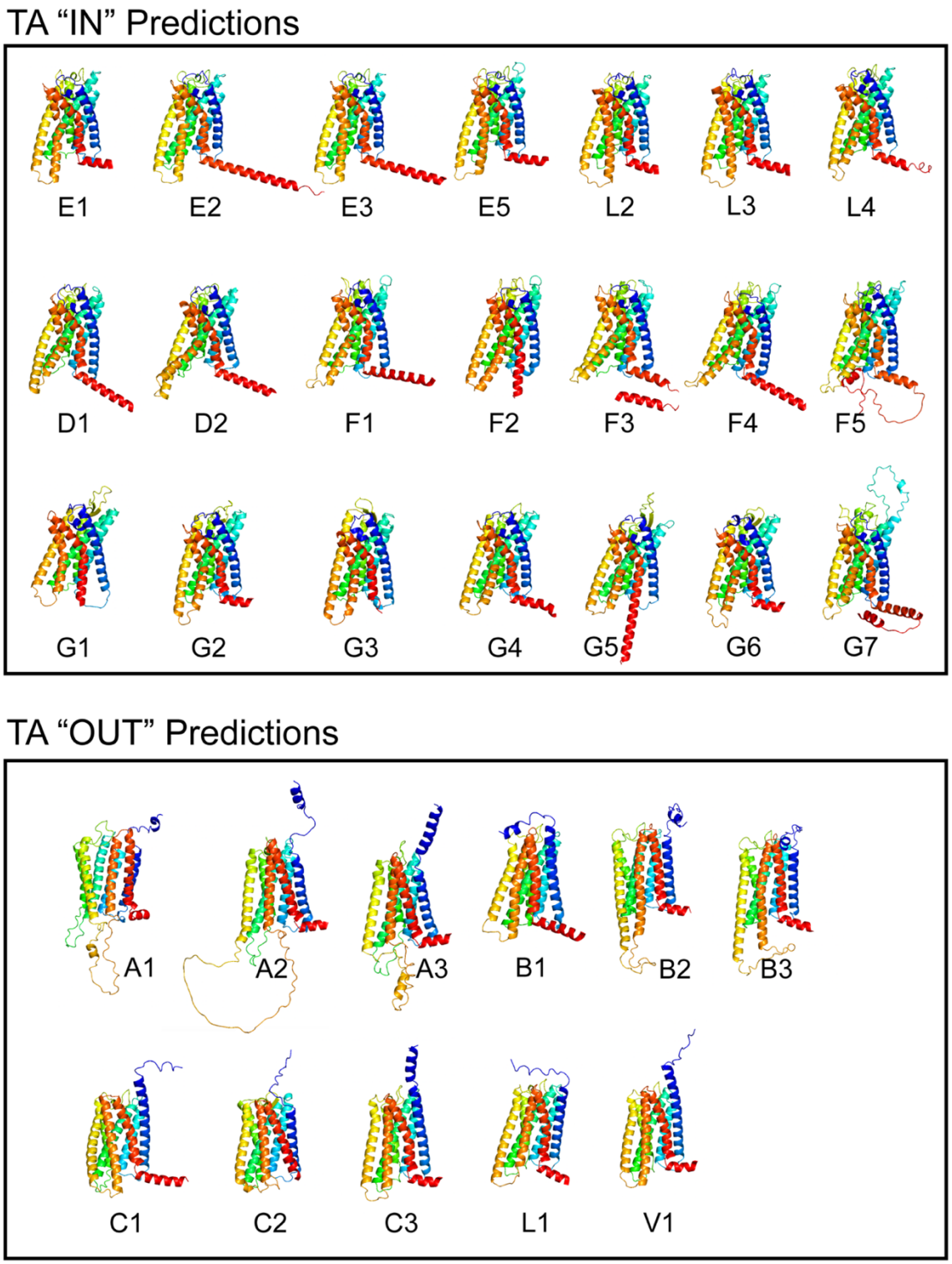
Receptor	Mass of MBP-CTF (kDa)
A1	104.4
A2	105.5
A3	108.1
B1	115.7
B2	117
B3	118.2
C1	104.8
C2	104.6
C3	129.2
D1	79.5
D2	79.2
E1	77.2
E2	71.5
E3	78.6
E4	75.2
E5	77.7

Receptor	Mass of MBP-CTF (kDa)
F1	81.8
F2	74.2
F3	79
F4	78.1
F5	82.6
G1	78
G2	89.6
G3	76.9
G4	84.1
G5	76.8
G6	86.7
G7	86.8
L1	113
L2	114.6
L3	111.4
L4	75.5
V1	89.5

Fig. S5 – Expression of human aGPCR proteins evaluated in this study. Western blot detection of N-terminal HA and C-terminal FLAG epitope tags to evaluate total protein expression for MBP-CTF and MBP-CTFΔTA variants of all 33 human aGPCRs (Note: ADGRE4 is a predicted pseudogene and no mRNA transcripts encoding the 7TM domain have been detected in humans). Blotting for GAPDH was used to control for protein loading.


Gs
Gq
Gi1
G13
EV A1 A2 A3 B1 B2 B3 C1 C2 C3 D1 D2 E1 E2 E3 E5 F1 F2 F3 F4 F5 G1 G2 G3 G4 G5 G6 G7 L1 L2 L3 L4 V1
ΔBRET Ratio
(EV normalized)
MBP-CTF (Filled Bars)
MBP-CTFΔTA (Unfilled Bars)

762 **Fig. S6 – Empty vector-normalized resting BRET ratios for all proteins tested in four G**
763 **protein biosensor assays.** Relative BRET ratios given by the human MBP-CTF (filled bars) and
764 MBP-CTFΔTA (unfilled bars) proteins across the four G protein biosensors tested. For each
765 protein, ΔBRET values are normalized to the resting BRET ratio established by cells transfected
766 with empty vector. Data are presented as mean ± SEM of three independent biological replicates
767 (n=3). Individual biological replicates are represented by circle, square, and diamond-shaped
768 data points.



769 **Fig. S7 – ColabFold structural predictions for the human CTF proteins.** Structural
770 predictions for each CTF protein sequence from ColabFold. Models were manually sorted into
771 two clusters based on the position of the TA peptide (in or out of the orthosteric pocket of the
772 7TM domain) and the presence of the kink in TM6, which is colored orange.

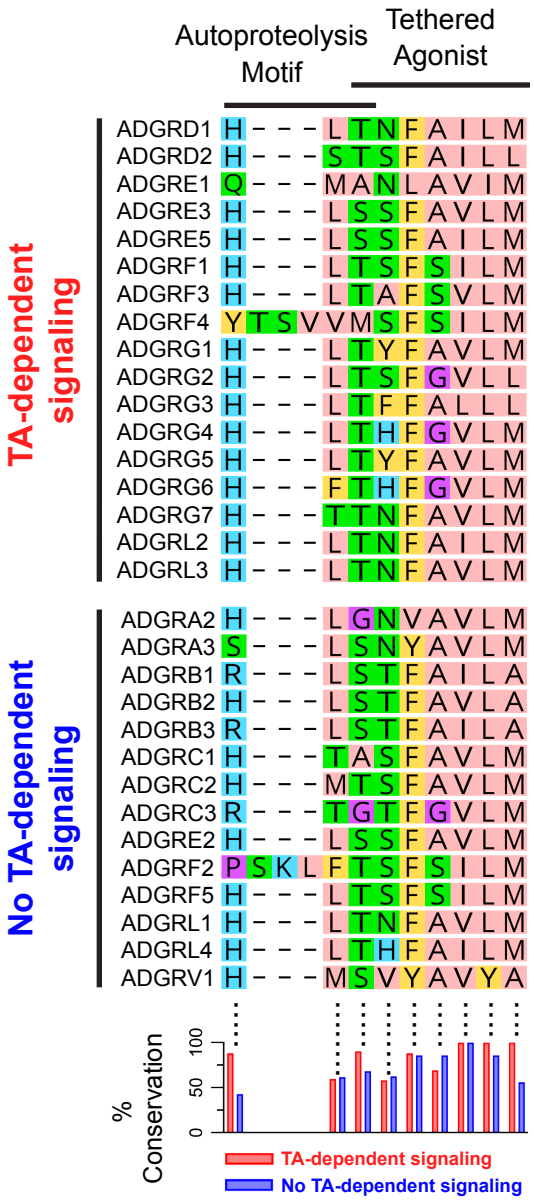


Fig. S8 – Conservation of the TA sequence and autoproteolysis motif among aGPCRs. Multiple sequence alignment of aGPCRs around the autoproteolysis motif and the TA, with aGPCRs separated based on TA-signaling competence in our MBP fusion system. The amount of conservation as a function of position is plotted below the sequence alignment.

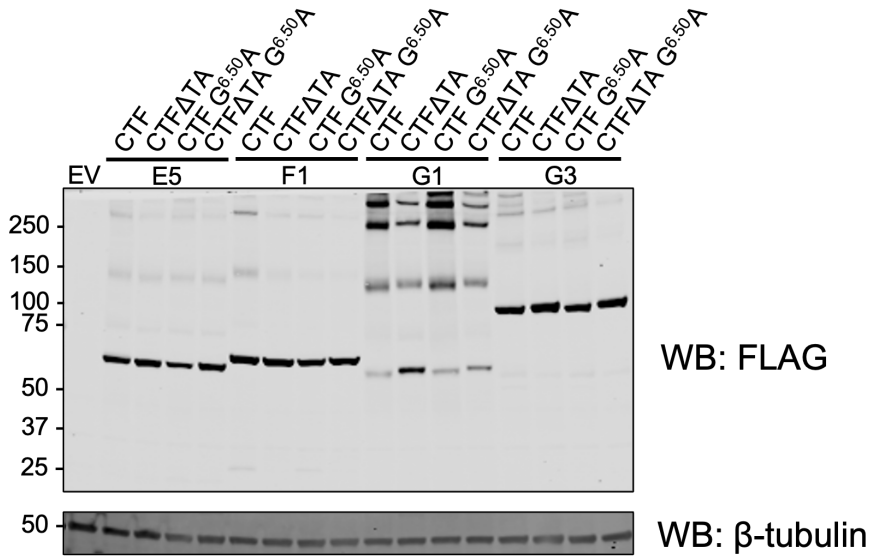


Fig. S9 – Expression of $G^{6.50}$A mutants. Western blot detection of C-terminal FLAG epitope tags for constructs evaluated in Fig. 6B. Blotting for β-tubulin was used to control for protein loading.

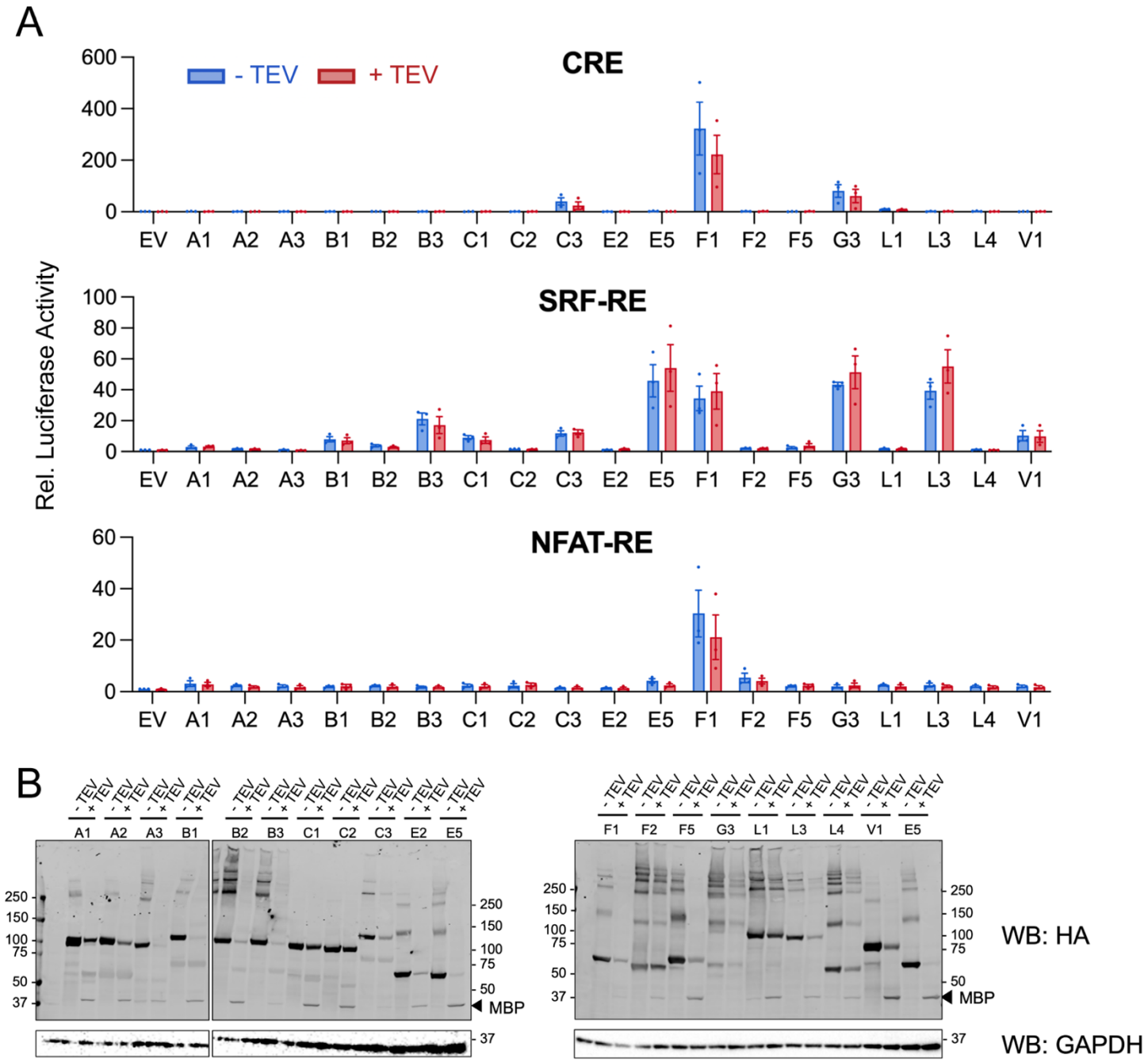


Fig. S10 – Effects of TEV-mediated MBP removal on signaling activity of CTF proteins. (A) CRE, SRF-RE, and NFAT-RE reporter gene assays for cells co-expressing various MBP-CTF proteins with (red) and without (blue) a TEV protease variant engineered for activity in the secretory pathway (secTEV). Data are normalized to the mean value for empty vector without secTEV expression and are presented as mean ± SEM of three independent biological replicates (n=3). (B) Western Blot detection of N-terminal HA epitope tags on MBP-CTF proteins co-expressed with and without secTEV protease. Blotting for GAPDH was used to control for protein loading.