Tethered agonist activated ADGRF1 structure reveals molecular preference for $G\alpha_q$ signalling

Daniel T. D. Jones[1], Andrew N. Dates[1], Shaun D. Rawson[1], Maggie M. Burruss[1], Colin H. Lipper[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

*Lead contact. Correspondence: stephen_blacklow@hms.harvard.edu and daniel_jones@hms.harvard.edu.

Abstract

Adhesion G-Protein Coupled Receptors (aGPCRs) have evolved an activation mechanism to translate extracellular force into liberation of a tethered agonist (TA) to modulate cell signalling. We report here that ADGRF1 is the first class B GPCR shown to signal through all major G-protein classes and identify the structural basis for its $G\alpha_q$ preference by cryo-EM. Our structure shows that $G\alpha_q$ over $G\alpha_s$ preference in ADGRF1 derives from tighter packing at the conserved F569 of the TA, altering contacts between TM helix I and VII, with a concurrent rearrangement of TM helices VII and VIII at the site of $G\alpha$ recruitment. $G\alpha_s$ signalling is also more sensitive to mutation of TA or binding site residues than $G\alpha_q$. Our work advances the understanding of aGPCR TA activation in molecular detail, identifying structural features that potentially explain preferential signal modulation.

Main

Adhesion G protein coupled receptors (aGPCRs) are the second largest class of GPCRs, constituting 33 members across 9 subfamilies [1]. aGPCRs control a multitude of cellular processes involved in organ development and tissue homeostasis, responding to external stimuli that sense extracellular physical forces, such as binding to protein ligands presented on cells, components of extracellular matrix and shear flow[2]. Mutations in aGPCRs are genetically responsible for human developmental disorders, such as ADGRV1 in Usher syndrome[3], and EMR2 in vibratory urticaria[4].

aGPCRs have a unique modular architecture, with an N-terminal ectodomain containing extracellular adhesive modules, followed by a G-protein coupled receptor autoinducing (GAIN) domain linked to a seven-transmembrane (7TM) bundle that relays extracellular events to the cell. A distinguishing hallmark of aGPCRs is that the GAIN domain undergoes autoproteolysis at a short hairpin turn between the final two beta-strands, separating the aGPCR into non-covalently attached N-terminal and C-terminal fragments, named NTF and CTF respectively[5,6]. Currently, models for aGPCR activation suggest that mechanical force applied to the adhesive modules in aGPCRs induces dissociation of the NTF from the CTF[7,8], enabling the newly liberated N-terminal end of the CTF to act as an intramolecular agonist to the 7TM domain. Intramolecular ligation of this tethered agonist (TA) to the 7TM domain relays the initial extracellular cue to the cell through G-protein coupling or beta-arrestin activity.

The first structure of an aGPCR 7TM domain was solved for ADGRG3 (GPR97) bound to glucocorticoids in the orthosteric pocket, revealing key residues involved in ligand recognition, conformational switches important for receptor activation, and residues for coupling to G-protein[9]. Cryo-EM structures were then reported for seven different aGPCR family members activated by their TAs, including representative GPCR-G-protein complexes for all major families of Gα, except for Gα_q [10–13]. These studies showed that the TA acts an intramolecular ligand for the 7TM domain and identified a canonical binding pose for a conserved hydrophobic interaction motif (HIM) of the TA in the orthosteric site.

ADGRF1 (GPR110) is an aGPCR belonging to aGPCR subfamily VI [1,14]. ADGRF1 was identified as an oncogene overexpressed in lung and prostate cancer[15], and has been implicated in synaptamide dependent neural outgrowth and repair[16,17], though this activity remains

50 controversial[18]. ADGRF1 was used as a prototypical aGPCR to determine that membranes
51 expressing ADGRF1 treated with urea at high concentration induced shedding of the NTF,
52 liberating the TA to activate the 7TM and recruit $G\alpha_q$ preferentially over other $G\alpha$ proteins[8].
53 Further work has also showed that mouse ADGRF1 can be activated by synthetic TA peptides when
54 added in trans, signalling through both $G\alpha_s$ and $G\alpha_q$[19].

55 Recent structures of the TA-engaged state of the ADGRF1 CTF coupled to either mini$G\alpha_s$
56 and mini$G\alpha_i$ also identified a lysophosphatidylcholine (LPC) lipid bound at the intracellular
57 convergence of TM helices II, III and IV[13]. Interestingly, fragments of ADGRF1 expressed with the
58 GAIN domain included also yielded TA-engaged structures of ADGRF1 coupled to miniG but lacked
59 density for the GAIN domain, suggesting that the CTF may have dissociated from the NTF portion
60 of the protein during purification.

61 To understand the molecular events driving TA activation of ADGRF1 and how its structure
62 determines $G\alpha$ class preference, we determined a cryo-EM structure of TA-activated ADGRF1
63 coupled to mini$G\alpha_{s/q}$-$\beta_1\gamma_2$, stabilised by Nb35. Combining structural insights with extensive
64 functional studies, we identify key residues of the orthosteric binding site that are essential for
65 $G\alpha_s$ signalling but dispensable for $G\alpha_q$ signalling. Overall, our work complements other recent
66 aGPCR structures and extends them by visualizing ADGRF1 in complex with its most relevant G-
67 protein partner, elucidating a functional and structural landscape by which ADGRF1 is activated
68 and uncovering how its structure determines $G\alpha$ class preference. Furthermore, the mechanistic
69 insights determined here regarding TA engagement to strong and weakly associated G-proteins
70 offers valuable insight into how TA mediated activation of aGPCRs may be functionally modulated.

71 Results

72 ADGRF1 signals through all G-protein classes

73 We recently developed a platform for analysis of the signalling activity of aGPCRs, in which the CTF
74 of interest is expressed as a protein fusion with an N-terminal IL2 signal sequence followed by
75 maltose binding protein (MBP) and a tobacco etch virus cleavage site (denoted MBP-CTF) [20]
76 (Extended Data Fig. 1). Using this approach to measure signalling activity in reporter assays, we
77 observed transcriptional responses with both CRE and NFAT reporters, indicative respectively of
78 $G\alpha_s$ and $G\alpha_q$ signalling, consistent with previous studies[13,21]. We also observed transcriptional
79 responses with the SRE and SRF-RE reporters, which suggests that ADGRF1 can also signal through
80 $G\alpha_{i/o}$ and $G\alpha_{12/13}$ (Fig. 1a).

81 To further evaluate the potential for signalling through $G\alpha_{12/13}$ (a structure of $G\alpha_i$ has been
82 reported[13]) we used HEK293T cells with deletion of six Gα subunits ($G\alpha_s$ short, $G\alpha_s$ long, $G\alpha_q$,
83 $G\alpha_{11}$, $G\alpha_{16}$, $G\alpha_{12}$ and $G\alpha_{13}$, gift from Associate Prof. A. Inoue, Tohoku university, Japan)[22], and
84 deletion complementation by transient transfection to determine which Gα subunits support
85 ADGRF1 signalling in the SRF-RE reporter assay. ADGRF1 MBP-CTF signals strongly through $G\alpha_q$
86 and $G\alpha_{11}$, and also exhibits a smaller significant signal for $G\alpha_{12}$ and $G\alpha_{13}$ (Fig. 1b). These data show
87 that ADGRF1 has the functional capacity to signal through all Gα classes, making ADGRF1 the first
88 of the class B GPCRs shown to have this capability[23].

89

90 Protein engineering of ADGRF1 and Cryo-EM structure determination

91 To obtain an active state ADGRF1 G-protein protein complex amenable for structural and
92 functional investigation, we probed complex assembly with a suite of miniGα proteins[24,25], using
93 a NanoBiT recruitment assay[26, 27]. We found that LgBiT-miniGα_s, miniG$\alpha_{s/q}$ and miniGα_o were all
94 recruited to C-terminal SmBiT-tagged ADGRF1-CTF as judged by increased luminescence upon
95 SmBiT-LgBiT complementation, whereas LgBiT-miniGα_{12}, and matched CΔ5 controls were not
96 (Extended Data Fig. 2a). To assess tractibility for structure determination, we co-expressed these
97 constructs in Expi293T suspension cells, and also observed a similar pattern of Gα recruitment to
98 ADGRF1 in these cells (Extended Data Fig. 2b, left). To further test the stability of the complex, we
99 solubilised cells with mild detergent supplemented with cholesterol in the presence of apyrase.

100 The only complex that retained luminescence under these lysis conditions was the combination of
101 SmBiT-tagged ADGRF1-CTF with LgBiT-miniG$\alpha_{s/q}$ (Extended Data Fig. 2b, right). In addition, only
102 miniG$\alpha_{s/q}$ specifically purified with ADGRF1-CTF using anti-FLAG M2 magnetic beads (Extended
103 Data Fig. 2c), indicating that ADGRF1-Gα complex stability is greater for miniG$\alpha_{s/q}$ compared to
104 miniGα_s and miniGα_o. Furthermore, miniG$\alpha_{s/q}$ co-expression with ADGRF1-CTF appears to
105 increase the amount of receptor purified, suggesting that the miniG$\alpha_{s/q}$ may improve receptor
106 expression by facilitating the folding of the receptor, increasing its stability, or by blocking toxic
107 signalling through endogenous signalling pathways.
108

109 **Structure of the ADGRF1-miniG$\alpha_{s/q}$ heterotrimeric G protein complex with the Nb35 nanobody**

110 To elucidate the basis for selective recruitment of G$\alpha_{s/q}$, we complexed ADGRF1 CTF with
111 a miniG$\alpha_{s/q}$ heterotrimeric G protein complex and the Nb35 nanobody, purified the complex by
112 size-exclusion chromatography (Fig. 2a), and determined its structure by cryo-EM to 3.4 Å
113 resolution (Extended Data Fig. 3-4). The cryo-EM map was of sufficient resolution to permit the
114 building of an atomic model for the receptor, G-protein complex, and nanobody (Fig. 2b, Extended
115 Data Fig. 5). The cryo-EM map in the orthosteric pocket on the extracellular face of the receptor
116 permitted unambiguous modelling of the TA (Extended Data Fig. 5), which adopts a short alpha-
117 helical motif with a loop capping the TA *en route* to TM1, akin to other resolved TA structures of
118 aGPCRs[10–13]. The polypeptide backbone of the 7TM domain of ADGRF1 bound to miniG$\alpha_{s/q}$ has an
119 RMSD of 0.6 Å when aligned to the 7TM domain of ADGRF1 bound to either miniGα_s and miniGα_i,
120 highlighting that the overall organization of these structures is highly similar despite their different
121 G-protein partners. An LPC lipid is present at the intracellular interface of TM helices II, III and IV,
122 as seen in ADGRF1 structures bound to miniGα_s or miniGα_i[13]. The acyl chain is similarly positioned
123 in our structure, whereas the headgroup has the phosphate embedded deeper in the detergent
124 micelle (Fig. 2b, Extended Data Fig. 5).
125

126 **Structural differences between the ADGRF1 miniG$\alpha_{s/q}$ and ADGRF1 miniGα_s complexes**

127 MiniG$\alpha_{s/q}$ was engineered as a chimera with six key Gα_q residues on the C-terminal α5 helix
128 that distinguish Gα_q coupling over Gα_s, even though the bulk of the chimera is derived from

129 miniGα_s[24]. We also observe a preference for miniG$\alpha_{s/q}$ recruitment to the CTF of ADGRF1,
130 consistent with the previously observed preference for Gα_q in GTPγS activity assays[8], arguing that
131 the preference for Gα_q is driven largely by the differences in its C-terminal α5 helix. At the ADGRF1
132 G-protein interface we observe a canonical pose of the C-terminal α5 helix at the open cavity at
133 the intracellular opening of TA activated ADGRF1 (Fig. 2b and Fig. 3a). The N387 amide side chain
134 nitrogen of miniG$\alpha_{s/q}$ approaches within hydrogen bonding distance of the backbone carbonyl of
135 ADGRF1 R685$^{3.57}$, and the side chain nitrogen of N392 of miniG$\alpha_{s/q}$ approaches within hydrogen
136 bonding distance of the receptor side chain carboxylate of D842$^{8.47}$. The analogous residues of
137 miniGα_s are H387 and E392, which do not form the same hydrogen bonding arrangement in the
138 Gα_s complex. In the Gα_s complex, H387 does not make any polar interactions, and instead E392
139 approaches within hydrogen bonding distance of the backbone amine of S843$^{8.48}$ (Extended Data
140 Fig. 6a). Given the close proximity of E392 to D842$^{8.47}$, It is likely that charge repulsion destabilises
141 the miniGα_s-bound state by comparison to the Gα_q-bound state. These interactions are associated
142 with a shift in the positioning of the segment linking TM helix VII to helix VIII relative to the Gα
143 subunit, with the Cα of D392 in the α5 helix of miniGα_s situated 1.9 Å closer to the Cα of ADGRF1
144 S843$^{8.48}$ (Fig. 3c). The arrangement seen in the miniGα_s complex is also present in the miniGα_i
145 structure. Superposition of helices I and VII of miniG$\alpha_{s/q}$ on miniGα_s shows a 3.2 Å displacement
146 of the L384 Cα miniG$\alpha_{s/q}$ relative to that of N384 of miniGα_s (Extended Data Fig. 6b).
147 At the orthosteric site, we also observe a 1.2 Å displacement of L593$^{1.47}$ towards F569$^{P3'}$ in
148 the ADGRF1 complex with miniG$\alpha_{s/q}$, compared with the miniGα_s and miniGα_i complexes (Fig. 3d).
149 This inward movement leads to a rearrangement of the TM I alpha helix, allowing the conserved
150 S596$^{1.50}$ hydroxyl to approach within hydrogen bond distance of the backbone carbonyl of A828$^{7.48}$
151 at the kink in the TM VII helix. This interaction of S569$^{1.50}$ with TM helix VII is not observed in either
152 the miniGα_s or miniGα_i structures.
153
154 **Distinguishing preferential Gα_q over Gα_s tethered agonism at the orthosteric site**
155 Because ADGRF1 shows preference for Gα_q over other Gα partners[8], we investigated
156 whether Gα_q and Gα_s signalling exhibited differential sensitivity to alanine mutants of the receptor
157 stalk residues (Fig. 4a), using NFAT and CRE reporter gene assays for Gα_q and Gα_s coupling,

158 respectively (Fig. 4b). Positions 574-581, corresponding to P8'-P15', were largely unaffected by
159 mutation to alanine in NFAT and CRE assays, suggesting this region is not directly important for TA
160 dependent agonism, and therefore functioning mainly as a linker connecting the TA to TM1 in
161 ADGRF1.

162 Alanine substitutions at $F569^{P3'}$, $L572^{P6'}$ and $M573^{P7'}$ of the HIM affected TA-dependent
163 signalling. The $L572^{P6'}A$ and $M573^{P7'}A$ mutations both reduced the response of the CRE reporter,
164 whereas these substitutions either did not affect ($L572^{P6'}A$) or modestly increased ($M573^{P7'}A$) the
165 response of the NFAT reporter, arguing that $G\alpha_s$, but not $G\alpha_q$, selectively requires bulky
166 hydrophobic residues at positions P6' and P7' for TA-dependent signalling in ADGRF1.

167 Strikingly, we found that the lack of signalling activity for $F569^{P3'}A$ in both NFAT and CRE is
168 accompanied by greatly decreased receptor abundance, as judged by the lack of a band on
169 Western blot when probing for the N-terminal 3X-FLAG tag with an anti-Flag antibody (Fig. 4c). An
170 antibody raised to a C-terminal portion (residues 831-880) of ADGRF1 showed a similar reduction
171 in Western blot signal for the $F569^{P3'}A$ mutant (Extended Data Fig. 7), confirming that the loss of
172 anti-FLAG reactivity was not due to proteolysis of the N-terminal tag. We propose that failure of
173 the mutated TA to engage the 7TM leads to the exposure of the TA hydrophobic sequence,
174 creating a neoepitope that stimulates heat-shock protein binding in the secretory pathway and
175 subsequent degradation (*i.e.*, the inactivity of the TA is the cause of the reduced protein amount,
176 and not a consequence of it).

177 We explored the tolerance of $F569^{P3'}$ more extensively by testing a larger range of amino
178 acid substitutions, including a conservative leucine substitution, a methionine substitution, and a
179 suite of charged amino acids (Fig. 4d, e). Among these substitutions, only the $F569^{P3'}L$ mutation
180 produced a detectable signal (Fig. 4d), and only this substitution resulted in similar amounts of
181 expressed protein as the wild-type CTF as judged by Western blot (Fig. 4e).

182 Because of the variation in protein abundance for most of the $F569^{P3'}$ mutants, we also
183 assessed the role of the P3', P6', and P7' hydrophobic buried residues in activation using *trans*
184 addition of TA peptide sequences. For these assays, we used a form of the 7TM domain lacking
185 the TA, which we refer to as stalk-less, or Δstalk. We incubated cells expressing the Δstalk form of
186 ADGRF1 with a 12mer TA mimicking peptide (P12) at 100 μM or 10 μM, and with analogous

187 peptides containing alanine substitutions at P3', P6' or P7', and measured signalling activity using
188 the NFAT reporter. We found that the Δstalk construct signaled robustly through NFAT with the
189 P12 peptide at 100 μM, but not 10 μM (Fig. 4f, g). Together, our data show that a large bulky
190 hydrophobic residue is required at P3' for ADGRF1 tethered agonism, that $G\alpha_q$ signalling in TA
191 mode is more tolerant of substitutions at the positions at P6' or P7' in the HIM motif than $G\alpha_s$
192 signalling, and that signal activation by peptides added in *trans* is intolerant of alanine substitutions
193 at any of the HIM positions (P3', P6', or P7') even for $G\alpha_q$ signalling.

194 To further investigate tethered agonism of ADGRF1 through $G\alpha_s$ and $G\alpha_q$, we performed
195 extensive alanine mutagenesis of the TA binding site in the 7TM domain and measured the
196 signalling response of the NFAT and CRE reporters. The signalling activity of the mutated proteins
197 clustered into three distinct categories: mutations that did not affect the CRE or NFAT response,
198 those that were detrimental to both CRE and NFAT, and those that were detrimental to CRE only
199 (Fig. 5a, Extended Data Fig. 8).

200 Residues where alanine substitutions reduced signalling in both the NFAT and CRE
201 reporters include conserved residues $W734^{45.51}$ and W737 of ECL2, and several residues that
202 surround $F569^{P3'}$ in the 7TM core, including conserved 'switch' residue $W804^{6.53}$, neighboring
203 residues $F823^{7.41}$ and $N827^{7.45}$, and $T589A^{1.43}$, $L593A^{1.47}$, $A828L^{7.46}$ and $D638A^{2.61}$ (Fig. 5b-d). These
204 sets of mutations outline the importance of ECL2, the binding site arrangement around the
205 'switch' residue, and the conserved $F569^{P3'}$ position for general ADGRF1 tethered agonism. We
206 also found that positions $V812^{6.61}$ and $S814^{6.63}$ lowered signalling in both NFAT and CRE assays,
207 but this was likely due to a profound decrease in protein expression (Extended Data Fig. 8)

208 Strikingly, mutations that decreased CRE signalling but not NFAT signalling were in two
209 separate clusters that interact with positions other than $F569^{P3'}$ of the TA sequence. The first
210 cluster contains residues within van der Waals contact of $L572^{P6'}$ and $M573^{P7'}$ at positions $Y668^{3.40}$,
211 $F672^{3.44}$ and $F747^{5.39}$ on TM helices III and V, with the $F672^{3.44}A$ substitution actually enhancing the
212 NFAT response (Fig 5b-d). The second cluster consists of $T810^{7.28}$ and $H820^{7.38}$; $T810^{7.28}$ interacts
213 with $M573^{P7'}$ and $S574^{P8'}$, and $H820^{7.38}$ approaches within H-bonding distance of $S570^{P4}$. These
214 results show that CRE signalling is more sensitive that NFAT signalling upon disruption of the
215 pocket proximal to P6' and P7', as well as upon disruption of interactions with $S570^{P4'}$ and $S574^{P8'}$.

216 Discussion

217 For a typical G-protein in which the agonist is a soluble ligand, G-protein coupling in a
218 biological system is guided by intrinsic G-protein binding preferences, agonist selectivity for G-
219 protein, and the local reservoir of G-proteins in that cellular environment. aGPCRs are unique in
220 that the 7TM bundle and its TA are encoded in the same polypeptide chain. As a result, aGPCR G-
221 protein signalling is simplified to a combination of two factors – the G-protein preference of the
222 TA-engaged CTF and the G-protein context of the cellular environment.

223 Here, we combine cryo-EM structural studies with mutational analysis and cell-based
224 signalling assays to elucidate structure-function relationships in ADGRF1 and deduce its G-protein
225 coupling profile preferences. Our work establishes that ADGRF1 prefers coupling to $G\alpha_q$, as
226 reported previously[8], provides a structure-based rationale for this coupling preference, and, most
227 strikingly, identifies a cluster of residues in the TA and in the binding pocket that selectively
228 suppress $G\alpha_s$ coupling upon mutation while retaining $G\alpha_q$ coupling (Fig. 6).

229 While the structure of the ADGRF1-mini$G\alpha_{s/q}$ complex shows that the overall architecture
230 of the complex and the pose of the TA in the orthosteric site are similar to recently observed
231 structures of ADGRF1 when complexed to mini$G\alpha_s$ and mini$G\alpha_i$[13], there are subtle differences in
232 the complexes that explain why $G\alpha_q$ coupling is more robust than $G\alpha_s$ coupling. We also establish
233 that ADGRF1 can signal through $G\alpha_{12}$ or $G\alpha_{13}$, which to our knowledge makes ADGRF1 the first
234 class B GPCR to show the functional capacity to signal through all major G-protein classes[28].

235 Among the structural differences, tighter packing of $L593^{1.47}$ with the highly conserved
236 phenylalanine (P3') of the HIM appears to be critical for $G\alpha_q$ preference, because the positional
237 shift of $L593^{1.47}$ propagates its effects down helix I, introducing a hydrogen bond between $S596^{1.50}$
238 and $A828^{7.48}$ to stabilize the kink in TM helix VII. Our functional data suggests that these
239 interactions contribute towards an increased stability of the ADGRF1-$G\alpha_q$ complex relative to the
240 $G\alpha_s$ and $G\alpha_i$ complexes. Both the TA and the binding pocket of the 7TM domain show increased
241 mutational lability for $G\alpha_s$ signalling at several residues, whereas $G\alpha_q$ signalling remains
242 unaffected by those same mutations, showing that $G\alpha_q$ signalling is more tolerant to mutations
243 than $G\alpha_s$. Our data also establish the critical importance of the $F569^{P3'}$ residue for ADGRF1 G-
244 protein signalling, because a bulky hydrophobic residue at the P3' position is essential for

245 expression and TA-dependent activation of the ADGRF1-CTF and mutations of residues that
246 surround the F569$^{P3'}$ pocket are detrimental in both NFAT and CRE assays.
247 Unlike TA signaling of ADGRF1, which is tolerant of substitutions at positions other than
248 F569$^{P3'}$, we found that *trans* addition of 12mer peptides to a stalkless form of ADGRF1 mimics the
249 TA only when the wild-type sequence is used at the relatively high concentration of 100 µM.
250 Mutations at the P6' and P7' positions of the peptides, which are tolerated in TA signaling assays,
251 failed to activate the stalkless ADGRF1 protein. This observation indicates that the development
252 of peptidomimetic agonists for ADGRF1 will require extensive optimization if high potency is to be
253 achieved in a soluble ligand.
254 Finally, our data elucidating a molecular basis for the preference of ADGRF1 for coupling
255 to $G\alpha_q$ over $G\alpha_s$ identifies key sites in the 7TM that influence G protein coupling preferences.
256 These findings should also inform the study of G protein coupling preferences and signaling biology
257 for other aGPCRs.
258
259 Methods
260 Cloning and construct design
261 MBP-CTF and signalling assay plasmids
262 DNA sequences encoding full length (20-910) and CTF (567-910) forms of ADGRF1 (Uniprot
263 Q5T601) were subcloned into the pFuse-hIgG1-Fc2 vector behind an N-terminal IL2-signal
264 sequence followed by a HA epitope tag, maltose binding protein and TEV protease site[29]. The CTF
265 sequence was also subcloned to create the M-CTF protein with the N-terminal tags and mannose-
266 binding protein omitted. Firefly reporter plasmids (CRE, NFAT, SRE and SRF-RE) were assembled
267 as described[29]. The pRL-TK plasmid (Promega) was used to express *Renilla* luciferase.
268
269 Receptor and G-protein plasmids
270 For recruitment assays, sequences encoding full-length (20-910) and CTF (567-910) ADGRF1 were
271 subcloned into pcDNA3.1(+) with an N-terminal haemagglutinin signal sequence followed by a
272 3XFLAG purification tag and a TEV protease cleavage site. For the CTF construct the TEV protease
273 site P1' residue encoded a threonine so that the correct residue would occupy T567 in the CTF

274 (e.g., ENLYFQ/TSFSILM…). The C-terminus has a "GGGGSGGGGSSG" linker followed by a SmBiT tag
275 "VTGYRLFEEIL"[30]. For proteins probed by Western blot the C-terminal SmBiT tag was replaced with
276 a HA epitope tag. Mutagenesis was performed with a Q5 site directed mutagenesis kit (New
277 England Biolabs) or using a modified QuikChange™ method[31].
278 MiniG protein sequences (miniGα_s, miniGα_o, miniG$\alpha_{s/q}$ and miniGα_{12}) with capacity to
279 complex with G$\beta_1\gamma_2$ (e.g. 399 variants[24]) were subcloned into pcDNA3.1(+) with an N-terminal
280 LgBiT fusion followed by a "GGGGSGGGGSSGEF" linker that includes a translated EcoRI site. For
281 structure determination and mutagenesis studies of ADGRF1, the constructs described above for
282 recruitment assays were used.
283
284 **Protein expression, purification and complexation**
285 ADGRF1 and miniG$\alpha_{s/q}$ were co-expressed in Expi293F cells. Cells were seeded at a density of 3.3
286 x 10^6 cells/mL in 900 mL of Expi293 expression media. To transfect cells, equal amounts of receptor
287 (250 ug) and miniG (250 ug) DNA were resuspended in 50 mL opti-MEM, FectoPro (Polyplus)
288 transfection reagent (0.5 mL) was resuspended in 50 mL opti-MEM, the DNA and FectoPro
289 solutions were mixed to give a final 1:1 DNA/FectoPro ratio, and the mixture incubated at room
290 temperature for 10 minutes before adding to cells. Approximately 24 hours later, filter-sterilised
291 Valproic acid sodium salt (Sigma-Aldrich) was added to 5 mM, along with 10 mL of 45 % D-(+)-
292 Glucose solution (Sigma-Aldrich). Cells were cultured for a further 24 hours before they were
293 harvested by centrifugation at 4000 *g* for 15 minutes. The pellet was flash frozen and then stored
294 at -80 ^{o}C until purification.
295 Cells were thawed and resuspended in 250 mL of ice-cold hypotonic lysis buffer (10 mM
296 HEPES, pH 7.5, containing Roche protease inhibitor tablets and 0.1 mU/mL apyrase to prevent
297 separation of the GPCR-miniG complex). The pellet was stirred on ice for 30 minutes and lysed by
298 dounce homogenization. The broken cells were pelleted by centrifugation at 50,000 *g* and
299 resuspended in 225 mL solubilisation buffer (10 mM HEPES, pH 7.5 containing 1 mM $MgCl_2$, 2 mM
300 $CaCl_2$, and 100 mM NaCl) with 0.1 mU/mL apyrase and 1:100,000 Benzonase (v/v). The
301 resuspended material was dounce homogenized, and DDM-CHS (10:1 pre-mix, Anatrace) was then
302 added to a final concentration of 1 % (w/v). This solution was stirred for 75 min minutes at 4 ^{o}C,

303 clarified by centrifugation at 50,000 *g* for 1h, and passed through a glass microfibre filter. The
304 filtrate was then loaded onto M2 anti-FLAG antibody affinity resin by gravity flow. The resin was
305 washed with 30 column volumes of wash buffer (10 mM HEPES, pH 7.5 containing 1 mM $MgCl_2$, 2
306 mM $CaCl_2$, 100 mM NaCl, and 0.1 % (w/v) LMNG-CHS), and the protein was eluted by the
307 application of wash buffer containing 100 µg/mL of 3XFLAG peptide. The eluted protein was
308 concentrated using a 100 kDa MWCO centrifugal concentrator (Amicon).

309 Nb35 and human $G\beta_1\gamma_2$ were purified according to previously published protocols[32].
310 ADGRF1-mini$G\alpha_{s/q}$, Nb35 and $G\beta_1\gamma_2$ were mixed in a 1:1.15:1.15 ratio and incubated overnight at
311 4 ^{o}C with rotation. The protein mixture was centrifuged at 21,000 *g* for 20 minutes at 4 ^{o}C, and
312 then spun through a Durapore® PVDF 0.1 µm column (Millipore). The protein mixture was injected
313 onto a Superdex S200 10/300 GL column equilibrated in SEC buffer (10 mM HEPES pH 7.5
314 containing 1 mM $MgCl_2$, 2 mM $CaCl_2$, 100 mM NaCl, and 0.005 % (w/v) LMNG-CHS). The peak
315 fraction containing the complex was collected and concentrated using a 100 kDa MWCO
316 centrifugal concentrator to approximately 8 mg/mL.
317

318 Cryo-EM sample preparation and image acquisition

319 The complex was vitrified on QUANTIFOIL® holey carbon grids (400-mesh, copper, R1.2/1.3,
320 Electron Microscopy Sciences) using a FEI Vitrobot Mark IV (FEI, Hillsboro). Grids were glow-
321 discharged, and 3 µL of sample loaded to the grid in a chamber at 22 ^{o}C and 100 % humidity.
322 Samples were applied at a force of 15 and blotted for 5-7 seconds before plunge-freezing in liquid
323 ethane.

324 Data were collected using a FEI Titan Krios at 300 kV with a Gatan Quantum Image Filter
325 with K3 Summit direct electron detection camera in counting mode with a total exposure dose of
326 approximately 54 $e^{-}/Å^{-2}$. 50 frames per movie were collected at magnification of 105,000x,
327 corresponding to 0.825 Å per pixel. Micrographs were collected at defocus values ranging from -
328 0.5 to -1.5 um. The processing scheme is summarised in Extended Data Figure 4, in brief, motion
329 corrected and dose-weighted using the RELION motion correction implementation[33] and contrast
330 transfer function parameters estimated by CTFFIND4[34]. Low quality micrographs were removed

331 using micassess[35] prior to particle picking in topaz [36]. ~4.9M particles were initially identified,
332 leading to ~2.9M particles after 2D classification in RELION [37]. Two independent 3D classifications
333 in RELION were then performed with different initial references, with the best classes from each
334 subsequently combined and duplicates removed to leave ~1.2M particles corresponding to
335 ADGRF1. Following further 3D classification in RELION ~660k particles underwent particle polishing
336 in RELION [33] and further 2D classification followed by masked 3D classification with no alignment
337 using a T value of 20 in RELION. 73,903 particles with strong transmembrane density were
338 identified which resulted in a 3.6Å reconstruction from CryoSPARC non-uniform refinement with
339 contrast transfer function refinement [38]. Subsequent local non-uniform refinement with a mask
340 excluding the micelle improved the reconstruction to 3.44Å.
341

342 **Atomic modelling and model refinement**

343 Model building was carried out in Coot using overlaid non-uniform refinement and local filtered
344 maps generated in CryoSPARC[39], as well as deepEMhancer[40] maps. The ADGRF1-miniGα_s
345 coordinate file (PDB: 7WU3) was used as an initial model. The miniGα_s sequence was changed in
346 Coot to generate miniG$\alpha_{s/q}$[24]. The coordinates were then manually rebuilt using Coot, and refined
347 using ISOLDE[41], Phenix Real-Space Refine[42] and Servalcat[43]. The final models were evaluated using
348 MolProbity. Statistics of the map reconstruction and model refinement are presented in
349 Supplementary Table 1. Structural biology applications used in this project (except CryoSPARC)
350 were compiled and configured by SBGrid[44].
351

352 **Cell Culture**

353 HEK293T and HEKΔ6 cells were cultured in DMEM supplemented with 10% FBS and 1% penicillin-
354 streptomycin at 37°C in a 5% CO_2 humidified incubator.
355

356 **Dual luciferase reporter assays**

357 HEK293T cells were seeded in DMEM in clear-bottom white 96-well plates pre-coated with poly-
358 D-lysine the day before transfection. At roughly 70 % cell confluency, the medium in each well was
359 replaced with fresh DMEM. Cells were transfected using Lipofectamine™ 2,000 (Invitrogen) at a

360 ratio of 3 µL per 1 µg DNA in Opti-MEM using the manufacturer's instructions. Each well was co-
361 transfected with 50 ng firefly reporter plasmid, 1 ng pRL-TK plasmid, and the indicated amounts
362 of receptor DNA. Cells were cultured in the transfection mixture for 6 hours before the media was
363 replaced with complete media. 24 hours after transfection, the media was replaced with fresh
364 DMEM, and the cells were incubated for another 8 hours prior to assay readout. For SRF-RE and
365 SRE read-outs the media was replaced with DMEM lacking FBS. Before measuring luminescence,
366 cells were washed once with PBS, 20 µL of 1X Passive Lysis Buffer was added to each well, and the
367 lysates were incubated with shaking for 15 minutes. Cell lysates were serially pipetted, and the
368 assay plate was centrifuged at 400 *g* for 1 minute. Dual-Glo luciferase measurements (Promega)
369 were read on a GloMax luminometer. The ratio of Firefly: Renilla signal was calculated for each
370 well and normalized to the mean of the EV-transfected controls to give Relative Luminescence
371 Units (RLU). Each assay was performed in technical triplicate wells and then repeated in biological
372 triplicate on different days.

373 For peptide experiments, transfection was performed as above. Peptides were dissolved
374 in DMSO to give 10 mM stocks and diluted to 10x concentration in Opti-MEM immediately before
375 use. The final DMSO concentration was 1 % (v/v) in all wells. Peptides were added *in trans* at the
376 initial media change 6 hours after transfection, and luminescence readings were acquired 24 hours
377 after transfection. Each assay was performed in technical triplicate wells and then repeated in
378 biological duplicate on different days.

379 For Gα rescue assays, HEKΔ6 cells were seeded as above the day before transfection.
380 Varying amounts of Gα subunit plasmid (0.3 ng of $G\alpha_{12}$ and $G\alpha_{13}$; 20 ng of $G\alpha_s$, $G\alpha_q$, $G\alpha_{11}$, and
381 $G\alpha_{16}$) were co-transfected with 30 ng firefly reporter plasmid and 0.6 ng pRL-TK plasmid (SRF-RE),
382 and total DNA load per well was balanced with pcDNA3.1. Media was changed 6 hours after
383 transfection mixes were added to the wells as in the standard assay format. At 24 hours after
384 transfection, the media was replaced with DMEM without FBS, and the cells were incubated for
385 another 8 hours prior to assay readout. Each assay was performed in technical triplicate wells and
386 then repeated in biological duplicate on different days.

387

388 G-protein NanoBiT recruitment assays

389 HEK293T cells were seeded in DMEM in clear-bottom white 96-well plates pre-coated with poly-
390 D-lysine the day before transfection. Cells were transfected using GeneJuice™ (Sigma-Aldrich) at
391 a ratio of 3 µL per 1 µg DNA in Opti-MEM using manufacturer's instructions. Equal amounts of
392 receptor and miniG DNA (50 ng) were used per well. Cells were incubated for 48 hours and washed
393 with PBS prior to assay readout. Cells were resuspended in opti-MEM and the Nano-Glo®
394 Luciferase Assay was read out on a GloMax luminometer. Each assay was performed in biological
395 triplicate.

396 Author Contributions:

397 D.T.D.J. and S.C.B. conceived of the research. D.T.D.J., M.M.B. and A.N.D. designed and cloned
398 constructs. M.M.B. performed mutagenesis of TA and binding site residues. M.M.B., A.N.D. and
399 D.T.D.J. performed signaling assays. D.T.D.J. performed NanoBiT recruitment assays. D.T.D.J.
400 expressed, purified and formed the ADGRF1-MiniG$\alpha_{s/q}$-G$\beta_1\gamma_2$-Nb35 complex. D.T.D.J. and C.H.L.
401 vitrified ADGRF1 complex on grids. S.D.R. and D.T.D.J. processed cryo-EM data, and S.D.R.
402 produced the final map. D.T.D.J. built the ADGRF1 complex model. D.T.D.J. and S.C.B. wrote the
403 manuscript.
404

405 Acknowledgements:

406 We thank members of the Blacklow and Kruse laboratories for helpful discussions. Cryo-EM data
407 were collected at the Harvard Cryo-EM Centre for Structural Biology at Harvard Medical School.
408 We thank Sarah Sterling and Megan Meyer for helpful comments and advice during grid screening,
409 and Megan Meyer for data collection.

410 Funding:

411 This work was supported by R35 CA220340 (SCB), and a grant from the Warren Alpert Foundation
412 (SCB). A.N.D. was also supported in part by a Fujifilm fellowship.

413

414 Competing Interests:

415 S.C.B. receives funding for an unrelated project from Novartis, is on the scientific advisory board
416 for Erasca, Inc., is an advisor to MPM Capital, and is a consultant for IFM, Scorpion Therapeutics
417 and Ayala Pharmaceuticals for unrelated projects.

418 References

419 1. Bjarnadóttir, T. K. *et al.* The human and mouse repertoire of the adhesion family of G-
420 protein-coupled receptors. *Genomics* **84**, 23–33 (2004).
421 2. Bondarev, A. D. *et al.* Opportunities and challenges for drug discovery in modulating
422 Adhesion G protein-coupled receptor (GPCR) functions.
423 *https://doi.org/10.1080/17460441.2020.1791075* 1291–1307 (2020).
424 doi:10.1080/17460441.2020.1791075
425 3. Weston, M. D., Luijendijk, M. W. J., Humphrey, K. D., Möller, C. & Kimberling, W. J.
426 Mutations in the VLGR1 gene implicate G-protein signalling in the pathogenesis of Usher
427 syndrome type II. *Am. J. Hum. Genet.* **74**, 357–366 (2004).
428 4. Boyden, S. E. *et al.* Vibratory Urticaria Associated with a Missense Variant in ADGRE2 . *N.*
429 *Engl. J. Med.* **374**, 656–663 (2016).
430 5. Krasnoperov, V. G. *et al.* α-Latrotoxin Stimulates Exocytosis by the Interaction with a
431 Neuronal G-Protein-Coupled Receptor. *Neuron* **18**, 925–937 (1997).
432 6. Araç, D. *et al.* A novel evolutionarily conserved domain of cell-adhesion GPCRs mediates
433 autoproteolysis. *EMBO J.* **31**, 1364–1378 (2012).
434 7. Liebscher, I., Schö, J., Monk, K. R. & Schö Neberg Correspondence, T. A Tethered Agonist
435 within the Ectodomain Activates the Adhesion G Protein-Coupled Receptors GPR126 and
436 GPR133. (2014). doi:10.1016/j.celrep.2014.11.036
437 8. Stoveken, H. M., Hajduczok, A. G., Xu, L. & Tall, G. G. Adhesion G protein-coupled
438 receptors are activated by exposure of a cryptic tethered agonist. *Proc. Natl. Acad. Sci. U.*
439 *S. A.* **112**, 6194–6199 (2015).
440 9. Ping, Y. Q. *et al.* Structures of the glucocorticoid-bound adhesion receptor GPR97–Go
441 complex. *Nat. 2020 5897843* **589**, 620–626 (2021).
442 10. Xiao, P. *et al.* Tethered peptide activation mechanism of the adhesion GPCRs ADGRG2 and
443 ADGRG4. *Nat. 2022 6047907* **604**, 771–778 (2022).
444 11. Ping, Y.-Q. *et al.* Structural basis for the tethered peptide activation of adhesion GPCRs.
445 *Nat. 2022 6047907* **604**, 763–770 (2022).
446 12. Barros-Álvarez, X. *et al.* The tethered peptide activation mechanism of adhesion GPCRs.

447 *Nat. 2022 6047907* **604**, 757–762 (2022).
448 13. Qu, X. *et al.* Structural basis of tethered agonism of the adhesion GPCRs ADGRD1 and
449 ADGRF1. *Nat. 2022* **604**, 1–8 (2022).
450 14. Fredriksson, R., Lagerström, M. C., Höglund, P. J. & Schiöth, H. B. Novel human G protein-
451 coupled receptors with long N-terminals containing GPS domains and Ser/Thr-rich regions.
452 *FEBS Lett.* **531**, 407–414 (2002).
453 15. Lum, A. M. *et al.* Orphan receptor GPR110, an oncogene overexpressed in lung and
454 prostate cancer. *BMC Cancer* **10**, 40 (2010).
455 16. Lee, J. W. *et al.* Orphan GPR110 (ADGRF1) targeted by N-docosahexaenoylethanolamine in
456 development of neurons and cognitive function. *Nat. Commun.* **7**, (2016).
457 17. Kwon, H. *et al.* Ligand-Induced GPR110 Activation Facilitates Axon Growth after Injury. *Int.*
458 *J. Mol. Sci.* **22**, (2021).
459 18. Stoveken, H. M., Larsen, S. D., Smrcka, A. V. & Tall, G. G. Gedunin- and khivorin-derivatives
460 are small-molecule partial agonists for adhesion G protein-coupled receptors
461 GPR56/ADGRG1 and GPR114/ADGRG5. *Mol. Pharmacol.* **93**, 477–488 (2018).
462 19. Demberg, L. M. *et al.* Activation of Adhesion G Protein-coupled Receptors: AGONIST
463 SPECIFICITY OF STACHEL SEQUENCE-DERIVED PEPTIDES*. *J. Biol. Chem.* **292**, 4383 (2017).
464 20. Dates, A. N. *et al.* A Fusion Protein Platform for Analyzing Tethered Agonism in the
465 Adhesion Family of G Protein-Coupled Receptors. *bioRxiv* 2022.07.14.500097 (2022).
466 doi:10.1101/2022.07.14.500097
467 21. Demberg, L. M. *et al.* Activation of adhesion G protein-coupled receptors: Agonist
468 specificity of Stachel sequence-derived peptides. *J. Biol. Chem.* **292**, 4383–4394 (2017).
469 22. Grundmann, M. *et al.* Lack of beta-arrestin signalling in the absence of active G proteins.
470 *Nat. Commun. 2018 91* **9**, 1–16 (2018).
471 23. Hauser, A. S. *et al.* Common coupling map advances GPCR-G protein selectivity. *Elife* **11**,
472 (2022).
473 24. Nehmea, R. *et al.* Mini-G proteins: Novel tools for studying GPCRs in their active
474 conformation. *PLoS One* **12**, e0175642 (2017).
475 25. Carpenter, B. & Tate, C. G. Engineering a minimal G protein to facilitate crystallisation of G

476 protein-coupled receptors in their active conformation. *Protein Eng. Des. Sel.* **29**, 583
477 (2016).
478 26. Hall, M. P. *et al.* Engineered luciferase reporter from a deep sea shrimp utilizing a novel
479 imidazopyrazinone substrate. *ACS Chem. Biol.* **7**, 1848–1857 (2012).
480 27. Wan, Q. *et al.* Mini G protein probes for active G protein–coupled receptors (GPCRs) in
481 live cells. *J. Biol. Chem.* **293**, 7466 (2018).
482 28. Hauser, A. S. *et al.* Common coupling map advances GPCR-G protein selectivity. *Elife* **11**,
483 (2022).
484 29. Dates, A. N. *et al.* A Fusion Protein Platform for Analyzing Tethered Agonism in the
485 Adhesion Family of G 5 Protein-Coupled Receptors 6 7 8 9.
486 doi:10.1101/2022.07.14.500097
487 30. Dixon, A. S. *et al.* NanoLuc Complementation Reporter Optimized for Accurate
488 Measurement of Protein Interactions in Cells. *ACS Chem. Biol.* **11**, 400–408 (2016).
489 31. Liu, H. & Naismith, J. H. An efficient one-step site-directed deletion, insertion, single and
490 multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.* **8**, 91 (2008).
491 32. Rasmussen, S. G. F. *et al.* Crystal structure of the β2 adrenergic receptor–Gs protein
492 complex. *Nat. 2011 4777366* **477**, 549–555 (2011).
493 33. Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion
494 correction in cryo-EM single-particle analysis. *urn:issn:2052-2525* **6**, 5–17 (2019).
495 34. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron
496 micrographs. *J. Struct. Biol.* **192**, 216 (2015).
497 35. Li, Y., Cash, J. N., Tesmer, J. J. G. & Cianfrocco, M. A. High-Throughput Cryo-EM Enabled by
498 User-Free Preprocessing Routines. *Structure* **28**, 858-869.e3 (2020).
499 36. Bepler, T. *et al.* Positive-unlabeled convolutional neural networks for particle picking in
500 cryo-electron micrographs. *Nat. Methods 2019 1611* **16**, 1153–1160 (2019).
501 37. Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM structure
502 determination. *J. Struct. Biol.* **180**, 519 (2012).
503 38. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization
504 improves single-particle cryo-EM reconstruction. *Nat. Methods 2020 1712* **17**, 1214–1221

505 (2020).
506 39. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid
507 unsupervised cryo-EM structure determination. *Nat. Methods 2017 143* **14**, 290–296
508 (2017).
509 40. Sanchez-Garcia, R. *et al.* DeepEMhancer: a deep learning solution for cryo-EM volume
510 post-processing. *Commun. Biol. 2021 41* **4**, 1–8 (2021).
511 41. Croll, T. I. ISOLDE: A physically realistic environment for model building into low-resolution
512 electron-density maps. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 519–530 (2018).
513 42. Afonine, P. V. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography.
514 *urn:issn:2059-7983* **74**, 531–544 (2018).
515 43. Yamashita, K., Palmer, C. M., Burnley, T. & Murshudov, G. N. Cryo-EM single-particle
516 structure refinement and map calculation using Servalcat. *Acta Crystallogr. Sect. D, Struct.*
517 *Biol.* **77**, 1282–1291 (2021).
518 44. Morin, A. *et al.* Collaboration gets the most out of software. *Elife* **2**, (2013).
519

Data Figures

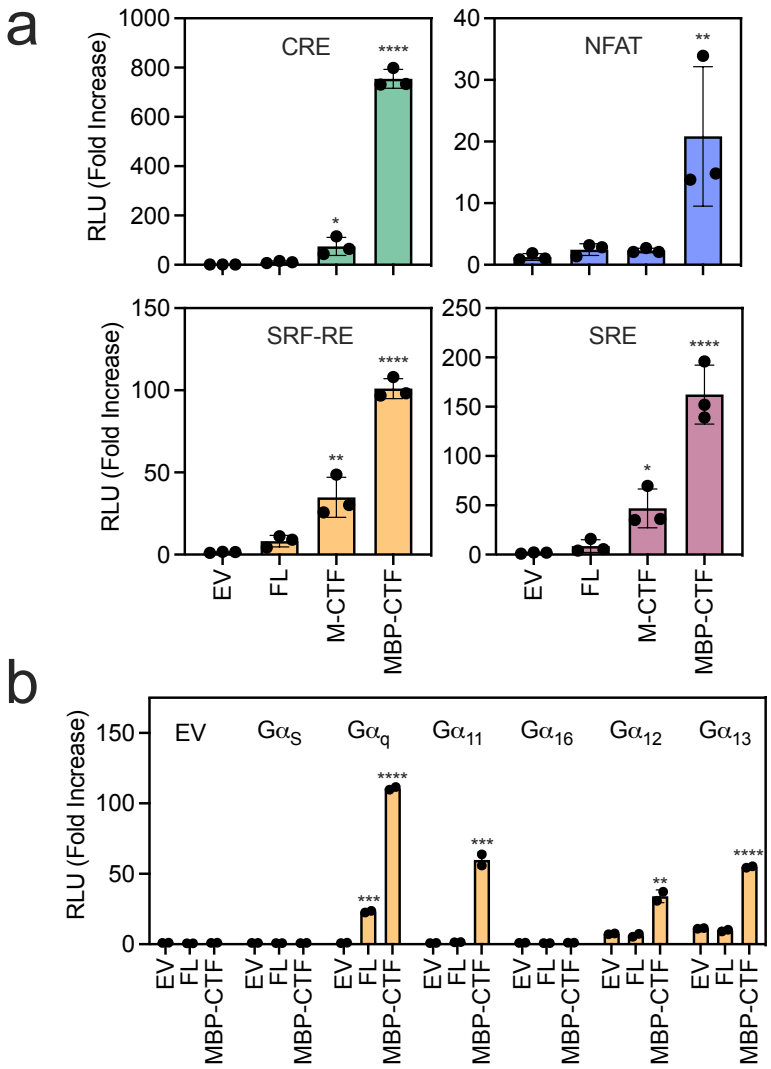
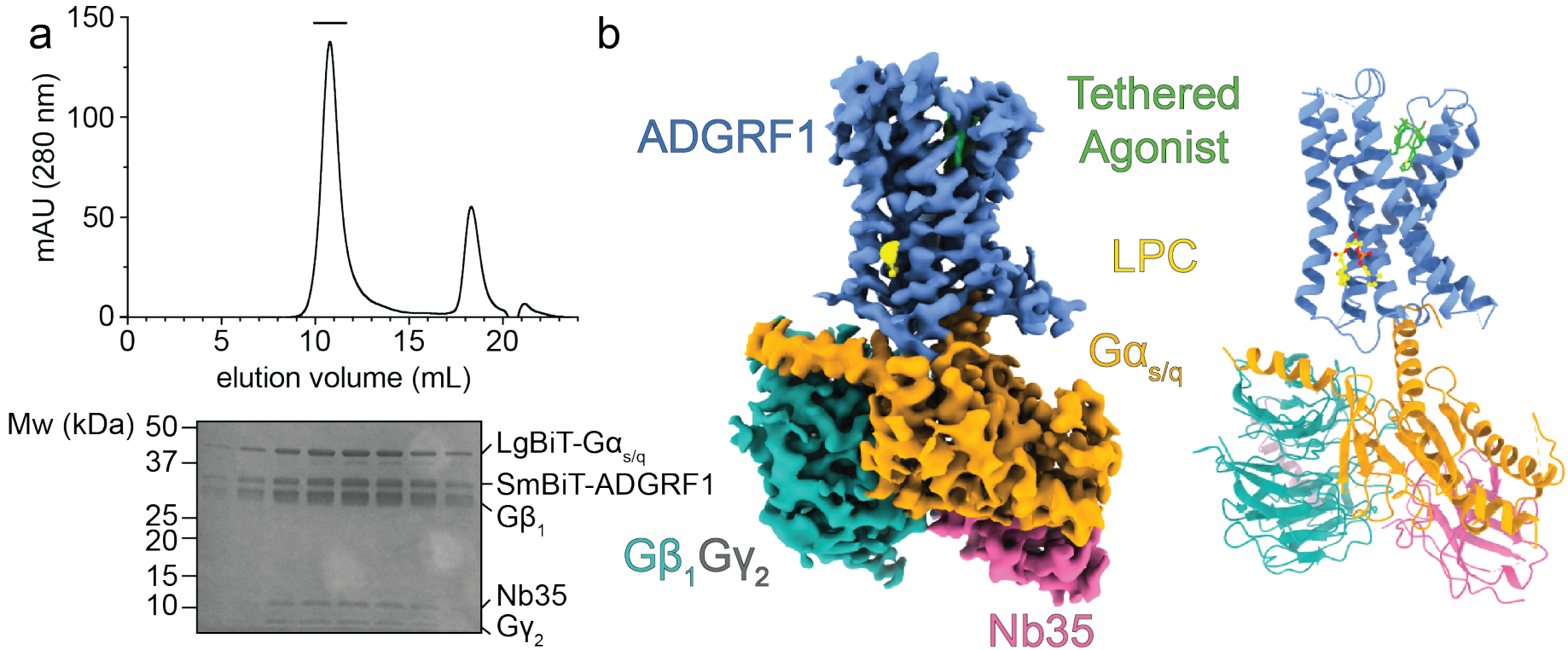


Figure 1 – ADGRF1 signals through all G-protein classes. (a) Transcription factor reporter assay responses for ADGRF1 FL, M-CTF or MBP-CTF relative to empty vector for CRE (top left), NFAT (top right), SRF-RE (bottom left) and SRE (bottom right). Cells were transfected with 30 ng DNA/well for SRE and SRF-RE, and 10 ng DNA/well for CRE and NFAT. Data are normalized to empty vector as relative luminescence units (RLU). Data are presented as mean ± s.d. of three biological replicates. One-way analysis of variance (ANOVA) was used with Dunnett's multiple-comparison post-hoc test to compare the difference between samples to each reporter's empty vector control (*P≤0.05, **P≤0.01, and ****P≤0.0001). (b) Gα transient transfection complementation assay. Activity of the SRF-RE reporter was measured in HEKΔ6 cells transfected with a combination of ADGRF1 and different Gα subunits. The Gα subunit or Gα empty vector control was co-transfected with receptor empty vector, ADGRF1 FL or MBP-CTF. Values are reported as the ratio of the raw luminescence values of firefly and renilla luciferases. Data are presented as mean ± s.d. of two biological replicates. Each point represents the mean value of three technical replicates. In all panels, One-way analysis of variance (ANOVA) was used with Dunnett's multiple-comparison post-hoc test to compare the difference between samples to each Gα empty vector control (*P≤0.05, **P≤0.01, ***P≤0.001 and ****P≤0.0001).

537



538 **Figure 2 – ADGRF1-miniGα$_{s/q}$-Nb35 complex purification and cryo-EM structure.** (a) ADGRF1 miniGα$_{s/q}$-Gβ$_1$γ$_2$-Nb35 purification. Top: elution profile
539 of the complex after size-exclusion chromatography as a final purification step. Bottom: peak fractions (bar above chromatogram) analyzed by
540 SDS-PAGE. (b) Structure of the ADGRF1-Gα$_{s/q}$-β$_1$γ$_2$-Nb35 complex. Left: cryo-EM map of the ADGRF1 complex. ADGRF1 (light blue), the TA (green),
541 Gα$_{s/q}$ (orange), Gβ$_1$ (teal), Gγ$_2$ (grey) and Nb35 (pink) are coloured as indicated. The density threshold was set at five standard deviations from the
542 mean. Right: cartoon rendering of the structure, with residue side chains of the tethered agonist rendered as sticks. The density threshold was set
543 at three standard deviations from the mean for the TA. The LPC lipid (yellow) molecule is shown with a ball-and-stick model.

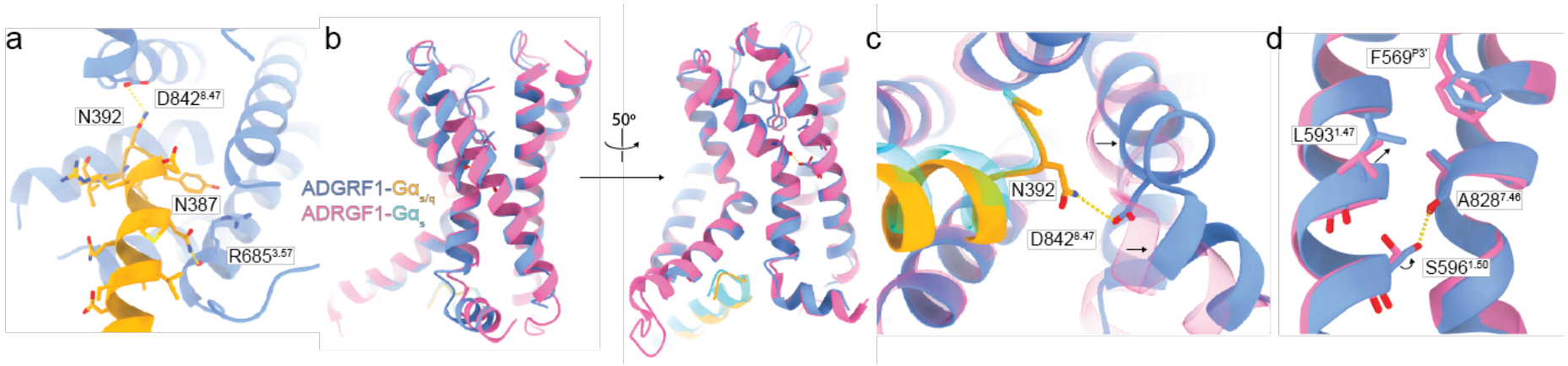


Figure 3 – The ADGRF1-miniG$\alpha_{s/q}$ interface and comparison to ADGRF1-miniGα_s. (a) Close-up of the ADGRF1 (blue) interface with G$\alpha_{s/q}$ (orange) along the Cα5 helix of G$\alpha_{s/q}$. Hydrogen bonding interactions between the receptor and Gα are indicated with yellow dashed lines. (b) Comparison of the ADGRF1 (blue) - G$\alpha_{s/q}$ (orange) complex with the ADGRF1 (pink) - miniGα_s (cyan) complex, aligned by least-squares superposition of the ADGRF1 subunits. (c) Zoomed in view focusing on D842$^{8.47}$ in the turn connecting helices VII and VIII. The hydrogen bond between N392 and the carboxylate of D842$^{8.47}$ is shown with a yellow dotted line. (d) Zoomed in view focusing on residues L593$^{1.47}$ and S596$^{1.50}$. Arrows indicate positional shifts of these residues in the miniG$\alpha_{s/q}$ structure relative to their positions in miniGα_s. The hydrogen bond between S596$^{1.50}$ and the backbone carbonyl of A828$^{7.46}$ is shown with a yellow dotted line.

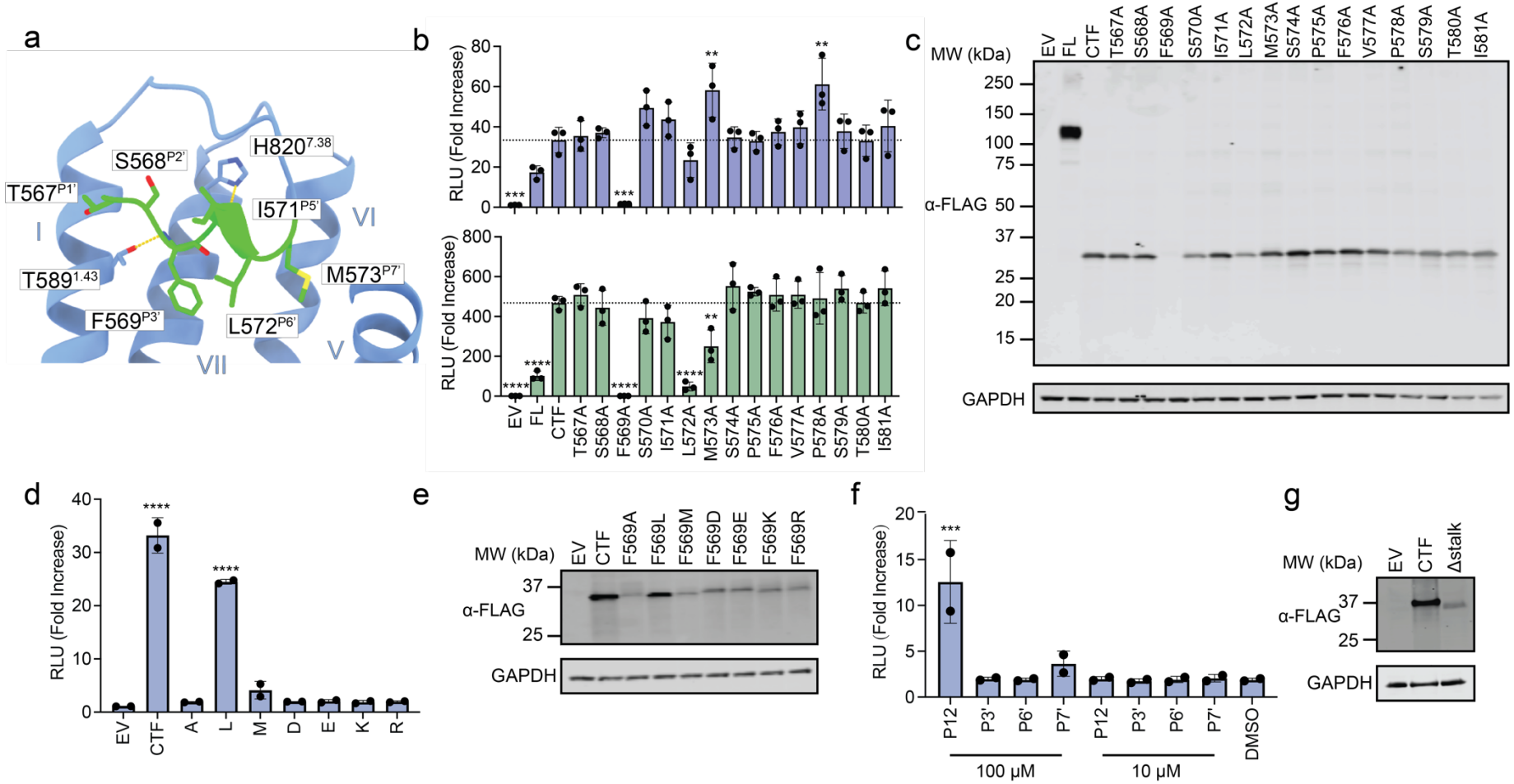


Figure 4 – Tethered agonism signalling determinants through NFAT and CRE. (a) Pose of the TA in the orthosteric site. TA sidechain atoms and main-chain atoms making polar contacts are shown as sticks. Side-chains of 7TM residues making polar contacts (yellow dashes) to the TA are also shown. TA superscript labels refer to the TA position relative to the GPS autoproteolysis site. Receptor residue superscript labels refer to Wootten numbering. (b) Scanning alanine mutagenesis of the TA. NFAT and CRE reporter data were acquired using ADGRF1 CTF as a reference sequence. Cells were transfected with 20 ng of receptor DNA/well. Data are normalized to empty vector and reported as relative luminescence units (RLU). The mean for the CTF is indicated as a horizontal dotted line. *Figure legend continued overleaf.*

557 (c) α-FLAG Western blot comparing steady-state amounts of wild-type and mutated ADGRF1 CTF proteins tested in panel b. An α-GAPDH blot was
558 used as a loading control. (d) NFAT signalling data and (e) α-FLAG Western blot of wild-type and P3' mutated forms (L, M, R, K, D and E) of the
559 ADGRF1 CTF. (f) NFAT reporter gene activity of the ADGRF1 CTFΔstalk protein stimulated by *trans* addition of a 12-residue TA peptide (P12). The
560 response to the wild-type peptide or to peptides with alanine mutations at the P3', P6' or P7' position are shown at doses of 10 or 100 μM. (g) α-
561 FLAG Western blot comparing steady-state amounts of wild-type and Δstalk ADGRF1 CTF proteins tested in panel f. An α-GAPDH blot was used as
562 a loading control. For panel (b), data are presented as mean ± s.d. of three biological replicates. Each point represents the mean value of three
563 technical replicates. For panels (d and f) data are presented as mean ± s.d. of two biological replicates. Each point represents the mean value of
564 three technical replicates. One-way analysis of variance (ANOVA) was used with Dunnett's multiple-comparison post-hoc test to compare the
565 difference between CTF to all other samples individually in b and d, and to DMSO in f (**P<0.01;***P<0.001;****P<0.0001).

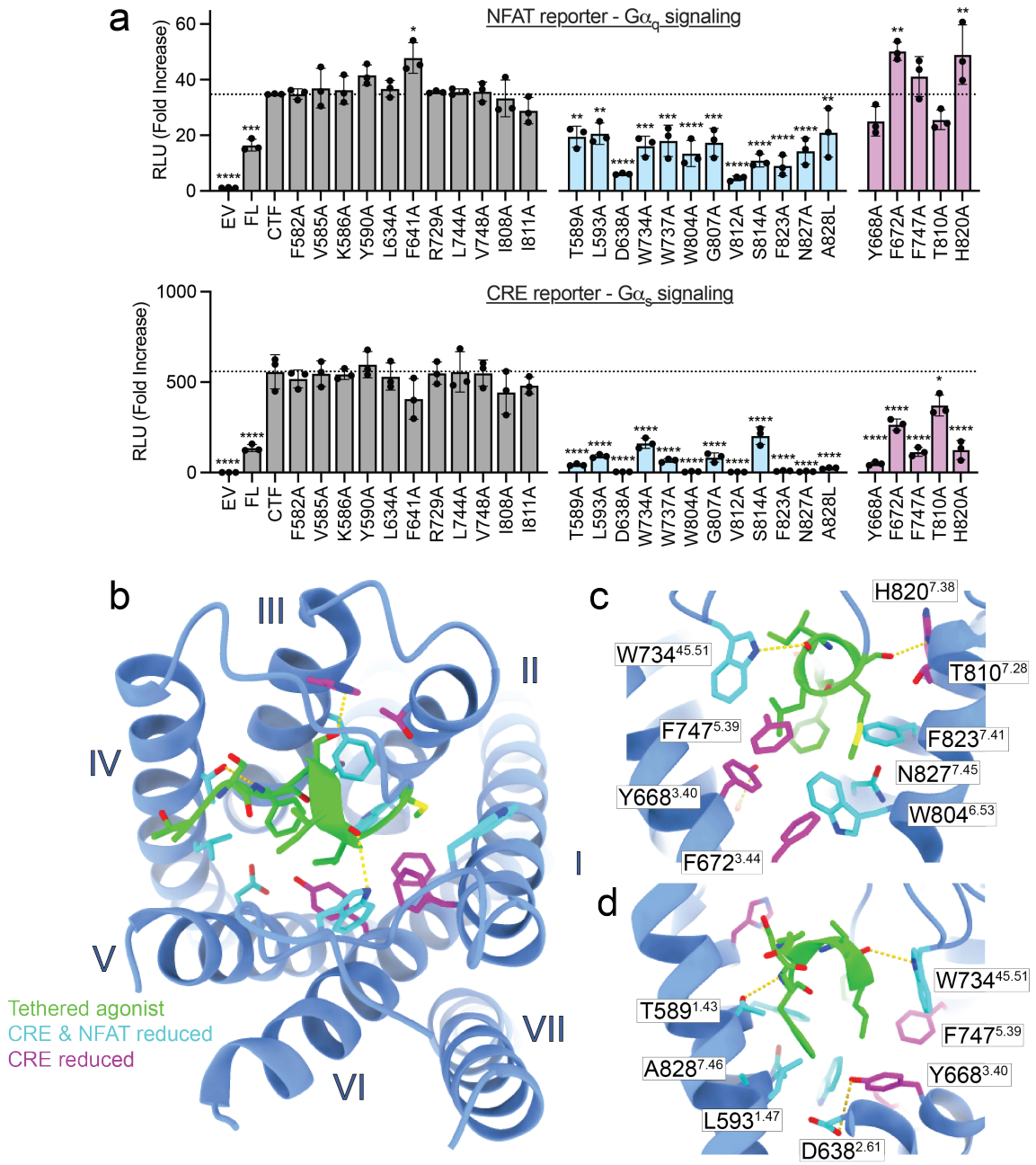
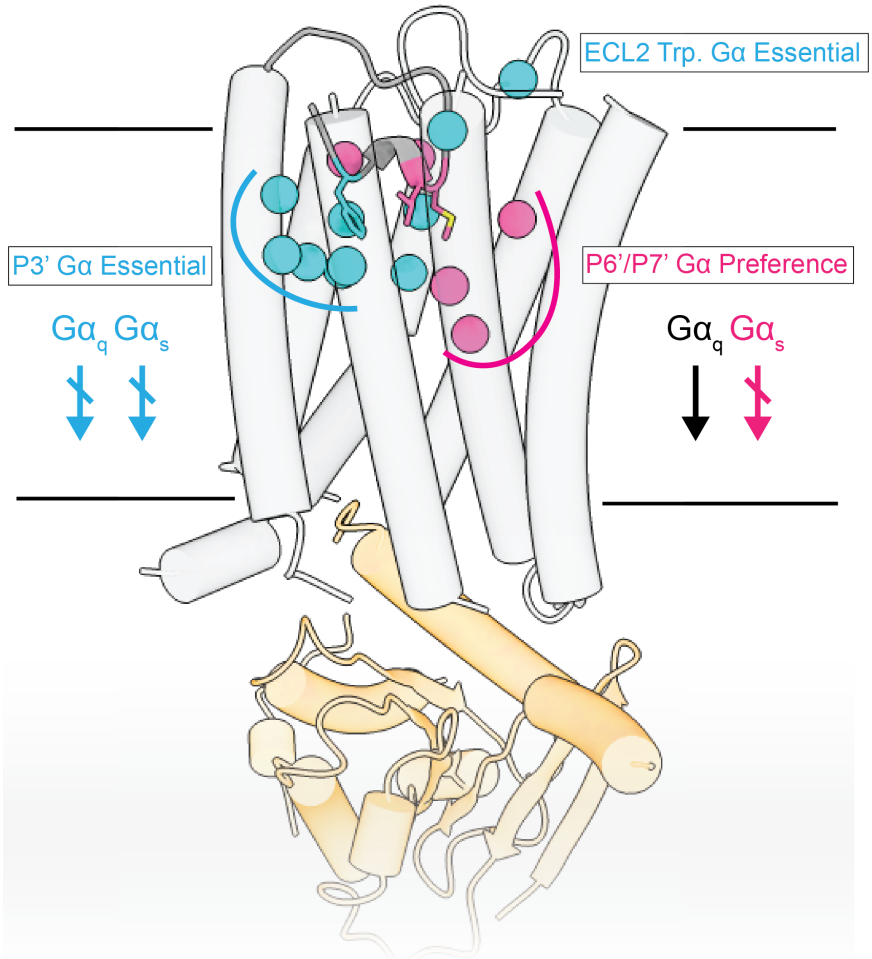


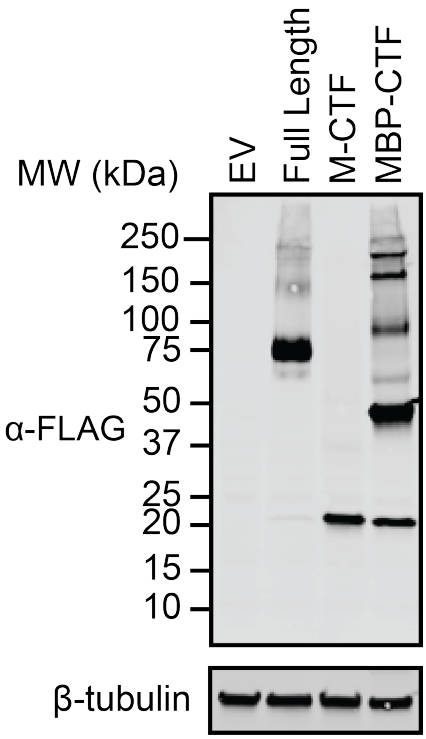
Figure 5 – Binding site signalling determinants through NFAT and CRE. (a) NFAT and CRE signalling data for TA binding site mutants in the context of the ADGRF1 CTF protein. Cells were transfected with 20 ng of receptor DNA/well. Data are normalized to empty vector and reported as relative luminescence units (RLU). Data are grouped according to their effects on NFAT and CRE signals relative to the wild-type ADGRF1 CTF protein. Mutations without significant effects on either reporter are grey, mutations that significantly reduce both NFAT and CRE signals are cyan, and mutations that significantly reduce only the CRE signal are magenta. Data are presented as mean ± s.d. of three biological replicates. The mean for the CTF is indicated as a horizontal dotted line. Each point represents the mean value of three technical replicates. One-way analysis of variance (ANOVA) was used with Dunnett's multiple-comparison post-hoc test to compare the difference between CTF to all other samples individually (*P≤0.05, **P≤0.01, ***P≤0.001 and ****P≤0.0001). *Figure legend continued overleaf.*

(b) Top view of the TA in the orthosteric site. TA sidechain atoms and main-chain atoms making polar contacts (yellow dashes) are shown as sticks. Side-chains of 7TM residues that affect signalling when mutated to alanine are shown as sticks and are coloured according to the grouping in (a). (c) View from the P7' end and (d) the P3' end of the TA. Side-chains of 7TM residues that affect signalling when mutated to alanine are shown as sticks, coloured according to the grouping in (a), and annotated by residue with Wooten superscripts.

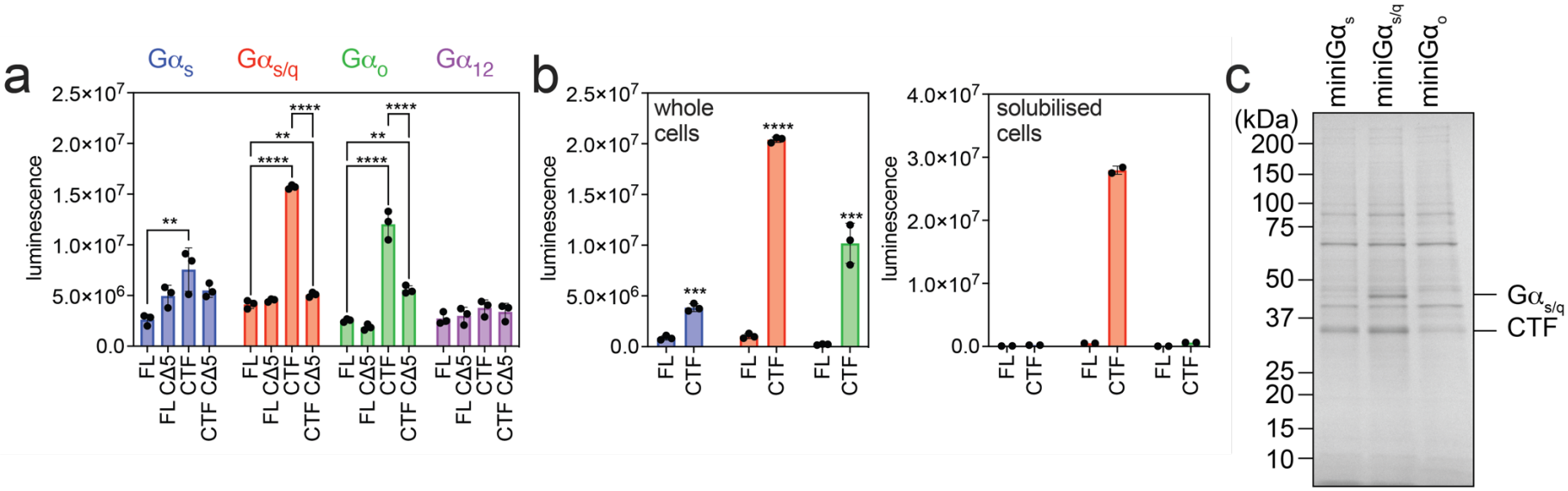


585 **Figure 6 – Mutations of the TA and associated binding site cluster by G-protein class**
586 **preference.** ADGRF1 coupled to miniG$\alpha_{s/q}$ represented as tubed helices, colored light grey and
587 orange respectively. Residues that reduced signaling activity are shown, with TA residues
588 rendered as sticks and with Cα atoms of residues in the TA binding site rendered as spheres.
589 Residues that reduce reduce both Gα_s and Ga_q signalling are colored cyan, whereas residues
590 that only reduce Gα_s signaling are magenta.
591

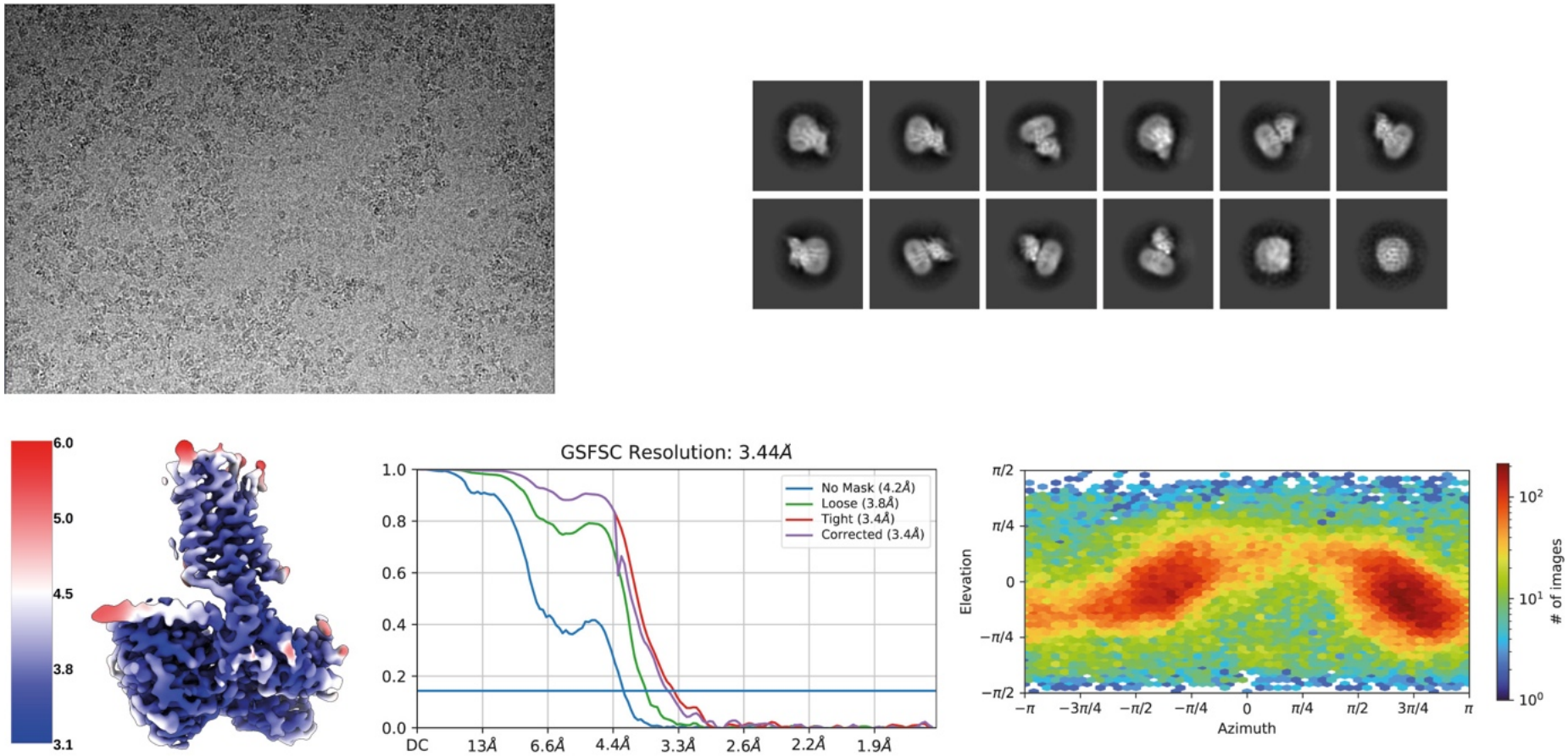
592 Extended Data



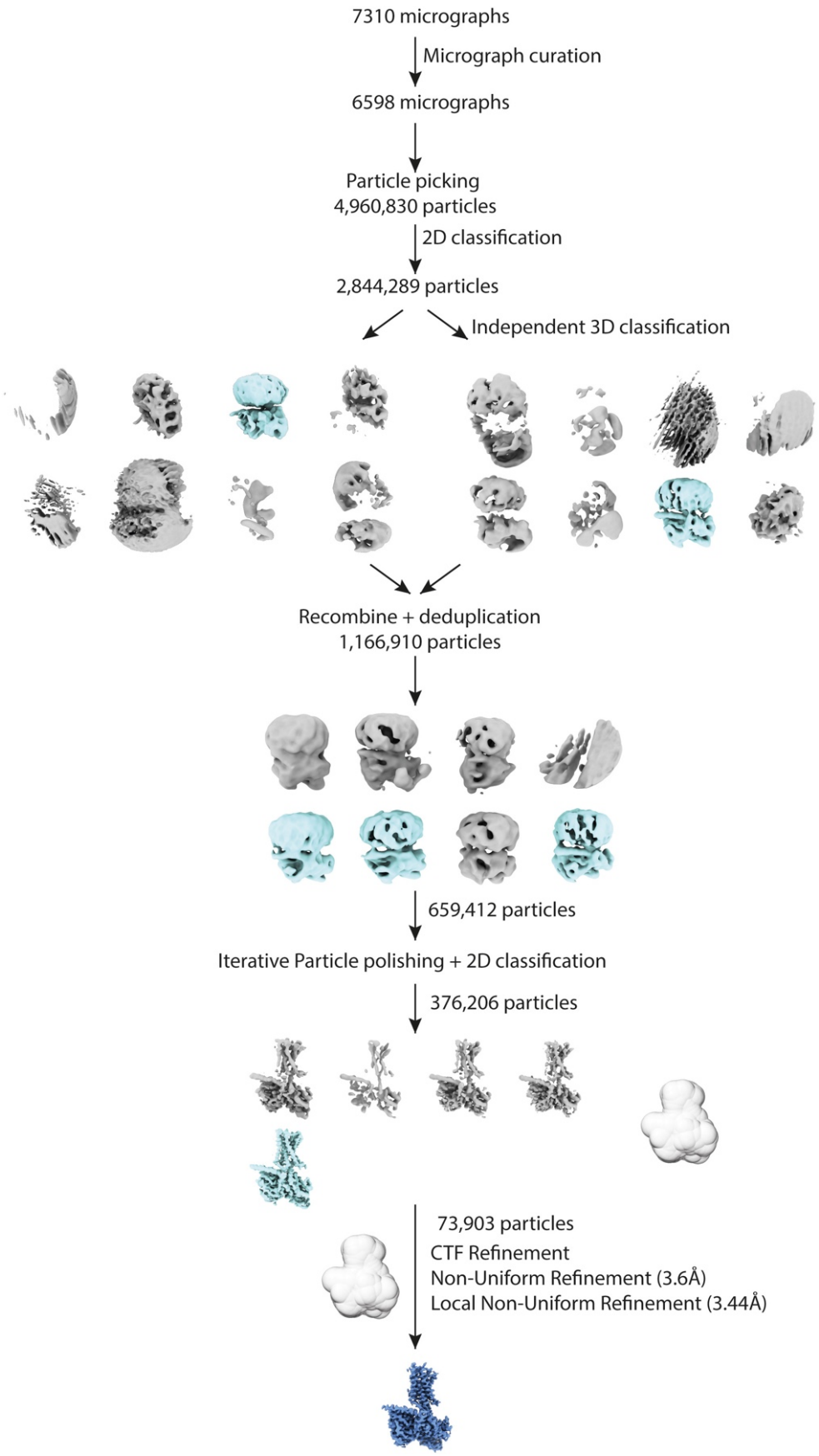
593 **Extended Data Figure 1 – Expression of ADGRF1 proteins used in signalling experiments with**
594 **NFAT, CRE, SRE and SRF-RE reporters.** α-FLAG Western blot of ADGRF1 proteins used in
595 signalling assays to identify G-protein coupling partners. A Western blot with an anti-β-tubulin
596 antibody was used as a loading control.



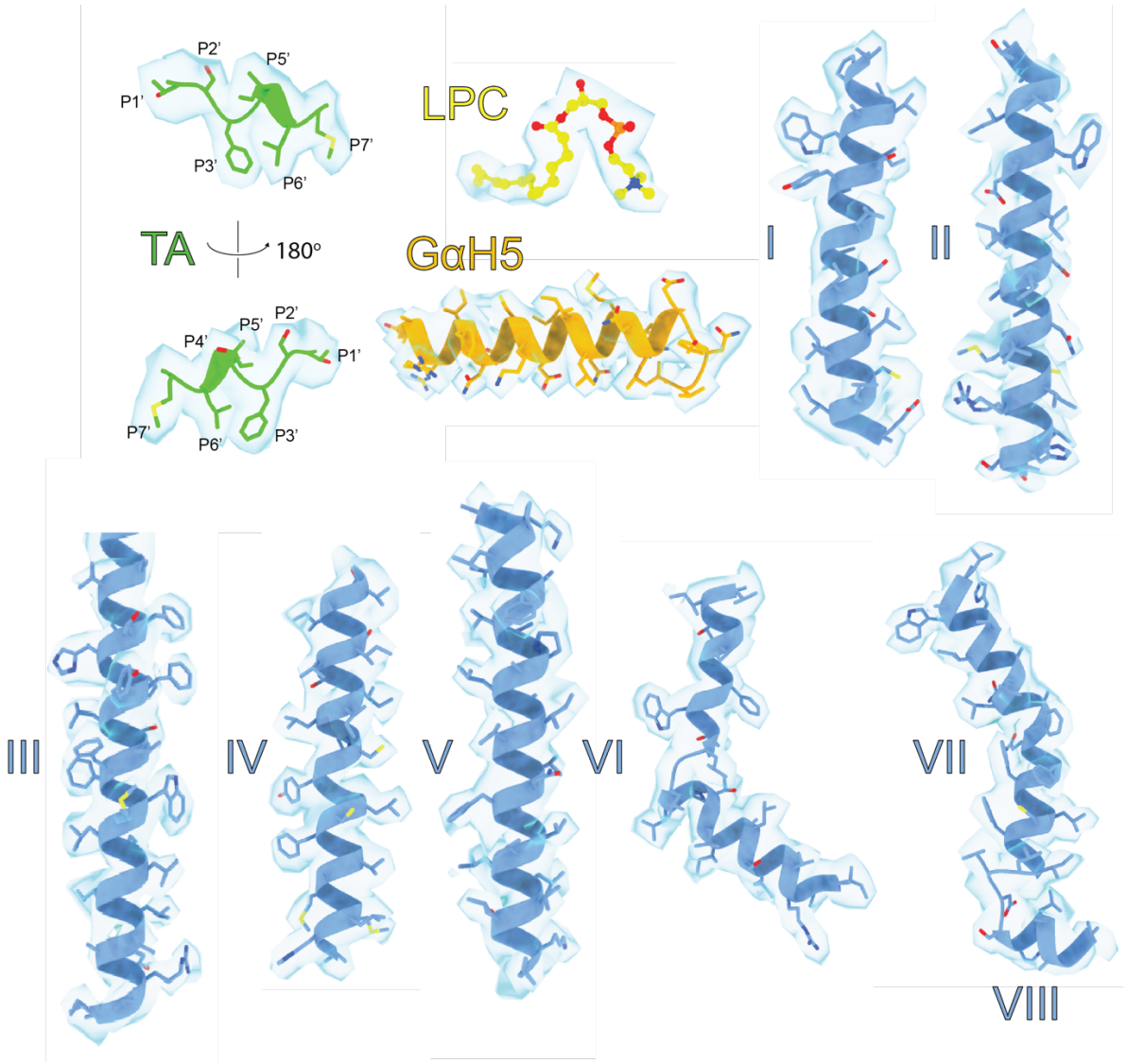
Extended Data Figure 2 – NanoBiT complementation of GPR110-CTF with miniGα proteins. (a) NanoBiT recruitment assays for assessment of receptor-miniG coupling efficiency. Adherent HEK293T cells were transfected with equal amounts of DNA for receptor and miniG proteins and luminescence recorded after the addition of substrate. Full-length (FL) and C-terminal fragment (CTF) variants of the receptor were combined with either functional miniG or non-functional miniG CΔ5 controls. Luminescence resulting from luciferase reconstitution is plotted, with bars grouped and coloured by miniG class. One-way analysis of variance (ANOVA) was used with Dunnett's multiple-comparison post-hoc test to compare the difference between FL and CΔ5 to all other samples within each miniG group individually (*P≤0.05, **P≤0.01, ***P≤0.001 and ****P≤0.0001). (b) Suspension Expi293T cells were transfected with equal amounts of DNA for receptor and miniG proteins in whole cell (left) and solubilised formats (right). Luminescence resulting from luciferase reconstitution is plotted, with bars coloured by miniG class as in (a). Unpaired two-tailed T-test was used to compare the difference between FL and CTF for whole cell luminescence (***P≤0.001 and ****P≤0.0001). (c) Coomassie blue staining after SDS-PAGE of ADGRF1-CTF complexes recovered on magnetic beads, assessing coupling efficiency to miniG proteins.



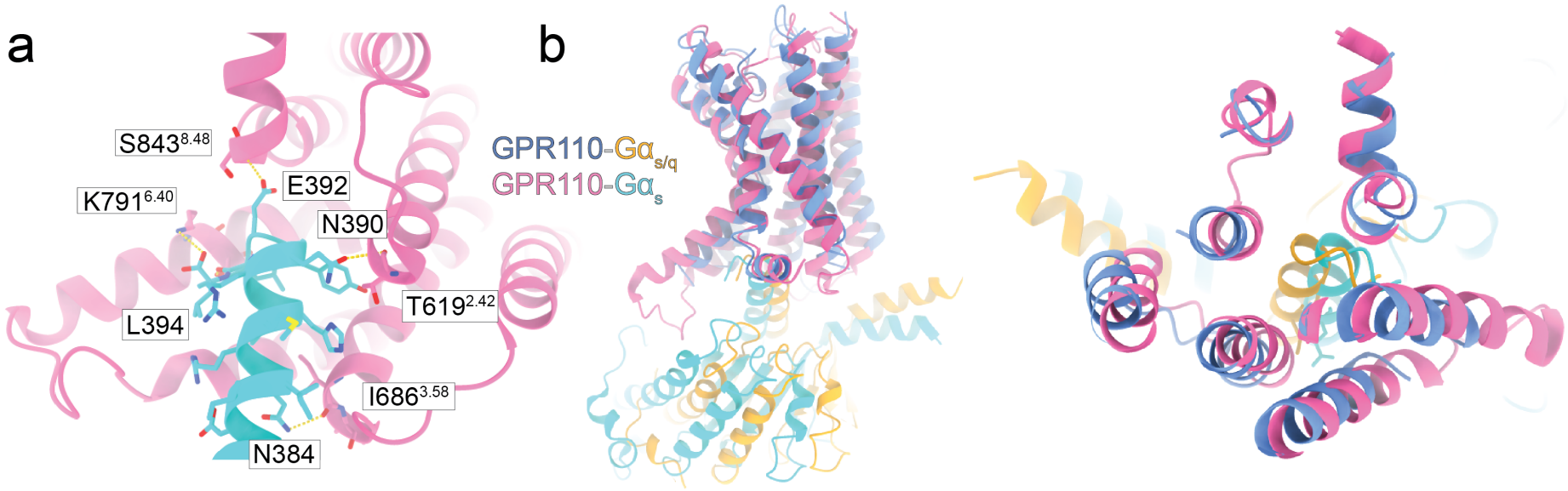
Extended Data Figure 3 – Cryo-EM data quality. (Top, left) A representative image of ADGRF1 complexes in vitreous ice visualized by cryo-EM on a Titan Krios microscope equipped with Gatan K3 detector. (Top, right) 2D-class averages showing a range of particle orientations. (Bottom, left) Local resolution map generated in Cryosparc (GS-FSC = 0.143 cut-off). (Bottom, centre) GS-FSC curves with default Cryosparc masks. (Bottom, right) Particle orientation distribution map of ADGRF1-miniG$\alpha_{s/q}$ complexes used for structure determination.



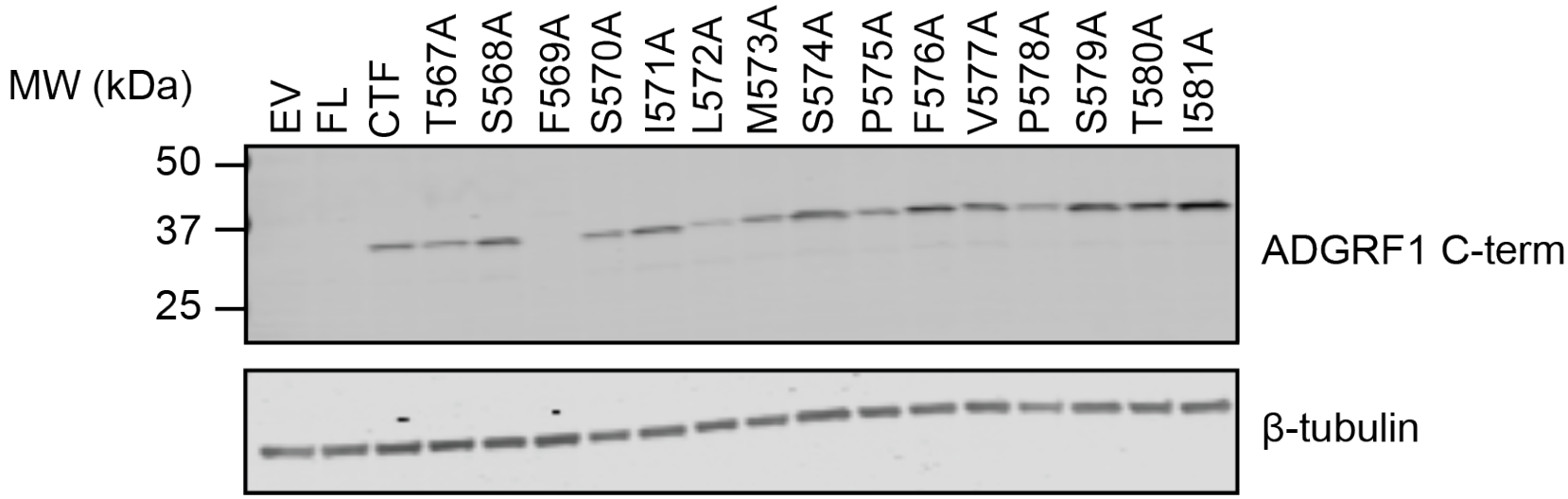
611 **Extended Data Figure 4 – Cryo-EM processing scheme for ADGRF1 reconstruction.** Selected
612 classes are blue and discarded classes are grey.



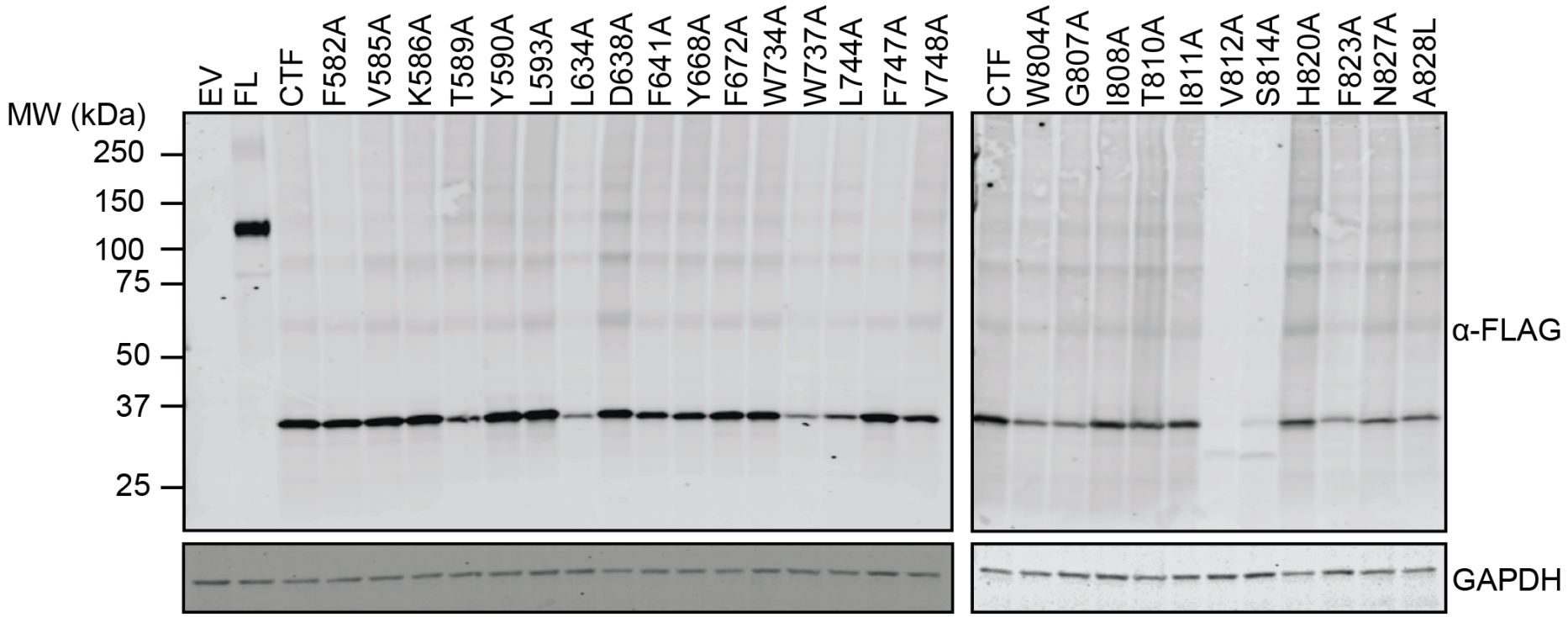
Extended Data Figure 5 – Map to model density of key ADGRF1 components. All protein structures are rendered in cartoon with residue side chains represented as sticks. The LPC molecule is rendered as ball-and-sticks. TA (green); tethered agonist; LPC (yellow), lysophosphatidylcholine (16:0), GαH5 (orange); C-terminal alpha helix of miniG$\alpha_{s/q}$, I-VII (blue) transmembrane helices I-VII, and VIII (blue); helix VIII. The density threshold was set at three standard deviations from the mean for all components.



Extended data Figure 6 – Comparison of ADGRF1 bound to miniGα$_{s/q}$ and miniGα$_{s}$. a) Close-up of the ADGRF1 (pink) interface with Gα$_{s}$ (cyan) along the Cα5 helix of Gα$_{s}$. Hydrogen bonds between the receptor and Gα are indicated. (b) Left: side-view, and Right: top-slice view of ADGRF1 bound to miniGα$_{s}$ or miniGα$_{s/q}$. The receptors are aligned by least-squares superposition of TM helices I and VII.



Extended Data Figure 7 – Amount of ADGRF1 CTF protein expressed by wild-type and alanine scanning mutants. Protein was detected by Western blotting with an α-ADGRF1 C-terminal antibody.



Extended Data Figure 8 – Amount of ADGRF1 CTF protein expressed by wild-type and alanine scanning mutants of the binding site. Protein was detected by Western blotting with an α-FLAG antibody, which recognizes an N-terminal FLAG tag. A Western blot with an anti-GAPDH antibody was used as a loading control. A Western blot with an anti-GAPDH antibody was used as a loading control.

627 Table 1 | Cryo-EM data collection, refinement and validation statistics.

ADGRF1 (EMD-XXX, PDB XXX)	
Data collection and processing	
Voltage (kV)	300
Electron exposure ($e^-/Å^{-2}$)	~53.5
Defocus range (μm)	0.5 – 1.5
Pixel size (Å)	0.825
Symmetry imposed	C1
Initial particle images (no.)	4,960,830
Final particle images (no.)	73,903
Map Resolution (Å)	3.44
FSC threshold	0.143
Map resolution range (Å)	3.1 - 6.0
Refinement	
Initial model used	7wu3
Model resolution (Å)	3.10
FSC threshold	0.143
Map-sharpening B-factor ($Å^2$)	-111.2
Model composition	
Non-hydrogen atoms	6,671
Protein residues	891
B factors ($Å^2$)	
Protein	74.17
Ligand	97.56
R.M.S.D. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.594
Validation	
MolProbity score	1.63
Clashscore	5.21
Poor rotamers (%)	0.00
Ramachandran plot	
Favoured (%)	94.88
Allowed (%)	5.12
Disallowed (%)	0.00