

# Tethered agonist activated ADGRF1 structure and signalling analysis reveal basis for G protein coupling

Received: 30 September 2022

Accepted: 14 April 2023

Published online: 29 April 2023

 Check for updates

Daniel T. D. Jones<sup>1</sup>✉, Andrew N. Dates<sup>1</sup>, Shaun D. Rawson<sup>1</sup>,  
Maggie M. Burruss<sup>1</sup>, Colin H. Lipper<sup>1</sup> & Stephen C. Blacklow<sup>1,2</sup>✉

Adhesion G Protein Coupled Receptors (aGPCRs) have evolved an activation mechanism to translate extracellular force into liberation of a tethered agonist (TA) to effect cell signalling. We report here that ADGRF1 can signal through all major G protein classes and identify the structural basis for a previously reported  $G\alpha_q$  preference by cryo-EM. Our structure shows that  $G\alpha_q$  preference in ADGRF1 may derive from tighter packing at the conserved F569 of the TA, altering contacts between TM helix I and VII, with a concurrent rearrangement of TM helix VII and helix VIII at the site of  $G\alpha$  recruitment. Mutational studies of the interface and of contact residues within the 7TM domain identify residues critical for signalling, and suggest that  $G\alpha_s$  signalling is more sensitive to mutation of TA or binding site residues than  $G\alpha_q$ . Our work advances the detailed molecular understanding of aGPCR TA activation, identifying features that potentially explain preferential signal modulation.

Adhesion G protein coupled receptors (aGPCRs) are the second largest class of GPCRs, constituting 33 members across 9 subfamilies<sup>1</sup>. 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<sup>2</sup>. Mutations in aGPCRs are genetically responsible for human developmental disorders, such as ADGRV1 in Usher syndrome<sup>3</sup>, and EMR2 in vibratory urticaria<sup>4</sup>.

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<sup>5,6</sup>. 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<sup>7,8</sup>. Current models for aGPCR activation posit that mechanical force applied to the adhesive modules in aGPCRs induces dissociation of the NTF from the CTF<sup>9,10</sup>, 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<sup>11</sup>. 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  $G\alpha$ <sup>12–16</sup>. 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 (TØFØØLM) of the TA in the GPCR orthosteric site.

ADGRF1 (GPR110) is an aGPCR belonging to aGPCR subfamily VI<sup>17</sup>. ADGRF1 was identified as an oncogene overexpressed in lung and prostate cancer<sup>18</sup>, and has been implicated in synaptamide dependent neural outgrowth and repair<sup>19,20</sup>, though this activity remains controversial<sup>21</sup>. ADGRF1 was used as a prototypical aGPCR to

<sup>1</sup>Department of Biological Chemistry and Molecular Pharmacology, Blavatnik Institute, Harvard Medical School, Boston, MA 02115, USA. <sup>2</sup>Department of Cancer Biology, Dana Farber Cancer Institute, Boston, MA 02215, USA. ✉e-mail: [daniel\\_jones@hms.harvard.edu](mailto:daniel_jones@hms.harvard.edu); [stephen\\_blacklow@hms.harvard.edu](mailto:stephen_blacklow@hms.harvard.edu)

determine that membranes expressing ADGRF1 treated with urea at high concentration induced shedding of the NTF, liberating the TA to activate the 7TM and recruit  $G\alpha_q$  preferentially over other  $G\alpha$  proteins<sup>10</sup>. Further work has also showed that mouse ADGRF1 can be activated by synthetic TA peptides when added in trans, signalling through both  $G\alpha_s$  and  $G\alpha_q$ <sup>22</sup>.

Recent structures of the TA-engaged state of the ADGRF1 CTF coupled to either mini $G\alpha_s$  or mini $G\alpha_q$  also identified a lysophosphatidylcholine (LPC) lipid bound at the intracellular convergence of TM helices II, III and IV<sup>15</sup>. Interestingly, fragments of ADGRF1 expressed with the GAIN domain included also yielded TA-engaged structures of ADGRF1 coupled to miniG but lacked density for the GAIN domain, suggesting that the CTF may have dissociated from the NTF portion of the protein during purification. Additionally, using a HiBiT tethering approach, structures of ADGRF1 have been solved bound to a representative miniG of each G protein class, raising the possibility that ADGRF1 is a promiscuous G protein coupler<sup>16</sup>, but the HiBiT tethering approach artificially recruits receptors to G proteins, and thus the ADGRF1 G protein signalling profile remains unclear.

Here, we extensively characterise ADGRF1 G protein signalling. First, using transcription factor assays and deletion complementation by transient transfection in HEK293T cells lacking certain G proteins, we deconvolute the G proteins responsible for CRE, NFAT-RE and SRF signalling in cells. As a more direct measurement of G protein coupling, we show using TRUPATH G protein biosensors, that ADGRF1 can convert a representative member of each G protein class to an active state. To understand the molecular events driving TA activation of ADGRF1 and how its structure determines  $G\alpha$  class preference, we determined a cryo-EM structure of TA-activated ADGRF1 coupled to mini $G\alpha_{s/q}\beta_1\gamma_2$ , stabilised by Nb35. Combining structural insights with extensive functional studies, we identify key residues of the orthosteric binding site that are essential for  $G\alpha_s$  signalling but dispensable for  $G\alpha_q$  signalling. Overall, our work complements other recent aGPCR structures<sup>12–16</sup> and extends them by visualizing ADGRF1 in complex with its most relevant G protein partner in a non-tethered complex, elucidating a functional and structural landscape by which ADGRF1 is activated and uncovering how its structure determines  $G\alpha$  class preference. Furthermore, the mechanistic insights determined here regarding TA engagement to strong and weakly associated G proteins offers valuable insight into how TA mediated activation of aGPCRs may be functionally modulated.

## Results

### ADGRF1 signals through all G protein classes

We recently developed a platform for analysis of the signalling activity of aGPCRs, in which the CTF of interest is expressed as a protein fusion with an N-terminal IL2 signal sequence followed by maltose binding protein (MBP) and a tobacco etch virus cleavage site (denoted MBP-CTF)<sup>23</sup> (Supplementary Fig. 1). A FLAG tag is also appended to the C-terminal end of the protein. Using this approach to measure signalling activity in transcription factor reporter assays, we observed transcriptional responses with both CRE and NFAT-RE reporters, consistent with previous studies<sup>15,22</sup>, as well as with SRE and SRF-RE reporters (Fig. 1a).

To further evaluate the coupling spectrum of ADGRF1, and to investigate crosstalk of G proteins as inducers of the various transcriptional responses tested, we used HEK293T cells with deletion of six  $G\alpha$  subunits ( $G\alpha_s$  short,  $G\alpha_s$  long,  $G\alpha_q$ ,  $G\alpha_{11}$ ,  $G\alpha_{12}$  and  $G\alpha_{13}$ ; gift from A. Inoue, Tohoku university, Japan)<sup>24</sup>, and performed deletion complementation by transient transfection to determine which  $G\alpha$  subunits support ADGRF1 signalling in the CRE, NFAT-RE and SRF-RE reporter assays (Fig. 1b).

In CRE reporter assays with these cells, we observed a significant increase in basal signalling only upon transfection of cells with 20 ng of the  $G\alpha_s$  subunit; all other  $G\alpha$  subunits could be dosed at 20 ng without an increase in basal signalling. We therefore titrated the dose of  $G\alpha_s$

DNA and found that 1 ng of DNA was the highest DNA dose that did not raise basal signalling compared to empty vector control. When CRE reporter activity was assessed, a strong, significant increase in activity was observed when the MBP-CTF form of ADGRF1 was co-transfected with 1 ng  $G\alpha_s$ , whereas a much smaller (but significant) increase was observed with 20 ng  $G\alpha_q$  and MBP-CTF (Fig. 1b, top). These data show that CRE reporter activity induced by the ADGRF1 CTF results primarily from  $G\alpha_s$  coupling.

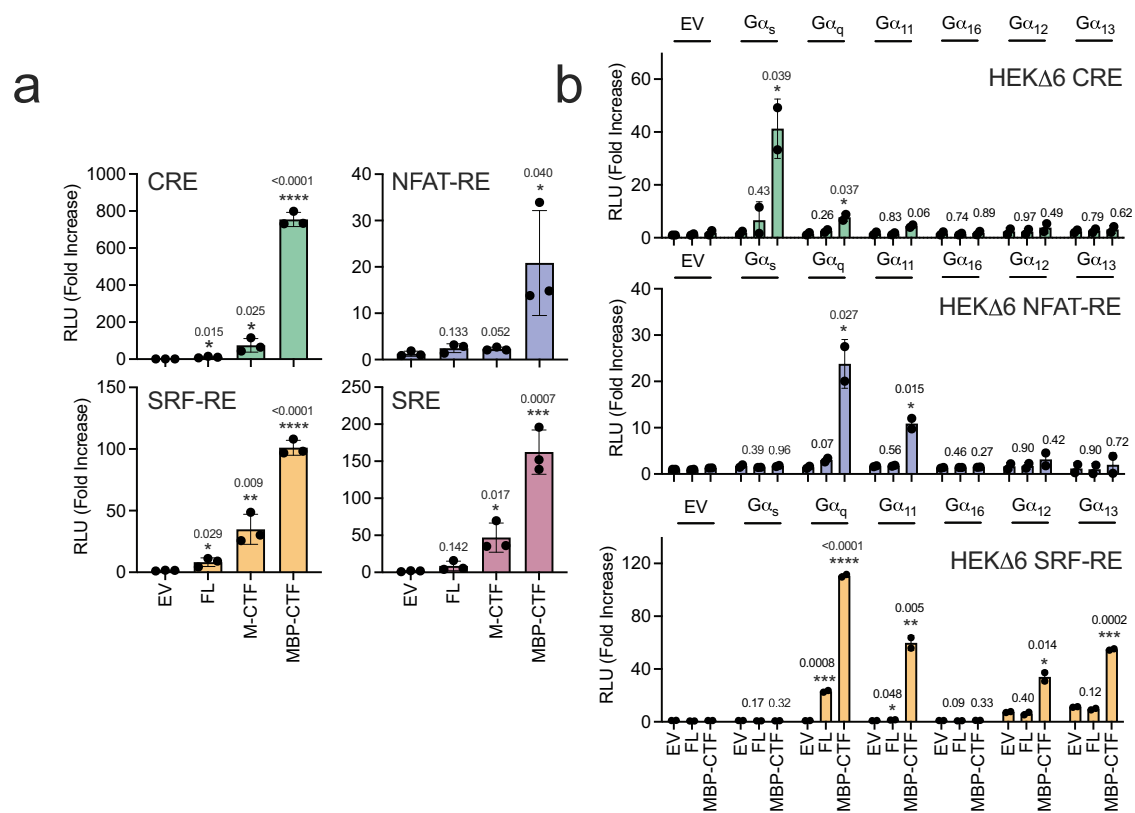
In NFAT-RE reporter assays all  $G\alpha$  subunits could be transfected at 20 ng DNA per well with no increase in basal signalling activity. When ADGRF1 MBP-CTF was co-transfected with either  $G\alpha_q$  or  $G\alpha_{11}$ , a robust, significant increase in signalling is observed, whereas no increase in basal signalling is seen upon co-transfection of the other  $G\alpha$  subunits (Fig. 1b, middle). Given that only co-transfection of  $G\alpha_q$  or  $G\alpha_{11}$  with MBP-CTF elicits a response, the NFAT-RE reporter shows fidelity to  $G\alpha_q$  signalling.

In SRF-RE reporter assays, a significant increase in basal signalling is observed when cells are transfected with 20 ng of plasmid encoding  $G\alpha_{12}$  or  $G\alpha_{13}$ , whereas plasmids encoding all other  $G\alpha$  subunits could be dosed at 20 ng without an increase in basal activity. Again, we titrated  $G\alpha_{12}$  and  $G\alpha_{13}$  DNA to identify the highest plasmid dose that did not result in an increase of basal signalling, which was 0.3 ng of DNA for each. When the optimized dose of each  $G\alpha$  subunit was co-transfected with MBP-CTF, significant increases in SRF-RE signalling were observed for  $G\alpha_q$ ,  $G\alpha_{11}$ ,  $G\alpha_{12}$ , and  $G\alpha_{13}$ , and a small but significant increase is also seen with ADGRF1 FL in the presence of  $G\alpha_q$  (Fig. 1b, bottom). SRF-RE therefore appears to report both on  $G\alpha_{q/11}$  and  $G\alpha_{12/13}$  signalling based on the responses seen upon complementation with these G proteins.

To evaluate ADGRF1 G protein signalling promiscuity further, we used the TRUPATH GPCR proximal BRET assay to interrogate G protein heterotrimer activation in cells<sup>25</sup>. In this assay, BRET from a  $G\alpha$ -Renilla-Luc8 fusion protein to  $G\gamma_2$ -GFP decreases when G protein turnover takes place and either shows minimal change or becomes elevated when the G protein pool remains in an inactive or complexed heterotrimeric state. We measured the BRET ratio in the TRUPATH assay for cells transfected with empty vector, full length ADGRF1, M-CTF, MBP-CTF, or MBP- $\Delta$ TA-CTF, from which the seven-residue sequence constituting the tethered agonist (TSFSILM) was deleted. We observed that G proteins representative of all major classes show a significant reduction in the BRET ratio for M-CTF and MBP-CTF relative to empty vector control (Fig. 2). We note that the BRET ratio decreases for all forms of ADGRF1 tested when  $G\alpha_{13}$  is the G protein partner, including the MBP- $\Delta$ TA-CTF protein that lacks the tethered agonist, suggesting some tethered agonist independent turnover of  $G\alpha_{13}$  is detectable for ADGRF1 in this assay. Nevertheless, the significance of BRET reduction when the tethered agonist is present is greater than in its absence, indicating a TA-dependent increase in response for  $G\alpha_{13}$  as well. Taken together with the transcription factor reporter assays, the TRUPATH results indicate that ADGRF1 has the capacity to transduce signals through  $G\alpha_s$ ,  $G\alpha_q$ ,  $G\alpha_i$  or  $G\alpha_{12/13}$ , and thus appears to be the first of the class B1 or B2 GPCRs shown to have the capability to signal through all G protein classes<sup>26</sup>.

### Protein engineering of ADGRF1 and Cryo-EM structure determination

To obtain an active state ADGRF1 G protein complex amenable for structural and functional investigation, we probed complex assembly with a suite of mini $G\alpha$  proteins<sup>27,28</sup>, using a NanoBiT recruitment assay<sup>29,30</sup>. Unlike HiBiT, which has a strong intrinsic affinity for LgBiT (~700 pM) and therefore forces the tethering of two proteins<sup>31</sup>, The SmBiT sequence used here has a very weak intrinsic affinity for LgBiT (~190  $\mu$ M), and requires a biologically relevant binding interface between a GPCR and  $G\alpha$  protein for formation of complexes<sup>31</sup>. We found that LgBiT-mini $G\alpha_s$ , mini $G\alpha_{s/q}$  and mini $G\alpha_o$  were all recruited to C-terminal SmBiT-tagged ADGRF1-CTF as judged by increased



**Fig. 1 | Analysis of G protein coupling of ADGRF1 using CRE, NFAT-RE and SRF-RE reporters in parental and HEKΔ6 knockout cells. a** Transcription factor reporter assay responses for ADGRF1 FL, M-CTF or MBP-CTF relative to empty vector (EV) for CRE (top left), NFAT-RE (top right), SRF-RE (bottom left) and SRE (bottom right). Cells were transfected with 30 ng receptor DNA/well for SRE and SRF-RE, and 10 ng DNA/well for CRE and NFAT-RE. Data are normalized to empty vector as relative luminescence units (RLU). Data are presented as mean  $\pm$  s.d. of three biological replicates. Each point represents the mean value of three technical replicates. Student's two-tailed T-test ( $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ , and  $****P < 0.0001$ ) was used to compare the means between biological samples to each reporter's empty vector control. **b** G $\alpha$  complementation assay. Activity of the CRE,

NFAT-RE and SRF-RE reporters were measured in HEKΔ6 cells transfected with a combination of EV, ADGRF1 FL or MBP-CTF, and either EV or one of several different G $\alpha$  subunits. All G $\alpha$  were transiently transfected at 20 ng DNA per well, except for G $\alpha_s$  in the CRE assay which was given at 1 ng, and G $\alpha_{12/13}$  which were given at 0.3 ng in the SRF-RE assay. Values are reported as the ratio of the raw luminescence values of firefly and renilla luciferases normalized to the matched G $\alpha$  empty vector (receptor) control. Data are presented as mean  $\pm$  s.d. of two biological replicates. Each point represents the mean value of three technical replicates. In all panels, Student's two-tailed T-test ( $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ , and  $****P < 0.0001$ ) was used to compare the means between biological samples to each reporter's empty vector control.

luminescence upon SmBiT-LgBiT complementation, whereas matched CΔ5 control forms of these mini G proteins and LgBiT-miniG $\alpha_{12}$  were not recruited (Supplementary Fig. 2a). To assess tractibility for structure determination, we co-expressed ADGRF1 with these miniG constructs in Expi293T suspension cells (ThermoFisher, A14527), and also observed a similar pattern of G $\alpha$  recruitment to ADGRF1 (Supplementary Fig. 2b, left). When we solubilised cells with mild detergent supplemented with cholesterol in the presence of apyrase, the only complex that retained luminescence under these lysis conditions was the combination of SmBiT-tagged ADGRF1-CTF with LgBiT-miniG $\alpha_{s/q}$  (Supplementary Fig. 2b, right). In addition, only miniG $\alpha_{s/q}$  specifically purified with ADGRF1-CTF using anti-FLAG M2 magnetic beads (Supplementary Fig. 2c), indicating that ADGRF1-G $\alpha$  complex stability is greater for miniG $\alpha_{s/q}$  than for miniG $\alpha_s$  or miniG $\alpha_q$ . Furthermore, miniG $\alpha_{s/q}$  co-expression with ADGRF1-CTF appears to increase the amount of receptor purified, suggesting that the miniG $\alpha_{s/q}$  may improve receptor expression by facilitating the folding of the receptor, increasing its stability, or by blocking toxic signalling through endogenous signalling pathways.

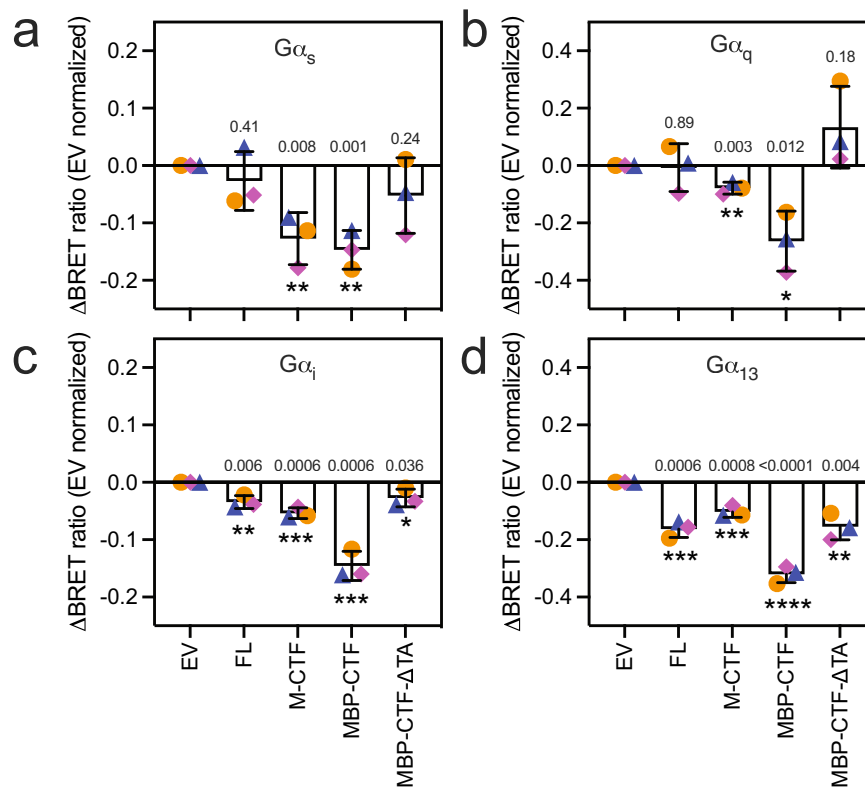
### Structure of the ADGRF1-miniG $\alpha_{s/q}$ heterotrimeric G protein complex with the Nb35 nanobody

To elucidate the basis for selective recruitment of G $\alpha_{s/q}$ , we complexed ADGRF1 CTF with a miniG $\alpha_{s/q}$  heterotrimeric G protein complex and

the Nb35 nanobody, purified the complex by size-exclusion chromatography (Fig. 3a), and determined its structure by cryo-EM to 3.4 Å resolution (Supplementary Figs. 3, 4). The cryo-EM map was of sufficient resolution to permit the building of an atomic model for the receptor, G protein complex, and nanobody (Fig. 3b, Supplementary Fig. 5). The cryo-EM map in the orthosteric pocket on the extracellular face of the receptor permitted unambiguous modelling of the TA (Supplementary Fig. 5), which adopts a short alpha-helical motif with a loop capping the TA en route to TML, akin to other resolved TA structures of aGPCRs<sup>12–15</sup>. The polypeptide backbone of the 7TM domain of ADGRF1 bound to miniG $\alpha_{s/q}$  has an RMSD of 0.6 Å when aligned to structures of the 7TM domain of ADGRF1 bound to either miniG $\alpha_s$  (PDB 7WU3) or miniG $\alpha_q$  (PDB 7WU4)<sup>15</sup> highlighting that the overall organization of these structures is highly similar despite their different G protein partners. An LPC lipid is present at the intracellular convergence of TM helices II, III, and IV, as also seen in ADGRF1 structures bound to miniG $\alpha_s$  or miniG $\alpha_q$ <sup>15</sup>. The acyl chain is similarly positioned in our structure, whereas the headgroup has the phosphate embedded deeper in the detergent micelle (Fig. 3b, Supplementary Fig. 5).

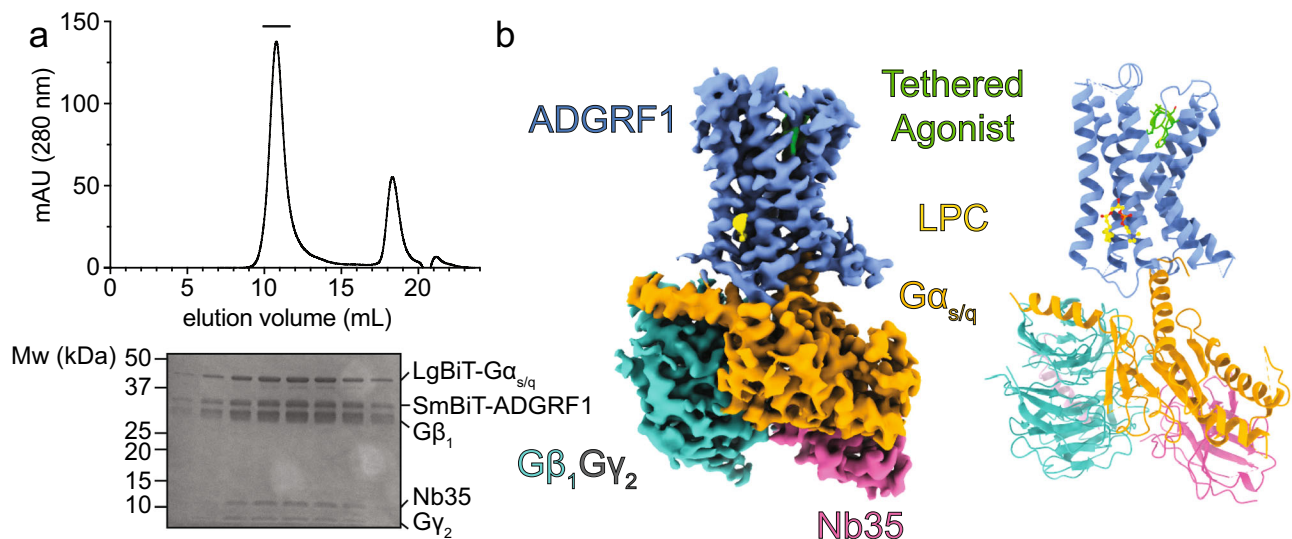
### Structural differences between the ADGRF1 miniG $\alpha_{s/q}$ and ADGRF1 miniG $\alpha_s$ complexes

MiniG $\alpha_{s/q}$  was engineered as a chimera with six key G $\alpha_q$  residues on the C-terminal  $\alpha 5$  helix that distinguish G $\alpha_q$  coupling over G $\alpha_s$ , even



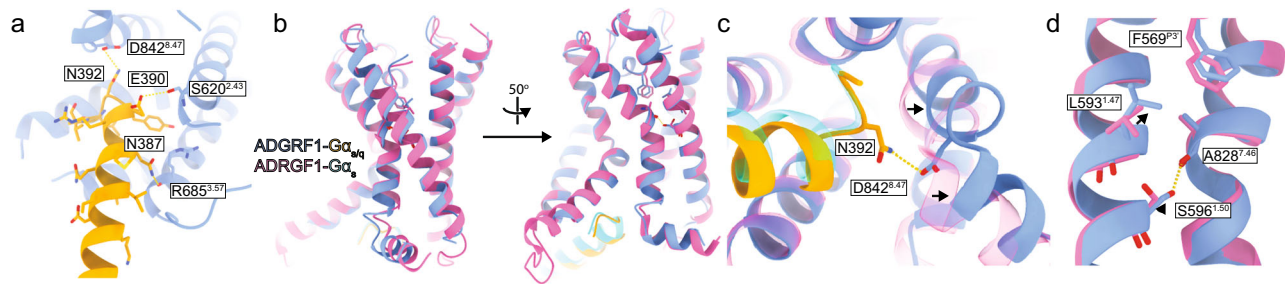
**Fig. 2 | Evaluating G protein coupling of TA-engaged ADGRF1 for all G protein classes using the TRUPATH BRET assay.** G protein activation assays, referenced to empty vector (EV), were performed for ADGRF1 FL, M-CTF, MBP-CTF, or MBP-ΔTA-CTF using TRUPATH  $G\alpha_s$  (a),  $G\alpha_q$  (b),  $G\alpha_i$  (c), and  $G\alpha_{13}$  (d) systems. Cells were transfected with 20 ng DNA/well of each receptor variant, and 80 ng DNA/well of the TRUPATH plasmid. BRET ratios are normalized to the empty vector control of

each receptor condition paired with each tested G protein (a–d). Data are presented as mean  $\pm$  s.d. of three biological replicates. Symbols and colours depict experiments from the same biological replicate (orange circles, blue triangles, magenta diamonds). Student's two-tailed T-test was used to compare the means between biological samples to each reporter's empty vector control (\* $P$  < 0.05, \*\* $P$  < 0.01, \*\*\* $P$  < 0.001, and \*\*\*\* $P$  < 0.0001).



**Fig. 3 | ADGRF1-mini $G\alpha_{s/q}$ -Nb35 complex purification and cryo-EM structure.** a ADGRF1 mini $G\alpha_{s/q}$ - $\beta_1\gamma_2$ -Nb35 purification. Top: elution profile of the complex after size-exclusion chromatography as a final purification step. Bottom: peak fractions (bar above chromatogram) analyzed by SDS-PAGE. b Structure of the ADGRF1- $G\alpha_{s/q}$ - $\beta_1\gamma_2$ -Nb35 complex. Left: cryo-EM map of the ADGRF1 complex. ADGRF1 (light blue), the TA (green),  $G\alpha_{s/q}$  (orange),  $\beta_1$  (teal),  $\gamma_2$  (grey), and Nb35

(pink) are coloured as indicated. The density threshold was set at five standard deviations from the mean. Right: cartoon rendering of the structure, with residue side chains of the tethered agonist rendered as sticks. The density threshold was set at three standard deviations from the mean for the TA. The LPC lipid (yellow) molecule is shown with a ball-and-stick model.



**Fig. 4 | The ADGRF1-miniG $\alpha_{s/q}$  interface and comparison to ADGRF1-miniG $\alpha_s$ .** **a** Close-up of the ADGRF1 (blue) interface with G $\alpha_{s/q}$  (orange) along the C $\alpha_5$  helix of G $\alpha_{s/q}$ . Hydrogen bonding interactions between the receptor and G $\alpha$  are indicated with yellow dashed lines. **b** Comparison of the ADGRF1 (blue)–G $\alpha_{s/q}$  (orange) complex with the ADGRF1 (pink)–miniG $\alpha_s$  (cyan) complex, aligned by least-squares superposition of the ADGRF1 subunits. **c** Zoomed in view focusing on D842<sup>8,47</sup> in the

turn connecting helices VII and VIII. The hydrogen bond between N392 and the carboxylate of D842<sup>8,47</sup> is shown with a yellow dotted line. **d** Zoomed in view focusing on residues L593<sup>1,47</sup> and S596<sup>1,50</sup>. Arrows indicate positional shifts of these residues in the miniG $\alpha_{s/q}$  structure relative to their positions in miniG $\alpha_s$ . The hydrogen bond between S596<sup>1,50</sup> and the backbone carbonyl of A828<sup>7,46</sup> is shown with a yellow dotted line.

though the bulk of the chimera is derived from miniG $\alpha_s$ <sup>27</sup>. Because we observe a preference for miniG $\alpha_{s/q}$  recruitment to the CTF of ADGRF1, consistent with the previously observed preference for G $\alpha_q$  in GTP $\gamma$ S activity assays<sup>10</sup>, the preference for G $\alpha_q$  is driven largely by the differences in its C-terminal  $\alpha 5$  helix.

At the ADGRF1 G protein interface we observe a canonical pose of the C-terminal  $\alpha 5$  helix at the open cavity at the intracellular opening of TA activated ADGRF1 (Figs. 3b, 4a). Three hydrogen bond interactions appear to stabilise the ADGRF1–G $\alpha_{s/q}$  interface in our structure. The N387 amide side chain nitrogen of miniG $\alpha_{s/q}$  approaches within hydrogen bonding distance of the backbone carbonyl of ADGRF1 R685<sup>3,57</sup>, E390 carboxylate approaches within hydrogen bonding distance of the hydroxyl group of S620<sup>2,43</sup> and the amide side chain nitrogen of N392 approaches within hydrogen bonding distance of D842<sup>8,47</sup> side chain carboxyl group (Fig. 4a).

The analogous residues of miniG $\alpha_s$  to N387, E390, N392 in miniG $\alpha_{s/q}$  are H387, Q390 and E392, which do not form the same hydrogen bonding arrangement as in the G $\alpha_{s/q}$  complex. In the G $\alpha_s$  complex, H387 does not make any polar interactions, Q390 approaches within hydrogen bonding distance of the backbone amide nitrogen of S617<sup>2,40</sup> and the S620<sup>2,43</sup> hydroxyl group and E392 approaches within hydrogen bonding distance of the backbone amine of S843<sup>8,48</sup> (Supplementary Fig. 6a). Given the close proximity of E392 to D842<sup>8,47</sup>, it is likely that charge repulsion destabilises the miniG $\alpha_s$ -bound state by comparison to the G $\alpha_q$ -bound state. These interactions are associated with a shift in the positioning of the segment linking TM helix VII to helix VIII relative to the G $\alpha$  subunit, with the C $\alpha$  of D392 in the  $\alpha 5$  helix of miniG $\alpha_s$  situated 1.9 Å closer to the C $\alpha$  of ADGRF1 S843<sup>8,48</sup> (Fig. 4c). The arrangement seen in the miniG $\alpha_s$  complex is also present in the miniG $\alpha_i$  structure. Superposition of helices I and VII of miniG $\alpha_{s/q}$  on miniG $\alpha_s$  shows a 3.2 Å displacement of the L384 C $\alpha$  miniG $\alpha_{s/q}$  relative to that of N384 of miniG $\alpha_s$  (Supplementary Fig. 6b).

At the orthosteric site, we also observe a 1.2 Å displacement of L593<sup>1,47</sup> towards F569<sup>P3'</sup> in the ADGRF1 complex with miniG $\alpha_{s/q}$ , compared with the miniG $\alpha_s$  and miniG $\alpha_i$  complexes (Fig. 4d). This inward movement leads to a rearrangement of the TM I alpha helix, allowing the conserved S596<sup>1,50</sup> hydroxyl to approach within hydrogen bond distance of the backbone carbonyl of A828<sup>7,48</sup> at the kink in the TM VII helix. This interaction of S596<sup>1,50</sup> with TM helix VII is not observed in either the miniG $\alpha_s$  or miniG $\alpha_i$  structures.

A structure of ADGRF1 bound to miniG $\alpha_{s/q}$  was also recently solved using the HiBiT tethering approach<sup>16</sup>. When the 7TM domain of our structure is aligned with the ADGRF1–miniG $\alpha_{s/q}$  tethered structure (PDB 7WXU), the RMSD over all atoms in the 7TM domain is 0.8 Å. However, the position of I839<sup>7,58</sup> is subtly different in the tethered structure, which affects the relative positions of TM helix VII and helix VIII to TM helix I (Supplementary Fig. 7). As a result the miniG $\alpha_{s/q}$  tethered structure

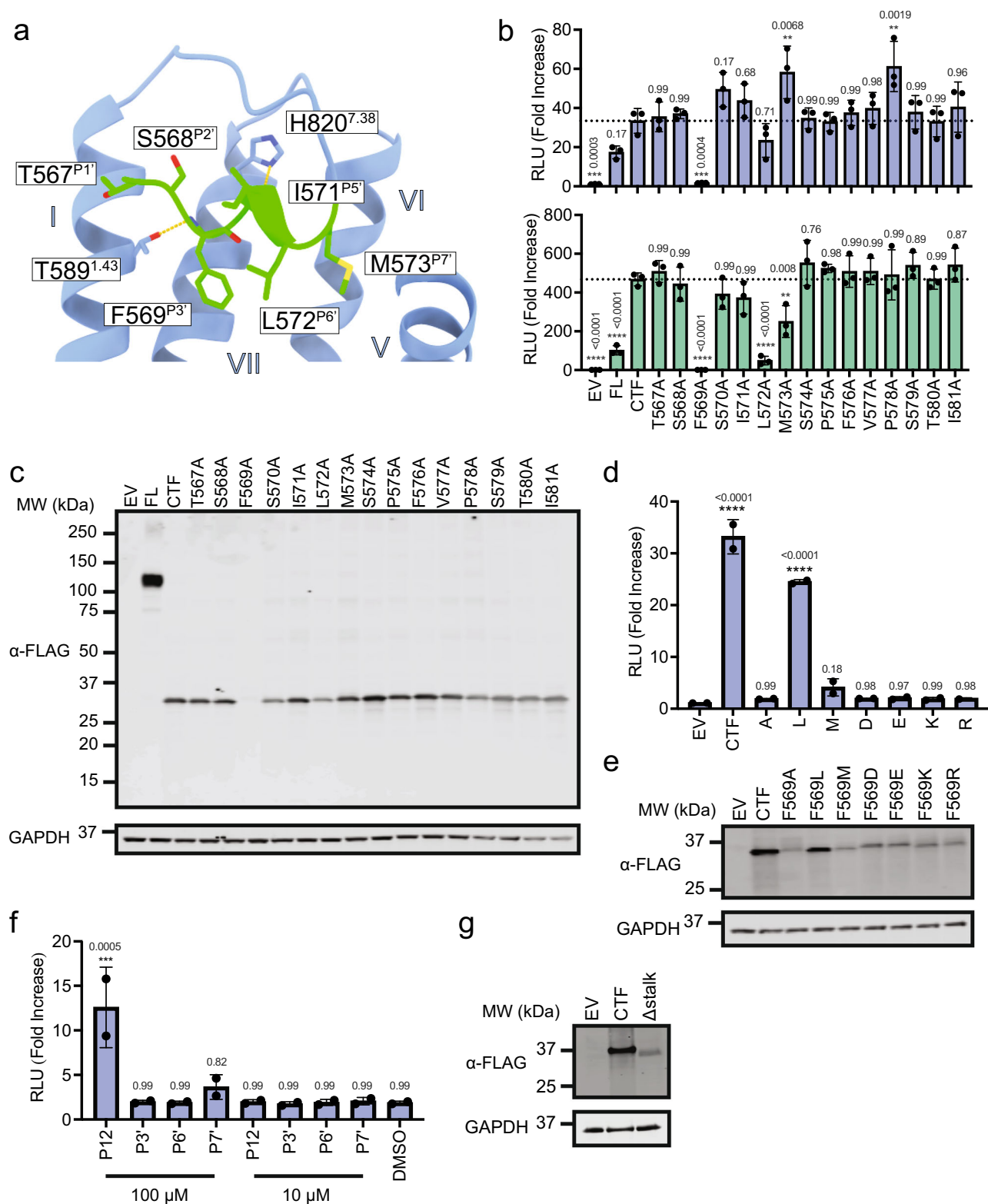
more closely resembles that of the untethered miniG $\alpha_s$  structure (PDB 7WU3)<sup>15</sup>. The side-chain interactions of N387, E390 and N392 of miniG $\alpha_{s/q}$  with R685<sup>3,57</sup>, S620<sup>2,43</sup> and D842<sup>8,47</sup>, respectively (described above) are also observed in the tethered structure<sup>16</sup>. The tethered structure additionally reports the presence of interactions between the sidechains of miniG $\alpha_{s/q}$  residues D381 and Q385 with the guanidinium group of R771<sup>5,63</sup>, as well as an interaction between V394 at the C-terminus with the side chain of K791<sup>6,40</sup>. In our structure, there is not appreciable density for R771<sup>5,63</sup>, suggesting flexibility in this region. Similarly, the side chain position of K791<sup>6,40</sup> is not well defined in our electron density map, suggesting that the position of this side chain is also flexible. Further analysis of coupling contacts to miniG $\alpha_i$ , miniG $\alpha_{12}$  and miniG $\alpha_{13}$  responsible for ADGRF1 G protein coupling promiscuity and comparisons among these structures, determined using the tethering approach, are available elsewhere<sup>16</sup>.

#### Distinguishing preferential G $\alpha_q$ over G $\alpha_s$ tethered agonism at the orthosteric site

Previous work using truncations and alanine substitutions at the P3' and P7' positions identified the functional importance of hydrophobic residues in TA-mediated signalling<sup>10</sup>. Because ADGRF1 shows preference for G $\alpha_q$  over other G $\alpha$  partners<sup>10</sup>, we investigated here whether G $\alpha_q$  and G $\alpha_s$  signalling exhibited differential sensitivity to alanine mutants of the receptor stalk residues (Fig. 5a), using NFAT-RE and CRE reporter gene assays for G $\alpha_q$  and G $\alpha_s$  coupling, respectively (Fig. 5b). Positions 574–581, corresponding to P8'–P15', were largely unaffected by mutation to alanine in NFAT-RE and CRE assays, indicating that this region is not directly important for TA dependent agonism, primarily functioning as a linker connecting the TA to TM1 of ADGRF1.

As in a previous analysis of the F569<sup>P3'</sup>A and M573<sup>P7'</sup>A mutations<sup>10</sup>, alanine substitutions at F569<sup>P3'</sup> and M573<sup>P7'</sup> of the TA affected TA-dependent signalling, as did an alanine substitution at L572<sup>P6'</sup>. The L572<sup>P6'</sup>A and M573<sup>P7'</sup>A mutations both reduced the response of the CRE reporter, whereas these substitutions either did not affect (L572<sup>P6'</sup>A) or modestly increased (M573<sup>P7'</sup>A) the response of the NFAT-RE reporter, arguing that G $\alpha_s$  activation, but not G $\alpha_q$  activation, selectively requires bulky hydrophobic residues at positions P6' and P7' for TA-dependent signalling in ADGRF1.

Strikingly, we found that the lack of signalling activity for F569<sup>P3'</sup>A in both NFAT-RE and CRE is accompanied by greatly decreased receptor abundance, as judged by the lack of a band on Western blot when probing for the N-terminal 3X-FLAG tag with an anti-Flag antibody (Fig. 5c). An antibody raised to a C-terminal portion (residues 831–880) of ADGRF1 showed a similar reduction in Western blot signal for the F569<sup>P3'</sup>A mutant (Supplementary Fig. 8), confirming that the loss of anti-FLAG reactivity was not due to proteolysis of the N-terminal tag. We propose that failure of the mutated TA to engage the 7TM



leads to the exposure of the TA hydrophobic sequence, creating a neoepitope that stimulates heat-shock protein binding in the secretory pathway and subsequent degradation (i.e., the inactivity of the TA is the cause of the reduced protein amount, and not a consequence of it).

We explored the tolerance of F569<sup>P3</sup> more extensively by testing a suite of amino acid substitutions, including a conservative leucine substitution, a methionine substitution, and all of the charged amino

acids (Fig. 5d, e). Among these substitutions, only the F569<sup>P3</sup>L mutation produced a detectable signal (Fig. 5d), and only this substitution resulted in similar amounts of expressed protein as the wild-type CTF as judged by Western blot (Fig. 5e). The ability of F569<sup>P3</sup>L to support TA-dependent signalling indicates that hydrophobic packing of this residue with F641<sup>2-61</sup> is sufficient for signalling activity, and that an aromatic  $\pi$ -stacking interaction is not required.

**Fig. 5 | Tethered agonism signalling determinants through NFAT-RE 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. Sidechains 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-RE 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. **c**  $\alpha$ -FLAG Western blot comparing steady-state amounts of wild-type and mutated ADGRF1 CTF proteins tested in panel b. An  $\alpha$ -GAPDH blot was used as a loading control. Blot shown is representative of three independent biological repeats. **d** NFAT-RE signalling data and **(e)**  $\alpha$ -FLAG Western blot of wild-type and P3' mutated forms (L, M, R, K, D and E) of the ADGRF1 CTF. An  $\alpha$ -GAPDH blot was used as a loading control. Blot shown is

representative of two independent biological repeats. **f** NFAT-RE reporter gene activity of the ADGRF1 CTF stalk protein stimulated by *trans* addition of a 12-residue TA peptide (P12). The response to the wild-type peptide or to peptides with alanine mutations at the P3', P6' or P7' positions are shown at doses of 10 or 100  $\mu$ M. **g**  $\alpha$ -FLAG Western blot comparing steady-state amounts of wild-type and  $\Delta$ stalk ADGRF1 CTF proteins tested in **f**. An  $\alpha$ -GAPDH blot was used as a loading control. Blot shown is representative of two independent biological repeats. For panel **(b)**, data are presented as mean  $\pm$  s.d. of three biological replicates. Each point represents the mean value of three technical replicates. For panels **d** and **f** data are presented as mean  $\pm$  s.d. of two biological replicates. 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 in **b** and **d**, and to DMSO in **f** (\*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ ).

Because of the variation in protein abundance for most of the F569<sup>P3'</sup> mutants, we also assessed the role of the P3', P6' and P7' hydrophobic buried residues in activation using *trans* addition of TA peptide sequences, analogous to previous studies<sup>10</sup>. For these assays, we used a form of the 7TM domain lacking the TA, which we refer to as stalkless, or  $\Delta$ stalk. We incubated cells expressing the  $\Delta$ stalk form of ADGRF1 with a 12mer TA mimicking peptide (P12) at 100  $\mu$ M or 10  $\mu$ M, and with analogous peptides containing alanine substitutions at P3', P6' or P7', and measured signalling activity using the NFAT-RE reporter. We found that the  $\Delta$ stalk construct signalled robustly through NFAT-RE with the P12 peptide at 100  $\mu$ M, but not 10  $\mu$ M (Fig. 5f, g). Together, our data show that a large bulky hydrophobic residue is required at P3' for ADGRF1 tethered agonism, that TA-dependent  $G\alpha_q$  signalling is more tolerant of substitutions at the positions at P6' or P7' of the TA than  $G\alpha_s$  signalling, and that signal activation by peptides added in *trans* is intolerant of alanine substitutions at any of the TA positions (P3', P6', or P7') even for  $G\alpha_q$  signalling.

To further investigate TA-dependent activation of  $G\alpha_s$  and  $G\alpha_q$  by ADGRF1, we performed extensive alanine mutagenesis of the TA binding site in the 7TM domain and measured the signalling response of the NFAT-RE and CRE reporters. The signalling activity of the mutated proteins clustered into three distinct categories: mutations that did not affect the CRE or NFAT-RE response, those that were detrimental to both CRE and NFAT-RE, and those that were detrimental to CRE only (Fig. 6a, see Supplementary Fig. 9 for analysis of protein abundance by Western blot).

Residues where alanine substitutions reduced signalling in both the NFAT-RE and CRE reporters include conserved residues W734<sup>45,51</sup> and W737 of ECL2, and several residues that surround F569<sup>P3'</sup> in the 7TM core, including conserved 'switch' residue W804<sup>6,53</sup>, neighbouring residues F823<sup>7,41</sup> and N827<sup>7,45</sup>, and T589A<sup>1,43</sup>, L593A<sup>1,47</sup>, A828L<sup>7,46</sup> and D638A<sup>2,61</sup> (Fig. 6b–d). These sets of mutations outline the importance of ECL2, the binding site arrangement around the 'switch' residue, and the conserved F569<sup>P3'</sup> position for general ADGRF1 tethered agonism. We also found that positions V812<sup>6,61</sup> and S814<sup>6,63</sup> lowered signalling in both NFAT-RE and CRE assays, but this was likely due to a profound decrease in protein expression (Supplementary Fig. 9).

Strikingly, mutations that decreased CRE signalling but not NFAT-RE signalling were primarily in two separate clusters that interact with positions other than F569<sup>P3'</sup> of the TA sequence. The first cluster contains residues within van der Waals contact of L572<sup>P6'</sup> and M573<sup>P7'</sup> at positions Y668<sup>3,40</sup>, F672<sup>3,44</sup> and F747<sup>5,39</sup> on TM helices III and V, with the F672<sup>3,44</sup>A substitution enhancing the NFAT-RE response (Fig. 6b–d). The second cluster consists of T810<sup>7,28</sup> and H820<sup>7,38</sup>; T810<sup>7,28</sup> interacts with M573<sup>P7'</sup> and S574<sup>P8'</sup>, and H820<sup>7,38</sup> approaches within H-bonding distance of S570<sup>P4</sup>. An additional site at which alanine substitution enhanced NFAT-RE signalling while reducing CRE signalling was F641<sup>2,64</sup> on TM helix II. These results show that CRE signalling is more sensitive than NFAT-RE signalling upon disruption of the pocket

proximal to P6' and P7', as well as upon disruption of interactions with S570<sup>P4</sup> and S574<sup>P8'</sup>.

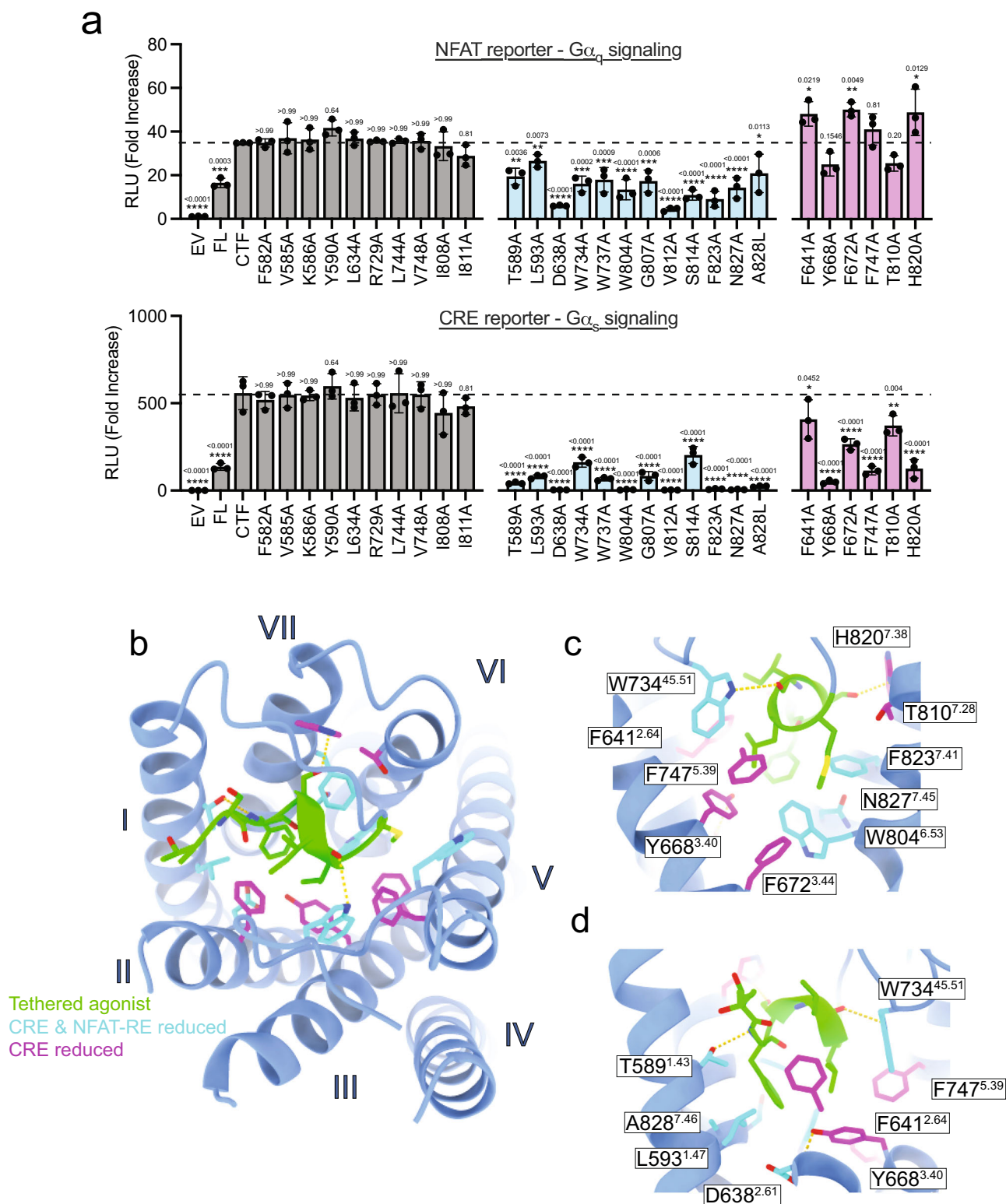
## Discussion

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

Here, we combine cryo-EM structural studies with mutational analysis and cell-based signalling assays to elucidate structure-function relationships in ADGRF1 and deduce its G protein coupling profile preferences. Our work confirms that ADGRF1 prefers coupling to  $G\alpha_q$ , as reported previously<sup>10</sup>, provides a structure-based rationale for this coupling preference, and, most strikingly, identifies a cluster of residues in the TA and in the binding pocket that selectively suppress  $G\alpha_s$  coupling upon mutation while retaining  $G\alpha_q$  coupling (Fig. 7).

While the structure of the ADGRF1-mini $G\alpha_{s/q}$  complex shows that the overall architecture of the complex and the pose of the TA in the orthosteric site are similar to recently observed structures of ADGRF1 when complexed to mini $G\alpha_s$ , mini $G\alpha_{12/13}$  and mini $G\alpha_{i5}$ , there are subtle differences in the complexes that explain why  $G\alpha_q$  coupling is more robust than  $G\alpha_s$  coupling. Through extensive analysis of ADGRF1 signalling, we find that it is competent to signal through each G protein class at both the G protein proximal, and deep transcription factor readout level. This is an important confirmation, as the TRUPATH assays determine that all G protein classes undergo ADGRF1 dependent heterotrimer disassociation, and the transcription factor reporter assays show that ADGRF1 dependent signalling is of sufficient strength to propagate at the transcriptional activation level. We also establish that ADGRF1 can signal through  $G\alpha_{12}$  or  $G\alpha_{13}$ , which to our knowledge makes ADGRF1 the first class B GPCR to show the capacity to signal through all major G protein classes<sup>26</sup>. Of note, these results are based on assays with overexpressed proteins. In future studies, it will be important to assess whether this capacity for ADGRF1 to couple to each G protein class leads to promiscuous signalling in specific physiological contexts.

Among the structural differences, tighter packing of L593<sup>1,47</sup> with the highly conserved phenylalanine (P3') of the TA appears to be critical for  $G\alpha_q$  preference, because the positional shift of L593<sup>1,47</sup> propagates its effects down helix I, introducing a hydrogen bond between S596<sup>1,50</sup> and A828<sup>7,48</sup> to stabilize the kink in TM helix VII. Our functional data suggests that these interactions contribute towards an increased stability of the ADGRF1- $G\alpha_q$  complex relative to the  $G\alpha_s$  and  $G\alpha_i$  complexes. Both the TA and the binding pocket of the 7TM domain show increased mutational lability for  $G\alpha_s$  signalling at several residues,



whereas  $G_{\alpha_q}$  signalling remains unaffected by those same mutations, showing that  $G_{\alpha_q}$  signalling is more tolerant to mutations than  $G_{\alpha_s}$ . Our data also confirm the critical importance of the F569<sup>P3</sup> residue for ADGRF1 G protein signalling, because a bulky hydrophobic residue at the P3' position is essential for expression and TA-dependent activation of the ADGRF1-CTF and mutations of residues that surround the F569<sup>P3</sup> pocket are detrimental in both NFAT-RE and CRE assays.

Unlike TA signalling of ADGRF1, which is tolerant of substitutions at positions other than F569<sup>P3</sup>, we found that *trans* addition of 12mer

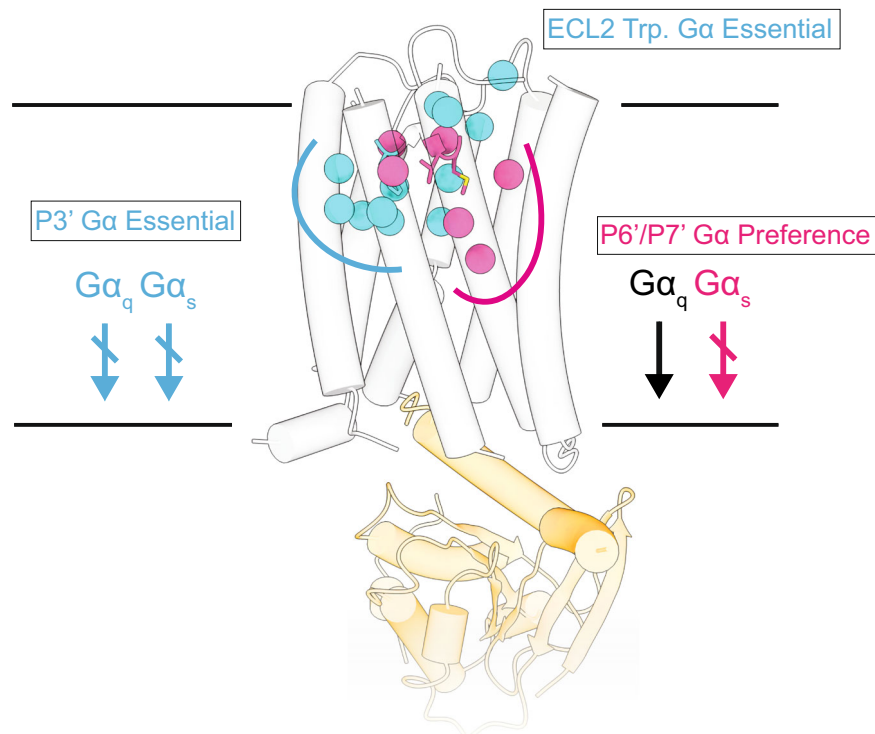
peptides to a stalkless form of ADGRF1 mimics the TA only when the wild-type sequence is used at the relatively high concentration of 100  $\mu$ M. Mutations at the P6' and P7' positions of the peptides, which are tolerated in TA signalling assays, failed to activate the stalkless ADGRF1 protein. This observation indicates that the development of peptidomimetic agonists for ADGRF1 will require extensive optimization if high potency is to be achieved in a soluble ligand.

Finally, our data elucidating a molecular basis for the preference of ADGRF1 for coupling to  $G_{\alpha_q}$  over  $G_{\alpha_s}$  identifies key sites in the 7TM



**Fig. 6 | Binding site signalling determinants through NFAT-RE and CRE.** a NFAT-RE 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-RE 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-RE and CRE signals are cyan, and mutations that significantly reduce only the CRE signal are magenta. Data are presented as mean  $\pm$  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  $****P < 0.0001$ ). b Top view of the TA in the orthosteric site. TA sidechain atoms making polar contacts (yellow dashes) are shown as sticks. Sidechains 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. Sidechains 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.



**Fig. 7 | Mutations of the TA and associated binding site cluster by G protein class preference.** ADGRF1 coupled to miniG $\alpha_{s/q}$  represented as tubed helices, coloured light grey and orange respectively. Residues that reduced signalling activity are shown, with TA residues rendered as sticks and with C $\alpha$  atoms of

residues in the TA binding site rendered as spheres. Residues that reduce both G $\alpha_s$  and G $\alpha_q$  signalling are coloured cyan, whereas residues that only reduce G $\alpha_s$  signalling are magenta.

that influence G protein coupling preferences. These findings should also inform the study of G protein coupling preferences and signalling biology for other aGPCRs.

## Methods

### MBP-CTF and signalling assay plasmids

DNA sequences encoding full length (20–910) and CTF (567–910) forms of ADGRF1 (Uniprot Q5T601) were subcloned into the pFuse-hlgG1-Fc2 vector behind an N-terminal IL2-signal sequence followed by a HA epitope tag, maltose binding protein and TEV protease site<sup>23</sup>. A FLAG tag was also inserted at the C-terminal end of the CTF before the stop codon. The CTF sequence was also subcloned to create the M-CTF protein with the N-terminal tags and mannose-binding protein omitted. Firefly reporter plasmids (CRE, NFAT-RE, SRE and SRF-RE) were assembled as described<sup>23</sup>. The pRL-TK plasmid (Promega) was used to express *Renilla* luciferase.

### Receptor and G protein plasmids

For recruitment assays, sequences encoding full-length (20–910) and CTF (567–910) ADGRF1 were subcloned into pcDNA3.1(+) with an N-terminal haemagglutinin signal sequence followed by a 3XFLAG

purification tag and a TEV protease cleavage site. For the CTF construct the TEV protease site P1' residue encoded a threonine so that the correct residue would occupy T567 in the CTF (e.g., ENLYFQ/TSFSLM...). The C-terminus has a "GGGGSGGGSSG" linker followed by a SmBiT tag "VTGYRLFEEIL"<sup>31</sup>. For proteins probed by Western blot the C-terminal SmBiT tag was replaced with a HA epitope tag. Mutagenesis was performed with a Q5 site directed mutagenesis kit (New England Biolabs) or using a modified QuikChange<sup>TM</sup> method<sup>32</sup>.

MiniG protein sequences (miniG $\alpha_s$ , miniG $\alpha_o$ , miniG $\alpha_{s/q}$  and miniG $\alpha_{12}$ ) with capacity to complex with G $\beta_1\gamma_2$  (e.g. 399 variants<sup>27</sup>) were subcloned into pcDNA3.1(+) with an N-terminal LgBiT fusion followed by a "GGGGSGGGSSGEF" linker that includes a translated EcoRI site. For structure determination and mutagenesis studies of ADGRF1, the constructs described above for recruitment assays were used.

### Protein expression, purification and complexation

ADGRF1 and miniG $\alpha_{s/q}$  were co-expressed in Expi293F cells. Cells were seeded at a density of  $3.3 \times 10^6$  cells/mL in 900 mL of Expi293 expression media. To transfect cells, equal amounts of receptor (250  $\mu$ g) and miniG (250  $\mu$ g) DNA were resuspended in 50 mL

opti-MEM, FectoPro (Polyplus) transfection reagent (0.5 mL) was resuspended in 50 mL opti-MEM, the DNA and FectoPro solutions were mixed to give a final 1:1 DNA/FectoPro ratio, and the mixture incubated at room temperature for 10 min before adding to cells. Approximately 24 h later, filter-sterilised Valproic acid sodium salt (Sigma-Aldrich) was added to 5 mM, along with 10 mL of 45 % D-(+)-Glucose solution (Sigma-Aldrich). Cells were cultured for a further 24 h before they were harvested by centrifugation at 4000 *g* for 15 min. The pellet was flash frozen and then stored at  $-80^{\circ}\text{C}$  until purification.

Cells were thawed and resuspended in 250 mL of ice-cold hypotonic lysis buffer (10 mM HEPES, pH 7.5, containing Roche protease inhibitor tablets and 0.1 mU/mL apyrase to prevent separation of the GPCR-miniG complex). The pellet was stirred on ice for 30 min and lysed by dounce homogenization. The broken cells were pelleted by centrifugation at 50,000 *g* and resuspended in 225 mL solubilisation buffer (10 mM HEPES, pH 7.5 containing 1 mM  $\text{MgCl}_2$ , 2 mM  $\text{CaCl}_2$ , and 100 mM NaCl) with 0.1 mU/mL apyrase and 1:100,000 Benzamide hydrochloride (v/v). The resuspended material was dounce homogenized, and DDM-CHS (10:1 pre-mix, Anatrace) was then added to a final concentration of 1% (w/v). This solution was stirred for 75 min at  $4^{\circ}\text{C}$ , clarified by centrifugation at 50,000 *g* for 1 h, and passed through a glass microfibre filter. The filtrate was then loaded onto M2 anti-FLAG antibody affinity resin by gravity flow. The resin was washed with 30 column volumes of wash buffer (10 mM HEPES, pH 7.5 containing 1 mM  $\text{MgCl}_2$ , 2 mM  $\text{CaCl}_2$ , 100 mM NaCl, and 0.1 % (w/v) LMNG-CHS), and the protein was eluted by the application of wash buffer containing 100  $\mu\text{g}/\text{mL}$  of 3XFLAG peptide. The eluted protein was concentrated using a 100 kDa MWCO centrifugal concentrator (Amicon).

Nb35 and human  $\text{G}\beta_1\gamma_2$  were purified according to previously published protocols<sup>33</sup>. ADGRF1-mini $\text{G}\alpha_{\text{s/q}}$ , Nb35 and  $\text{G}\beta_1\gamma_2$  were mixed in a 1:1.15:1.15 ratio and incubated overnight at  $4^{\circ}\text{C}$  with rotation. The protein mixture was centrifuged at 21,000 *g* for 20 min at  $4^{\circ}\text{C}$ , and then spun through a Durapore® PVDF 0.1  $\mu\text{m}$  column (Millipore). The protein mixture was injected onto a Superdex S200 10/300 GL column equilibrated in SEC buffer (10 mM HEPES pH 7.5 containing 1 mM  $\text{MgCl}_2$ , 2 mM  $\text{CaCl}_2$ , 100 mM NaCl, and 0.005 % (w/v) LMNG-CHS). The peak fraction containing the complex was collected and concentrated using a 100 kDa MWCO centrifugal concentrator to  $-8\text{ mg}/\text{mL}$ .

### Cryo-EM sample preparation and image acquisition

The complex was vitrified on QUANTIFOIL® holey carbon grids (400-mesh, copper, R1.2/L3, Electron Microscopy Sciences) using a FEI Vitrobot Mark IV (FEI, Hillsboro). Grids were glow-discharged, and 3  $\mu\text{L}$  of sample loaded to the grid in a chamber at  $22^{\circ}\text{C}$  and 100 % humidity. Samples were applied at a force of 15 and blotted for 5–7 s before plunge-freezing in liquid ethane.

Data were collected using a FEI Titan Krios at 300 kV with a Gatan Quantum Image Filter with K3 Summit direct electron detection camera in counting mode with a total exposure dose of  $-54\text{ e}^-/\text{\AA}^2$ . Thermo Scientific Smart EPU software was used for acquisition. 50 frames per movie were collected at magnification of 105,000x, corresponding to 0.825  $\text{\AA}$  per pixel. Micrographs were collected at defocus values ranging from  $-0.5$  to  $-1.5\text{ }\mu\text{m}$ . The processing scheme is summarised in Supplementary Fig. 4; in brief, motion corrected and dose-weighted using the RELION motion correction implementation<sup>34</sup> and contrast transfer function parameters estimated by CTFFIND4<sup>35</sup>. Low quality micrographs were removed using micassess<sup>36</sup> prior to particle picking in topaz<sup>37</sup>.  $-4.9\text{ M}$  particles were initially identified, leading to  $-2.9\text{ M}$  particles after 2D classification in RELION<sup>38</sup>. Two independent 3D classifications in RELION were then performed with different initial references, with the best classes from each subsequently combined and duplicates removed to leave  $-1.2\text{ M}$  particles corresponding to ADGRF1. Following further 3D classification in RELION - 660,000 particles underwent particle polishing in RELION<sup>34</sup> and further 2D classification followed by masked 3D classification with

no alignment using a T value of 20 in RELION. 73,903 particles with strong transmembrane density were identified which resulted in a 3.6  $\text{\AA}$  reconstruction from CryoSPARC non-uniform refinement with contrast transfer function refinement<sup>39</sup>. Subsequent local non-uniform refinement with a mask excluding the micelle improved the reconstruction to 3.44  $\text{\AA}$ .

### Atomic modelling and model refinement

Model building was carried out in Coot using overlaid non-uniform refinement and local filtered maps generated in CryoSPARC<sup>40</sup>, as well as deepEMhancer<sup>41</sup> maps. The ADGRF1-mini $\text{G}\alpha_{\text{s}}$  coordinate file (PDB: 7WU3) was used as an initial model. The mini $\text{G}\alpha_{\text{s}}$  sequence was changed in Coot to generate mini $\text{G}\alpha_{\text{s/q}}$ <sup>27</sup>. The coordinates were then manually rebuilt using Coot, and refined using ISOLDE<sup>42</sup>, Phenix Real-Space Refine<sup>43</sup> and Servalcat<sup>44</sup>. The final models were evaluated using MolProbity. Statistics of the map reconstruction and model refinement are presented in Table 1. Structural biology applications used in this project (except CryoSPARC) were compiled and configured by SBGrid<sup>45</sup>. Molecular graphics and analyses performed with UCSF

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

<b>ADGRF1</b>	
EMD-29684	
PDB 8G2Y	
<b>Data collection and processing</b>	
Voltage (kV)	300
Electron exposure ( $\text{e}^-/\text{\AA}^2$ )	-53.5
Defocus range ( $\mu\text{m}$ )	0.5–1.5
Pixel size ( $\text{\AA}$ )	0.825
Symmetry imposed	C1
Initial particle images (no.)	4,960,830
Final particle images (no.)	73,903
Map Resolution ( $\text{\AA}$ )	3.44
FSC threshold	0.143
Map resolution range ( $\text{\AA}$ )	3.1–6.0
<b>Refinement</b>	
Initial model used	7wu3
Model resolution ( $\text{\AA}$ )	3.10
FSC threshold	0.143
Map-sharpening B-factor ( $\text{\AA}^2$ )	-111.2
<b>Model composition</b>	
Non-hydrogen atoms	6,676
Protein residues	891
B factors ( $\text{\AA}^2$ )	
Protein	74.10
Ligand	98.71
<b>R.M.S.D. deviations</b>	
Bond lengths ( $\text{\AA}$ )	0.005
Bond angles ( $^{\circ}$ )	1.125
<b>Validation</b>	
MolProbity score	1.55
Clashscore	4.67
Poor rotamers (%)	0.00
<b>Ramachandran plot</b>	
Favoured (%)	95.47
Allowed (%)	4.53
Disallowed (%)	0.00

ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases<sup>46,47</sup>.

### Cell culture

HEK293T (ATCC, CRL-3216) and HEKΔ6 (from Asuka Inoue, Tohoku University) cells were cultured in DMEM supplemented with 10% FBS and 1% penicillin-streptomycin at 37 °C in a 5% CO<sub>2</sub> humidified incubator.

### Western blot analysis

To assess target protein expression and a loading control concurrently, we used two imaging channels on a LI-COR Odyssey CLx Fluorescent Imaging System. Antibodies were validated by both external sources and with comparison to empty vector controls in all western blots. To probe protein expression via the 3XFLAG epitope, we incubated membranes with 1:10,000-fold dilution of anti-FLAG M2 mouse antibody (Sigma-Aldrich F3165) and a 1:10,000 anti-GAPDH rabbit antibody (Cell Signalling Technology 14C10) as loading control. To probe protein expression via a C-terminal region of ADGRF1 (831–880) we incubated the membrane with 1:1,000-fold dilution of an anti-ADGRF1 rabbit antibody (Sigma-Aldrich SAB4501161) and 1:1,000-fold dilution of an anti-β-tubulin mouse antibody (Cell Signalling technology D3U1W) as a loading control. To visualise 1:10,000-fold dilutions of Donkey anti-rabbit IRDye<sup>®</sup>680 and goat anti-mouse IRDye<sup>®</sup>800 fluorophore conjugated secondary antibodies (LI-COR) were incubated with membranes.

### Dual luciferase reporter assays

HEK293T cells were seeded in DMEM in clear-bottom white 96-well plates pre-coated with poly-D-lysine the day before transfection. At roughly 70 % cell confluency, the medium in each well was replaced with fresh DMEM. Cells were transfected using Lipofectamine<sup>™</sup> 2000 (Invitrogen) at a ratio of 3 μL per 1 μg DNA in Opti-MEM using the manufacturer's instructions. Each well was co-transfected with 50 ng firefly reporter plasmid for CRE and NFAT-RE along with 1 ng pRL-TK plasmid. For SRE and SRF-RE each well was co-transfected with 30 ng of firefly reporter plasmid and 0.6 ng pRL-TK plasmid. Cells were also co-transfected with the indicated amounts of receptor DNA. Cells were cultured in the transfection mixture for 6 h before the media was replaced with complete media. 24 h after transfection, the media was replaced with fresh DMEM, and the cells were incubated for another 8 h prior to assay readout. For SRF-RE and SRE read-outs the media was replaced with DMEM lacking FBS. Before measuring luminescence, cells were washed once with PBS, 20 μL of 1X Passive Lysis Buffer was added to each well, and the lysates were incubated with shaking for 15 min. Cell lysates were serially pipetted, and the assay plate was centrifuged at 400 g for 1 min. Dual-Glo luciferase measurements (Promega) were read on a GloMax luminometer. The ratio of Firefly: Renilla signal was calculated for each well and normalized to the mean of the EV-transfected controls to give Relative Luminescence Units (RLU). Each assay was performed in technical triplicate wells and then repeated in biological triplicate on different days.

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

For Gα rescue experiments in HEKΔ6 cells, the standard CRE, NFAT-RE, and SRF-RE assay formats described for wild-type cells were

used, with the exception that DNA encoding Gα subunits was also included in the transfection procedure. Per well, Gα DNA loads are as follows. For the CRE assay: 1 ng of Gα<sub>s</sub>; 20 ng of Gα<sub>q</sub>, Gα<sub>11</sub>, Gα<sub>16</sub>, Gα<sub>12</sub>, and Gα<sub>13</sub>. For the NFAT-RE assay: 20 ng of Gα<sub>s</sub>, Gα<sub>q</sub>, Gα<sub>11</sub>, and Gα<sub>16</sub>; 1 ng of Gα<sub>12</sub>, and Gα<sub>13</sub>. For the SRF-RE assay: 20 ng of Gα<sub>s</sub>, Gα<sub>q</sub>, Gα<sub>11</sub>, and Gα<sub>16</sub>; 0.3 ng of Gα<sub>12</sub>, and Gα<sub>13</sub>. Total DNA per well was balanced with pcDNA3.1. Each assay was performed in technical triplicate wells and then repeated in biological duplicate on different days.

### G protein NanoBiT recruitment assays

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

### TRUPATH G protein signalling assays

TRUPATH plasmids were a generous gift from Dr. Justin English, University of Utah. TRUPATH plasmids encode the same protein sequences reported in the original paper<sup>25</sup>, but genes encoding Gα-*Renilla*, Gβ<sub>1</sub> and GFP-γ<sub>2</sub> are encoded on a single plasmid. Sequences encoding the Gα-*Renilla* and Gβ<sub>1</sub> genes have a triple T2A sequence between them; the Gβ<sub>1</sub> gene and GFP-γ<sub>2</sub> are separated by an intervening IRES sequence.

For the TRUPATH assays, HEK293T cells were seeded in DMEM in clear-bottom white 96-well plates pre-coated with poly-D-lysine the day before transfection. Cells were transfected with 80 ng of TRUPATH plasmid and 20 ng of receptor plasmid per well. The culture media was changed 4–6 h after transfection, and BRET measurements were taken 48 h after transfection. To record the BRET signal, the media was removed from the wells, and cells were then washed twice with Hank's Balanced Salt Solution (HBSS) containing 10 mM HEPES, pH 7.5. Coelenterazine 400a (NanoLight Technology) was reconstituted to 1 mM in anhydrous ethanol and frozen at –80 °C. Immediately prior to BRET measurement, coelenterazine 400a was diluted to 6 μM working solution in HBSS and 80 μL was applied to the cells. BRET was measured using a GloMax luminometer using ET405/40x and ET510lp filters (Chroma Technology Corp), for donor (*Renilla*) and acceptor (GFP) readouts respectively.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

Source data are provided with this paper. The Cryo-EM data, including unprocessed and processed maps, generated in this study have been deposited in the Electron Microscopy Databank under the accession code [EMD-29684](https://doi.org/10.25541/EMD-29684). The modelled protein structure generated in this study has been deposited at the Protein Data Bank under the accession code [SC2Y](https://doi.org/10.25541/SC2Y). All other data are available from the corresponding author on request. Source data are provided with this paper.

### References

1. Bjarnadóttir, T. K. et al. The human and mouse repertoire of the adhesion family of G-protein-coupled receptors. *Genomics* **84**, 23–33 (2004).
2. Bondarev, A. D. et al. Opportunities and challenges for drug discovery in modulating Adhesion G protein-coupled receptor (GPCR) functions. *Expert Opin. Drug Discov.* **15**, 1291–1307 (2020).
3. Weston, M. D., Luijendijk, M. W. J., Humphrey, K. D., Möller, C. & Kimberling, W. J. Mutations in the VLGR1 gene implicate G-protein

- signaling in the pathogenesis of Usher syndrome type II. *Am. J. Hum. Genet.* **74**, 357–366 (2004).
4. Boyden, S. E. et al. Vibratory urticaria associated with a missense variant in ADGRE2. *N. Engl. J. Med.* **374**, 656–663 (2016).
  5. Vizurraga, A., Adhikari, R., Yeung, J., Yu, M. & Tall, G. G. Mechanisms of adhesion G protein-coupled receptor activation. *J. Biol. Chem.* **295**, 14065–14083 (2020).
  6. Bjarnadóttir, T. K. et al. Identification of novel splice variants of Adhesion G protein-coupled receptors. *Gene* **387**, 38–48 (2007).
  7. Krasnoperov, V. G. et al. Latrotoxin stimulates exocytosis by the interaction with a neuronal G-protein-coupled receptor. *Neuron* **18**, 925–937 (1997).
  8. Araç, D. et al. A novel evolutionarily conserved domain of cell-adhesion GPCRs mediates autoprolysis. *EMBO J.* **31**, 1364–1378 (2012).
  9. Liebscher, I., Schö, J., Monk, K. R. & Schö Neberg Correspondence, T. A tethered agonist within the ectodomain activates the adhesion G protein-coupled receptors GPR126 and GPR133. *Cell Rep.* **9**, 2018–2026 (2014).
  10. Stoveken, H. M., Hajduczuk, A. G., Xu, L. & Tall, G. G. Adhesion G protein-coupled receptors are activated by exposure of a cryptic tethered agonist. *Proc. Natl Acad. Sci. USA* **112**, 6194–6199 (2015).
  11. Ping, Y. Q. et al. Structures of the glucocorticoid-bound adhesion receptor GPR97–Go complex. *Nat.* **2020** 589:7843 **589**, 620–626 (2021).
  12. Xiao, P. et al. Tethered peptide activation mechanism of the adhesion GPCRs ADGRG2 and ADGRG4. *Nat.* **2022** 604:7907 **604**, 771–778 (2022).
  13. Ping, Y.-Q. et al. Structural basis for the tethered peptide activation of adhesion GPCRs. *Nature* **604**, 763–770 (2022).
  14. Barros-Álvarez, X. et al. The tethered peptide activation mechanism of adhesion GPCRs. *Nat.* **2022** 604:7907 **604**, 757–762 (2022).
  15. Qu, X. et al. Structural basis of tethered agonism of the adhesion GPCRs ADGRD1 and ADGRF1. *Nature* **2022** **604**, 1–8 (2022).
  16. Zhu, X. et al. Structural basis of adhesion GPCR GPR110 activation by stalk peptide and G-proteins coupling. *Nat. Commun.* **13**, 1–11 (2022).
  17. Fredriksson, R., Lagerström, M. C., Höglund, P. J. & Schiöth, H. B. Novel human G protein-coupled receptors with long N-terminals containing GPS domains and Ser/Thr-rich regions. *FEBS Lett.* **531**, 407–414 (2002).
  18. Lum, A. M. et al. Orphan receptor GPR110, an oncogene over-expressed in lung and prostate cancer. *BMC Cancer* **10**, 40 (2010).
  19. Lee, J. W. et al. Orphan GPR110 (ADGRF1) targeted by N-docosaheptaenylethanolamine in development of neurons and cognitive function. *Nat. Commun.* **7**, 13123 (2016).
  20. Kwon, H. et al. Ligand-induced GPR110 activation facilitates axon growth after injury. *Int. J. Mol. Sci.* **22**, 7 (2021).
  21. Stoveken, H. M., Larsen, S. D., Smrcka, A. V. & Tall, G. G. Gedunin and khivorin-derivatives are small-molecule partial agonists for adhesion G protein-coupled receptors GPR56/ADGRG1 and GPR114/ADGRG5. *Mol. Pharmacol.* **93**, 477–488 (2018).
  22. Demberg, L. M. et al. Activation of adhesion G protein-coupled receptors: Agonist specificity of Stachel sequence-derived peptides. *J. Biol. Chem.* **292**, 4383–4394 (2017).
  23. Dates, A. N. et al. A fusion protein platform for analyzing tethered agonism in the adhesion family of G protein-coupled receptors. *bioRxiv* <https://doi.org/10.1101/2022.07.14.500097> (2022).
  24. Grundmann, M. et al. Lack of beta-arrestin signaling in the absence of active G proteins. *Nat. Commun.* **9**, 1–16 (2018).
  25. DiBerto, J. F., Olsen, R. H. J. & Roth, B. L. TRUPATH: an open-source biosensor platform for interrogating the GPCR transducerome. *Methods Mol. Biol.* **2525**, 185–195 (2022).
  26. Hauser, A. S. et al. Common coupling map advances GPCR-G protein selectivity. *eLife* **11**, e74107 (2022).
  27. Nehmea, R. et al. Mini-G proteins: Novel tools for studying GPCRs in their active conformation. *PLOS ONE* **12**, e0175642 (2017).
  28. Carpenter, B. & Tate, C. G. Engineering a minimal G protein to facilitate crystallisation of G protein-coupled receptors in their active conformation. *Protein Eng., Des. Selection* **29**, 583–593 (2016).
  29. Hall, M. P. et al. Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chem. Biol.* **7**, 1848–1857 (2012).
  30. Wan, Q. et al. Mini G protein probes for active G protein-coupled receptors (GPCRs) in live cells. *J. Biol. Chem.* **293**, 7466–7473 (2018).
  31. Dixon, A. S. et al. NanoLuc complementation reporter optimized for accurate measurement of protein interactions in cells. *ACS Chem. Biol.* **11**, 400–408 (2016).
  32. Liu, H. & Naismith, J. H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.* **8**, 91 (2008).
  33. Rasmussen, S. G. F. et al. Crystal structure of the  $\beta_2$  adrenergic receptor–Gs protein complex. *Nature* **477**, 549–555 (2011).
  34. Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *Int. Union Crystallogr.* **6**, 5–17 (2019).
  35. Rohou, A. & Grigorieff, N. CTFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
  36. Li, Y., Cash, J. N., Tesmer, J. J. G. & Cianfrocco, M. A. High-throughput Cryo-EM enabled by user-free preprocessing routines. *Structure* **28**, 858–869.e3 (2020).
  37. Bepler, T. et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* **16**, 1153–1160 (2019).
  38. Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
  39. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).
  40. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
  41. Sanchez-Garcia, R. et al. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* **4**, 1–8 (2021).
  42. Croll, T. I. ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. Sect. D: Struct. Biol.* **74**, 519–530 (2018).
  43. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. D: Struct. Biol.* **74**, 531–544 (2018).
  44. Yamashita, K., Palmer, C. M., Burnley, T. & Murshudov, G. N. Cryo-EM single-particle structure refinement and map calculation using Servalcat. *Acta Crystallogr. Sect. D: Struct. Biol.* **77**, 1282–1291 (2021).
  45. Morin, A. et al. Collaboration gets the most out of software. *eLife* **2**, e01456 (2013).
  46. Pettersen, E. F. et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
  47. Goddard, T. D. et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

## Acknowledgements

We thank members of the Blacklow and Kruse laboratories for helpful discussions. Cryo-EM data were collected at the Harvard Cryo-EM Centre for Structural Biology at Harvard Medical School. We thank Sarah Sterling

and Megan Meyer for helpful comments and advice during grid screening, and Megan Meyer for data collection. This work was supported by R35 CA220340 (S.C.B.), and a grant from the Warren Alpert Foundation (S.C.B.). A.N.D. was also supported in part by a Fujifilm fellowship.

### Author contributions

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 constructs. M.M.B. performed mutagenesis of TA and binding site residues. M.M.B., A.N.D. and D.T.D.J. performed signalling assays. D.T.D.J. performed TRUPATH and NanoBiT recruitment assays. D.T.D.J. expressed, purified and formed the ADGRF1-MiniGas/q-Gβ1γ2-Nb35 complex. D.T.D.J. and C.H.L. vitrified ADGRF1 complex on grids. S.D.R. and D.T.D.J. processed cryo-EM data, and S.D.R. produced the final map. D.T.D.J. built the ADGRF1 complex model. D.T.D.J. and S.C.B. wrote the manuscript.

### Competing interests

S.C.B. is on the board of directors for the 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 IFM, Scorpion Therapeutics, Odyssey Therapeutics, and Ayala Pharmaceuticals for unrelated projects. The remaining authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-38083-7>.

**Correspondence** and requests for materials should be addressed to Daniel T. D. Jones or Stephen C. Blacklow.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023