Structural Insights into Orphan G Protein-Coupled Receptors: A Focus on GPR161 and Proton-Sensing GPCRs

by
Nicholas Hoppe

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biophysics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Aashish Manglik
4C1E8A184D2E493...

Aashish Manglik

Chair

DocuSigned by:
Yifan Cheng
DocuSigned by:

Yifan Cheng

DocuSigned by:
Jason Cyster
DocuSigned by:

Jason Cyster

DocuSigned by:
David Julius
3CE818B8BD6645C...

David Julius

Committee Members

Copyright 2023

by

Nicholas Hoppe

Dedication

Colin, to our childhood selves, to the scientists we became, and to your memory.

Acknowledgements

My scientific training is the result of immense support from my colleagues, friends, and family.

To my advisor, Aashish Manglik. I was drawn to you and your lab because of the idealistic excitement and curiosity you have for research and the attentive way in which you cultivate those feelings in others. Although your lab was just a year old, it was clear you established an environment for training scientists who are not just thoughtful experts, but also humble learners who embrace a growth mindset. One example stands out. On a January morning, I was drinking coffee in the hallway procrastinating my first tissue extraction for maybe the third week. You called out my apprehension, told me to stand up, and without hesitation, walked me through the procedure step-by-step over the next two days. In that moment, the confidence I gained in you and in myself, set me up for success over the remainder of my PhD. Your ability to tactfully traverse long-term project planning to hands-on experiments alongside your students is just one reason you are an incredible scientist. I am grateful to have studied under you.

To my thesis committee, Jason Cyster, David Julius, and Yifan Cheng. Your groundbreaking science inspired and motivated my research. Having your guidance along the journey has been such an honor.

To my colleagues turned friends, Bryan Faust, Ben Barsi-Rhyne, Julian Harris, Chase Webb, Simone Harrison, Christian Billesbølle, Ishan Deshpande, Jiahao Liang, Tia Tummino, and Elissa Fink. From working early mornings and late nights, sharing failures and successes, and escaping

through outdoor adventures and cold drinks, my fondest UCSF memories are from my time with each of you.

To my Westconsinites, Kyle Nishikawa, Jason Wan, Andrew DuPlissis, Joel Stibbe, and Bridget Stollfus. I am grateful to have shared time in both Madison and the Bay Area with each of you and for the support you provided through weekend visits, dinners, and games.

To my parents, Mark and Mary Hoppe. Thank you for your unwavering support and your earnest efforts to learn about my science. Earning a PhD requires persistence and constant learning. Thank you for instilling these values in me.

Lastly, to my wife, Delia. You are the single most important person to me, and this has been especially true through my PhD. Thank you for sharing the ups and downs of this journey with me. I am excited to see where our next one takes us. I love you.

Contributions

Chapter 1 includes a reprint of "GPR161 structure uncovers the redundant role of sterol-regulated ciliary cAMP signaling in the Hedgehog pathway" written by Hoppe N, Harrison S, Hwang SH, Chen Z, Karelina M, Deshpande I, Suomivuori CM, Palicharla VR, Berry SP, Tschaikner P, Regele D, Covey DF, Stefan E, Marks DS, Reiter J, Dror RO, Evers AS, Mukhopadhyay S, Manglik A. *bioRxiv* (2023): 2023-05. Hoppe N and Harrison S contributed equally to this work. At the time of dissertation submission, these results are under revision at *Nature Structural & Molecular Biology*.

Chapter 2 includes unpublished results from a manuscript in preparation of submission written by Hoppe N, Howard M, Huang XP, Coyote-Maestas W, and Manglik A. Hoppe N and Howard M contributed equally to this work.

The co-authors and Aashish Manglik directed and supervised the research described in this dissertation. Nicholas Hoppe provided the experimentation and manuscript preparation for much of the following dissertation, which is a substantive contribution comparable to other dissertations in biochemistry and biophysics.

This material is based on work supported by the National Institutes of Health Ruth L. Kirschstein Fellowship (N.H.) under grant no. F31HL164045. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Institutes of Health.

Structural Insights into Orphan G Protein-Coupled Receptors:

A Focus on GPR161 and Proton-Sensing GPCRs

Nicholas Hoppe

Abstract

This thesis presents a comprehensive investigation into the structure and function of four orphan G protein-coupled receptors (GPCRs), GPR161 and the proton-sensing GPCRs, GPR4, GPR65, and GPR68. Orphan GPCRs, which lack known endogenous ligands, represent a challenging yet promising frontier in human biology and health due to their potential as novel therapeutic targets.

The first part of my thesis presents the structure-based deorphanization of GPR161, an orphan receptor involved in the Hedgehog signaling pathway. In this study, I and others study the activation mechanism of GPR161 to show a dual mode of activation that depends on both a self-interaction for stability and an allosteric cholesterol. While these features explain the cAMP signaling properties of GPR161, they unexpectedly did not explain the activity of GPR161 in the Hedgehog pathway. Our findings highlight the nuanced biology behind GPR161, and potentially other orphans, that cannot be known until a ligand is identified.

The second part of the thesis describes the activation mechanism for the proton-sensing GPCRs, GPR4, GPR65, and GPR68. Despite their physiological importance, the mechanism by which these receptors sense protons and transduce this signal into a cellular response has remained elusive. Our work identifies a distributed network of residues across the extracellular half of these

receptors that contribute to activation, providing a new framework for understanding proton-sensing membrane proteins.

Collectively, these studies provide insights into the activation mechanisms of these GPCRs and pave the way for the development of novel modulators targeting these receptors. In addition to the specific scientific findings for these receptors, I hope these biochemical and biophysical techniques are applied to other orphan GPCRs so that their biology and therapeutic potential is realized.

Table of Contents

List of Figures

List of Tables

Introduction

G protein-coupled receptors (GPCRs) constitute the largest family of membrane proteins in the human genome, playing pivotal roles across human physiology. Over one hundred of these receptors, lack known endogenous ligands and are known as orphan GPCRs. Orphan GPCRs present a significant challenge and huge opportunity in the field of drug discovery. Given their roles in various physiological processes and their accessibility on the cell surface, GPCRs are among the most successful drug targets. Deorphanizing GPCRs can lead to the development of novel therapeutics for a variety of diseases and understanding the function of orphan GPCRs sheds light on fundamental biological processes. My research at the University of California – San Francisco has focused on interrogating orphan GPCRs with biophysical methods, and my thesis focuses on my structural characterization of four orphan GPCRs, specifically GPR161 and the proton-sensing GPCRs, GPR4, GPR65, and GPR68.

Since the discovery that the β2 adrenergic receptor and opsins share a seven-transmembrane topology, GPCRs without known ligands have been a focal point of intense scientific research due to their potential as therapeutic targets[1]. This simple but profound observation enabled the development of two effective approaches to identify endogenous ligands for orphan GPCRs, known as deorphanization. First is expression cloning where a library of DNA sequences is iteratively selected for activity to a specific ligand [2]. Here the target ligand is known, but the potential orphan receptor is not. Second is reverse pharmacology, which works in the opposite direction. In reverse pharmacology, the orphan receptor is known and ligands, either from a high-throughput library or from tissue extracts, are screened for activity[3]. These methods have led to

numerous and important receptor-ligand pairings, including glucagon-like peptide-1, ghrelin, thyrotropin, and orexin[4–6]. These successes underscore the potential of orphan GPCRs as gateways to uncharted biological pathways and therapeutic possibilities.

Despite their promise, many GPCRs remain orphan because the availability of ligands is foundational to studying GPCRs, and the historical deorphanization methods, which depend on cell-based activity assays, may be exhausted[7]. In my research, I sought to apply advances in GPCR biophysics, including cryo-electron microscopy (cryo-EM) and native mass spectrometry, to open new avenues into studying orphan GPCRs. This as a challenging endeavor that opened more questions and possibilities than it answered. What follows are two mechanistic findings from my cryo-EM research on orphan GPCRs, and I am excited to see what future discoveries come for Orphan GPCRs, both from the Manglik Lab and the broader GPCR community.

References

1. Dixon, R. A. *et al.* Cloning of the gene and cDNA for mammalian beta-adrenergic receptor and homology with rhodopsin. *Nature* **321**, 75–79 (1986).
2. Lodes, M. J., Dillon, D. C., Houghton, R. L. & Skeiky, Y. A. W. Expression cloning. *Methods Mol. Med.* **94**, 91–106 (2004).
3. Libert, F., Vassart, G. & Parmentier, M. Current developments in G-protein-coupled receptors. *Curr. Opin. Cell Biol.* **3**, 218–223 (1991).
4. Thorens, B. Expression cloning of the pancreatic beta cell receptor for the gluco-incretin hormone glucagon-like peptide 1. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 8641–8645 (1992).
5. Parmentier, M. *et al.* Molecular cloning of the thyrotropin receptor. *Science* **246**, 1620–1622 (1989).
6. Sakurai, T. *et al.* Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. *Cell* **92**, 573–585 (1998).
7. Laschet, C., Dupuis, N. & Hanson, J. The G protein-coupled receptors deorphanization landscape. *Biochem. Pharmacol.* **153**, 62–74 (2018).

Chapter 1

GPR161 structure uncovers the redundant role of sterol-regulated ciliary cAMP signaling in the Hedgehog pathway

Contributing Authors

Nicholas Hoppe[1,2]*, Simone Harrison[1,2]*, Sun-Hee Hwang[3]*, Ziwei Chen[4,5], Masha Karelina[6,7,8,9,10], Ishan Deshpande[1], Carl-Mikael Suomivuori[7,8,9,10], Vivek R. Palicharla[3], Samuel P. Berry[11], Philipp Tschaikner[12,13], Dominik Regele[12], Douglas F. Covey[4,5,14,15], Eduard Stefan[12,13], Debora S. Marks[11], Jeremy Reiter[16], Ron O. Dror[6,7,8,9,10], Alex S. Evers[4,5,15], Saikat Mukhopadhyay[3#], Aashish Manglik[1,17,18#]

1. Department of Pharmaceutical Chemistry, University of California, San Francisco, CA, USA
2. Biophysics Graduate Program, University of California, San Francisco, CA, USA
3. Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, TX, USA
4. Department of Anesthesiology, Washington University School of Medicine, St. Louis, MO 63110, USA
5. Taylor Institute for Innovative Psychiatric Research, St. Louis, MO 63110, USA.
6. Biophysics Program, Stanford University, Stanford, CA 94305, USA
7. Department of Computer Science, Stanford University, Stanford, CA, USA
8. Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, CA, USA

9. Department of Structural Biology, Stanford University School of Medicine, Stanford, CA, USA
10. Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA, USA
11. Department of Systems Biology, Blavatnik Institute, Harvard Medical School, Boston, MA, USA.
12. Institute of Molecular Biology and Center for Molecular Biosciences, University of Innsbruck, Innsbruck 6020, Austria
13. Institute of Biochemistry and Center for Molecular Biosciences, University of Innsbruck; Tyrolean Cancer Research Institute (TKFI), Innsbruck 6020, Austria
14. Department of Psychiatry, Washington University School of Medicine, St. Louis, MO 63110, USA
15. Department of Developmental Biology, Washington University School of Medicine, St. Louis, MO 63110, USA
16. Department of Biochemistry and Biophysics, Cardiovascular Research Institute, University of California, San Francisco, San Francisco, California 94158
17. Department of Anesthesia and Perioperative Care, University of California, San Francisco, CA, USA
18. Chan Zuckerberg Biohub, San Francisco, CA, USA

[*]These authors contributed equally

[#]Correspondance to: Saikat Mukhopadhyay (saikat.mukhopadhyay@utsouthwestern.edu) or Aashish Manglik (aashish.manglik@ucsf.edu)

1.1 Abstract

The orphan G protein-coupled receptor (GPCR) GPR161 is enriched in primary cilia, where it plays a central role in suppressing Hedgehog signaling[1]. GPR161 mutations lead to developmental defects and cancers[2,3,4]. The fundamental basis of how GPR161 is activated, including potential endogenous activators and pathway-relevant signal transducers, remains unclear. To elucidate GPR161 function, we determined a cryogenic-electron microscopy structure of active GPR161 bound to the heterotrimeric G protein complex G_s. This structure revealed an extracellular loop 2 that occupies the canonical GPCR orthosteric ligand pocket. Furthermore, we identify a sterol that binds to a conserved extrahelical site adjacent to transmembrane helices 6 and 7 and stabilizes a GPR161 conformation required for G_s coupling. Mutations that prevent sterol binding to GPR161 suppress cAMP pathway activation. Surprisingly, these mutants retain the ability to suppress GLI2 transcription factor accumulation in cilia, a key function of ciliary GPR161 in Hedgehog pathway suppression. By contrast, a protein kinase A-binding site in the GPR161 C-terminus is critical in suppressing GLI2 ciliary accumulation. Our work highlights how unique structural features of GPR161 interface with the Hedgehog pathway and sets a foundation to understand the broader role of GPR161 function in other signaling pathways.

1.2 Introduction

Orphan G protein-coupled receptors (GPCRs) coordinate diverse signaling pathways to control many aspects of human physiology[5]. The orphan GPCR GPR161 has been characterized as a unique example of a constitutively active receptor that is located within the primary cilium of cells, an organelle that protrudes from the cell surface and locally organizes signaling components[1]. In its best understood signaling role, GPR161 is a critical negative regulator of the Hedgehog

pathway[1]. Knockout of *Gpr161* in mice is embryonically lethal, and the embryos display severe limb, facial, and early nervous system defects indicative of hyperactive Hedgehog signaling[13–17]. *GPR161* mutations in humans lead to developmental defects such as spina bifida[2,11,12], pituitary stalk interruption syndrome[20], and cancers such as medulloblastoma[10,14]. Overexpression of GPR161 has been linked to triple-negative breast cancer[4]. Like many other orphan GPCRs, however, fundamental mechanisms of GPR161 function remain unknown[14], including what stimulus gives rise to GPR161 constitutive activity and how signaling activity downstream of GPR161 impinges on its biological function.

The primary function of GPR161 has been framed by its discovery as a Hedgehog pathway regulator[1]. Hedgehog signaling during vertebrate embryogenesis mediates multicellular development, including the proper formation of limbs, the face, and the nervous system[15]. In the presence of the Hedgehog signal, GLI2/3 transcriptional factors accumulate in the primary cilia and form activators (GLI-A)[16]. In the absence of the Hedgehog signal, GLI2/3 are constitutively phosphorylated by protein kinase A (PKA), which leads to proteolytic conversion of these proteins into Hedgehog pathway repressors (GLI-R). Because PKA is canonically activated by the GPCR second messenger cyclic adenosine monophosphate (cAMP), current models propose that elevated ciliary cAMP levels activate PKA to suppress the Hedgehog pathway[17]. Although many GPCRs localize to the primary cilium[18,19], several observations have placed GPR161 as a unique Hedgehog pathway regulator. Loss of GPR161 function in mice and in fish causes phenotypes consistent with inappropriate Hedgehog pathway activation[13–17,20]. GPR161 functions both inside cilia and in the peri-ciliary endosomal compartments in regulating these phenotypes[1,6]. Furthermore, GPR161 is constitutively active in model cell lines and drives elevated cAMP via activation of G_s[1,6,20–22,23,24].

Upon Hedgehog pathway activation, GPR161 exits cilia by internalizing to the recycling endocytic compartment[1,21]. Finally, the C-terminus of GPR161 binds specifically PKA type I regulatory subunits through an A-kinase anchoring protein domain (AKAP), a unique feature of GPR161 among hundreds of GPCRs[25]. GPR161 is therefore thought to repress Hedgehog signaling by constitutive coupling to G_s, which elevates cAMP levels to drive PKA activity (Fig. 1.1a).

Several fundamental aspects of GPR161 function remain unclear - in particular, the potential stimulus that drives GPR161 activity remains unknown. The interdependent roles of G_s coupling and PKA binding, and their relative importance in Hedgehog signaling is also poorly defined. Here we use a combination of cryogenic-electron microscopy (cryo-EM), signaling studies, molecular dynamics simulations, and biochemical assays to determine the molecular mechanism of GPR161 activation and Hedgehog pathway repression. Our studies reveal that GPR161-induced G_s signaling is driven by a novel sterol-binding site. However, this signaling activity does not repress GLI2 ciliary accumulation, a key role of ciliary GPR161 in Hedgehog pathway repression**.** By contrast, the AKAP domain in GPR161 is necessary for repressing GLI2 accumulation in cilia. Together, these findings provide an activation mechanism for GPR161 and support PKA as a central downstream ciliary regulator of the Hedgehog pathway.

1.3 Results

Cryo-EM structure of active GPR161 bound to G_s heterotrimer

GPR161 is one of the most constitutively active GPCRs tested in the β-arrestin PRESTO-Tango assay[22], which uses nonciliated HEK293 cells (Fig. 1.1b). We reasoned that purification of GPR161 from HEK293 cells may allow us to determine a structure of the active signaling state

and may reveal potential stimuli for GPR161. Preparations of GPR161 in the absence of a signal transducer were of poor quality, suggesting that GPR161 alone may be structurally dynamic or otherwise unstable. To stabilize GPR161 in the active state and simultaneously increase the likelihood that the receptor would co-purify with an activating stimulus[26], we C-terminally fused the receptor to a minimal version of the $G\alpha_s$ protein. This minimized "miniG_s" construct retains the receptor-interacting GTPase domain of the $G\alpha_s$ subunit but is engineered to interact with a GPCR in the absence of a guanine nucleotide[27]. We purified GPR161-miniG_s to homogeneity, further complexed it with other heterotrimeric G protein subunits $G\beta_1$ and $G\gamma_2$ as well as nanobody 35 (Nb35) to stabilize the interaction between $G\alpha_s$ and $G\beta_1\gamma_2$ (Fig. 1.6)[35]. The resulting complex was imaged by cryogenic-electron microscopy to yield a reconstruction of the GPR161-G_s complex at 2.7 Å resolution (Fig. 1.1c,d and Fig. 1.7) and enabled model building for the seven transmembrane domain of GPR161, the G_s subunits, Nb35, and, most notably, a sterol-like molecule (Fig. 1.1c,d and Fig. 1.8).

Our structure of the GPR161-G_s complex is similar to many other activated Class A GPCRs bound to heterotrimeric G proteins, like the prototypical β_2-adrenergic receptor (β_2AR) (Fig. 1.9a). A key hallmark of Class A GPCR activation is outward displacement of transmembrane helix 6 (TM6) to accommodate the C-terminal α-helix of Gα subunit[36]. While we do not have an inactive structure of GPR161 for comparison, the conformation of TM6 in GPR161 bound to G_s is similar to the outward displaced conformation observed for β_2AR (Fig. 1.9a). We conclude that our structure of GPR161-G_s captures the G_s coupled, active conformation of the receptor.

GPR161 extracellular loop 2 is self-activating

The structure of active GPR161 revealed a unique conformation of extracellular loop 2 (ECL2) compared to the majority of ligand-activated Class A GPCRs. Notably, the ECL2 of GPR161 forms a beta hairpin that folds over the extracellular face of the receptor to completely occlude the canonical orthosteric ligand binding pocket observed for many other Class A GPCRs (Fig. 1.1c,d and Fig. 1.2a). The comparison to β_2AR, for example, highlights that the GPR161 ECL2 occupies the same space that adrenaline occupies in β_2AR (Fig. 1.2b). The conformation of ECL2 in GPR161 is reminiscent of the orphan GPCR GPR52, which contains an ECL2 that also occludes the extracellular face of the receptor (Fig. 1.2b)[37]. In GPR52, ECL2 serves as a key determinant of constitutive activity - in effect, GPR52 is "self-activated" by ECL2. Indeed structures of several other orphan GPCRs, including GPR21 and GPR17 have recently revealed similar conformations of ECL2 associated with self-activation (Fig. 1.9b)[38,39].

To first verify the constitutive activity of GPR161, we used two different cellular assays: GloSensor assay to assess cAMP production and a miniG_s recruitment assay[40] using an optimized NanoLuciferase fragment complementation termed "NanoBiT"[41]. Expression of GPR161 in suspension adapted HEK293 cells gave consistently high levels of cAMP under basal conditions relative to empty vector and β_2AR (13 and 8 fold, respectively) (Fig. 1.2c and Fig. 1.10). However, GPR161 produced markedly less cAMP than GPR52 (Fig. 1.2c), perhaps highlighting that self-interaction of ECL2 is not sufficient to drive high basal activity. In a miniG_s protein fragment complementation assay, GPR161 basally recruited more miniG_s than β_2AR, with levels that are more similar to GPR52 (Fig. 1.2d). The results from these two

orthogonal assays demonstrate that GPR161 is constitutively active, albeit to a lesser extent than the self-activating orphan receptor GPR52.

Inspired by the example of GPR52, we examined the possible role of ECL2 in GPR161 activation. Several hydrophobic residues of the GPR161 ECL2 protrude into a region that overlaps with the canonical orthosteric pocket in other Class A GPCRs (Fig. 1.2a). We targeted several of these ($M177^{ECL2}$, $V179^{ECL2}$, $W182^{ECL2}$) for mutagenesis experiments to understand whether the conformation of ECL2 in GPR161 is important for constitutive activity. We substituted each of these positions with either alanine (to test for simple loss of the side chain) or arginine (to introduce a large perturbation in local hydrophobic contacts). We also examined a $W182G^{ECL2}$ mutant, which has previously been associated with rare cases of spina bifida, a neural tube developmental defect[42]. Mutation of these hydrophobic residues in ECL2 to either alanine or arginine caused a near complete loss of cAMP generation by GPR161, suggesting that the in *cis* interaction with ECL2 is essential for GPR161 activation (Fig. 1.2e). To explore the possibility that the GPR161 ECL2 may be more dynamic in the absence of miniG_s, we performed molecular dynamics simulations of GPR161 without miniG_s. In six simulations, we observed that ECL2 remains stably in a similar conformation as observed in the cryo-EM structure (Fig. 1.2f,g and Fig. 1.13b). We therefore surmise that ECL2 contributes to self-activation of GPR161.

A sterol facilitates GPR161 coupling to G_s

A surprising finding in the cryo-EM map of GPR161-G_s is the presence of a sterol-like density located at an extrahelical site near the cytoplasmic ends of transmembrane helices 6 and 7 (TM6

and TM7). Although the exact identity of this sterol is unclear, we tentatively modeled a cholesterol molecule in this density. We next sought to understand whether sterols engage this site and whether sterol binding at this site leads to signaling output for GPR161. Given the importance of sterols in metazoan Hedgehog pathway signaling[43–45], we first examined whether residues surrounding the putative sterol are conserved in evolution. Several key interacting residues (I323$^{7.52}$, W327$^{7.56}$ and R332$^{8.51}$) are conserved from humans to the echinoderm *Strongylocentrotus purpuratus* (Fig. 1.3b and Supplementary Fig. 1.11b).

To determine whether the observed density is indeed a sterol, we performed photoaffinity labeling experiments with purified GPR161-miniG_s and two cholesterol analogs: LKM38 and KK231. These sterol analogs contain an ultraviolet light activated diazirine group, either on the B-ring of the steroid (LKM38) or the aliphatic tail (KK231), that rapidly forms covalent adducts with proximal residues. Previous studies have demonstrated that these sterol photoaffinity analogs enable identification of functionally-relevant sterol binding sites in diverse membrane proteins[46–49]. After photoaffinity labeling GPR161-miniG_s, adducted residues were identified by tryptic digestion followed by LC-MS/MS sequencing of the resultant peptides. We obtained 76% sequence coverage of GPR161, with full residue-level sequencing of six of the seven transmembrane helices (Fig. 1.12b). Consistent with our binding pose for cholesterol, the diazirine group in the B ring of LKM38 labeled K267$^{6.32}$ in TM6 while the similar functional group in the tail of KK231 labeled residue C319$^{7.48}$ in TM7 (Fig. 1.3c,d and Fig. 1.12a). To determine whether the observed photolabeling is specific, we repeated this experiment in the presence of a 33-fold molar excess of unlabeled cholesterol. For both LKM38 and KK231 the

presence of unlabeled cholesterol completely suppressed photoaffinity labeling, suggesting that cholesterol itself can bind at this site (Fig. 1.3e).

We next sought to understand whether cholesterol binding promotes interactions between GPR161 and G_s. In the absence of an inactive-state structure of GPR161, we turned to molecular dynamics simulations to assess whether the presence of G_s is required for stable cholesterol binding (Fig. 1.4a,b and Fig. 1.13a). We simulated GPR161 either restrained to remain in the miniG_s conformation on the intracellular side or without any restraints. Each condition was simulated with six replicate simulations, each 1 μs in length. When GPR161 is restrained in the miniG_s bound conformation, 5 of 6 simulations showed stable cholesterol association with $W327^{7.56}$ (Fig. 1.13d). By contrast, in the absence of any restraints, $W327^{7.56}$ flipped inward into the seven transmembrane core of GPR161, thereby removing a key contact for cholesterol at the extrahelical binding site. Additionally, this rotamer of $W327^{7.56}$ would occlude binding of the C-terminal α-helix of G_s (Fig. 1.13e). Indeed, in 5 of 6 simulations of unrestrained GPR161, we observed that cholesterol rapidly disengages the extrahelical binding site and remains unbound for the remainder of the simulation (Fig. 1.13c). GPR161 did not transition to an inactive-like conformation in unrestrained simulations. These simulations therefore suggested that cholesterol binding to GPR161 at the TM6/TM7 extrahelical site is cooperative with G_s binding.

If cholesterol potentiates G_s binding, we predicted that disrupting the cholesterol binding site would both decrease cholesterol binding and may decrease cAMP production downstream of GPR161-induced G_s activation. Closer examination of the conserved residues in the sterol site highlighted that $I323^{7.52}$ interacts with the iso-octyl tail of cholesterol, $W327^{7.56}$ binds cholesterol

central rings, and R332$^{8.51}$ engages the hydroxyl group (Fig. 1.4c). We generated a GPR161-miniG$_s$ construct substituting alanine at these conserved positions (GPR161-AAA$^{7.52, 7.56, 8.51}$-miniG$_s$) and tested the ability of purified receptor preparations to bind ^{3}H-cholesterol using a scintillation proximity assay (Fig. 1.4d). In this assay, GPR161-AAA$^{7.52, 7.56, 8.51}$-miniG$_s$ bound cholesterol less than as compared with GPR161-miniG$_s$. We observe some residual binding of ^{3}H-cholesterol to GPR161-AAA$^{7.52, 7.56, 8.51}$-miniG$_s$, suggesting that I323$^{7.52}$, W327$^{7.56}$ and R332$^{8.51}$ help mediate the sterol interaction. Residual binding of ^{3}H-cholesterol may reflect GPR161-AAA$^{7.52, 7.56, 8.51}$-miniG$_s$ being trapped by miniG$_s$ in the G$_s$-interacting conformation. Our attempts to conduct this binding experiment with GPR161 alone was limited by the inability to purify the receptor in the absence of miniG$_s$. Supporting the importance of the sterol binding site, alanine mutation of W327$^{7.56}$ and R332$^{8.51}$ showed decreased cAMP production in a GloSensor assay compared to wild-type, while the GPR161-AAA$^{7.52, 7.56, 8.51}$ mutant ablates cAMP production (Fig. 1.4e). Importantly the GPR161-AAA$^{7.52, 7.56, 8.51}$ mutant showed even lower levels of cAMP production compared to V129E$^{3.54}$, a mutant previously designed to directly disrupt the predicted GPR161-G$_s$ interaction[38]. Similarly, GPR161-AAA$^{7.52, 7.56, 8.51}$ mutant did not recruit miniG$_s$ in a NanoLuciferase complementation assay[40] (Fig. 1.5d) while GPR161-V129E showed a more moderate decrease in miniG$_s$ recruitment (Fig. 1.14d). Our combined biochemical, simulation, and signaling studies show that cholesterol, and potentially other sterols, can bind GPR161 to support interactions with G$_s$, thereby promoting cAMP production.

GPR161 induced cAMP is dispensable for repression of GLI2 ciliary accumulation

We next aimed to understand how activation of GPR161 leads to Hedgehog pathway repression in cilia. Upon Hedgehog pathway activation, GLI2 proteins accumulate at the tips of cilia[50–52]. We previously showed removal of GPR161 from cilia, either as a *Gpr161* gene knockout or from disruption of GPR161 trafficking to cilia results in accumulation of GLI2 in ciliary tips of resting cells[6]. We therefore used GLI2 ciliary accumulation as a sensitive test for GPR161 function in this cellular compartment.

GPR161 localizes to primary cilia in ciliated NIH 3T3 cells in the basal condition, assessed by co-localization with the ciliary markers acetylated tubulin (AcTub) and pericentrin (PCNT). Concordant with Hedgehog pathway inhibition, GLI2 does not accumulate in the primary cilium in the basal condition (DMSO treatment). Addition of the small molecule Hedgehog pathway agonist SAG leads to exit of GPR161 from cilia and accumulation of GLI2 in the ciliary tip (Fig. 1.5a). In *Gpr161*$^{-/-}$ NIH 3T3 cells, GLI2 accumulates in the ciliary tip in the basal condition and addition of SAG does not further increase ciliary GLI2 levels; this phenotype is consistent with loss of ciliary Hedgehog pathway repression by GPR161. Expression of wild-type GPR161 in *Gpr161*$^{-/-}$ NIH 3T3 cells rescues repression of Hedgehog pathway in cilia, as indicated by a low level of GLI2 positive cilia in the basal condition.

We first examined whether GPR161 ECL2 mutants, which are defective for cAMP production, can rescue GLI2 ciliary suppression. Here, we found that ECL2 mutants of GPR161 failed to accumulate in the primary cilium, suggesting that these mutants are either defective in biogenesis

or are impaired in normal trafficking to the primary cilium (Fig. 1.14b). Consistent with a lack of GPR161 localization in cilia, GPR161 ECL2 mutants also failed to suppress GLI2 in ciliary tips.

We next turned to the GPR161-AAA$^{7.52, 7.56, 8.51}$ and GPR161-V129E$^{3.54}$ mutants to understand whether cholesterol-dependent G_s activation is important for GPR161 repression of the Hedgehog pathway in cilia. Unlike the ECL2 mutants, GPR161-AAA$^{7.52, 7.56, 8.51}$ and GPR161-V129E$^{3.54}$ expressed in *Gpr161*$^{-/-}$ NIH 3T3 cells showed robust ciliary localization similar to wild-type GPR161 (Fig. 1.5a and Fig. 1.14a,b). Like wild-type GPR161, activation of the Hedgehog pathway by SAG led to exit of GPR161-AAA$^{7.52, 7.56, 8.51}$ from the primary cilium. By contrast, GPR161-V129E$^{3.54}$ did not exit cilia upon Hedgehog pathway activation (Fig. 1.14a,b). Our prior studies with GPR161-V129E$^{3.54}$ suggested that this mutant reduced cAMP production[1]; we found here that GPR161-V129E$^{3.54}$ did not suppress ciliary GLI2 when expressed in *Gpr161*$^{-/-}$ NIH 3T3 cells (Fig. 1.14a,c). However, the inability of GPR161-V129E$^{3.54}$ to exit cilia, combined with residual interaction with miniG_s, suggests that this mutation may have consequences beyond decreased cAMP production. Surprisingly, expression of the GPR161-AAA$^{7.52, 7.56, 8.51}$ mutant suppressed GLI2 accumulation, consistent with repression of the Hedgehog pathway in the basal state (Fig. 1.5a,b). This unexpected result suggests that a G_s signaling-independent function of GPR161 is the predominant mediator for GLI repression in cilia in NIH 3T3 cells.

GPR161 is unique among many GPCRs in that it contains a PKA-binding AKAP domain. Previous studies have identified an amphipathic helix in the C-terminus of GPR161 that directly binds PKA regulatory subunits type I (RI)[38]; introduction of a single point mutant, L465P$^{C\text{-}term}$,

breaks this amphipathic helix and prevents PKA type I holoenzyme binding. Compared to the sterol-binding site, this PKA-binding site in GPR161 is less well conserved (Fig. 1.11b). We assessed whether disruption of PKA binding by GPR161 influences Hedgehog pathway repression. In HEK293 cells, expression of GPR161-L465P$^{C\text{-}term}$ led to constitutive cAMP production, albeit to a lesser extent than wild-type GPR161 (Fig. 1.5c). Indeed, GPR161-L465P$^{C\text{-}term}$ robustly recruited miniG_s in a NanoLuciferase fragment complementation assay[40] (Fig. 1.5d), indicating that this mutation does not influence GPR161 interactions with G_s. Consistent with previous reports, GPR161-L465P$^{C\text{-}term}$ recruited less PKA-RI than wild-type GPR161 (Fig. 1.5e). By contrast, the cAMP deficient mutants GPR161-AAA$^{7.52, 7.56, 8.51}$ and GPR161-V129E$^{3.54}$ recruited PKA-RI, albeit to slightly lower levels than the wild-type GPR161 (Fig. 1.5e and Fig. 1.14e).

Having validated that GPR161-L465P$^{C\text{-}term}$ attenuates interaction with PKA-RI, we tested whether this mutant represses the ciliary Hedgehog pathway in *Gpr161*$^{-/-}$ NIH 3T3 cells. We validated that GPR161-L465P$^{C\text{-}term}$ was located in primary cilia in the basal condition (Fig. 1.5a). Like wild-type GPR161, GPR161-L465P$^{C\text{-}term}$ exited cilia upon activation of the Hedgehog pathway by SAG (Fig. 1.5a and Fig. 1.14b). However, GPR161-L465P$^{C\text{-}term}$ was unable to repress GLI2 localization to the ciliary tip in the basal condition, indicating that PKA anchoring by GPR161 is critical for ciliary Hedgehog pathway control. Indeed, a double mutant combining disruption of G_s coupling and PKA binding (GPR161-AAA$^{7.52, 7.56, 8.51}$-L465P$^{C\text{-}term}$) was also unable to repress ciliary GLI2 in the basal condition (Fig. 1.14c). We conclude that GPR161 binding to PKA-RI is essential for Hedgehog repression in the primary cilia, while GPR161-induced G_s signaling is dispensable.

1.4 Discussion

Our studies here illuminate fundamental mechanisms of GPR161 activation, and how these mechanisms relate to GPR161-mediated regulation of Hedgehog signal transduction. Our cryo-EM structure of GPR161 revealed two stimuli contributing to GPR161 constitutive activity: first, a self-activating ECL2; and second, a sterol-like density at a unique extrahelical site. We demonstrate that ECL2 occludes the canonical Class A GPCR orthosteric site and is required for GPR161 trafficking to the primary cilia and cAMP signaling. We show that cholesterol can bind at the sterol site, that sterol-binding site availability is dependent upon the G protein-bound conformation of GPR161, and that the sterol-binding site regulates cAMP signaling. These two features of GPR161 activation illuminate the basis for G_s-induced cAMP constitutive activity observed in previous studies[1,6,21].

With these fundamental activation properties of GPR161, we provide new context into how GPR161 regulates the Hedgehog signaling pathway. A central model for GPR161 function in Hedgehog pathway repression is the importance of constitutive cAMP generation[1]. Optogenetic and chemogenetic triggers that elevate ciliary cAMP levels repress Hedgehog signaling[17]. We directly tested the importance of GPR161-induced cAMP production in one aspect of Hedgehog pathway repression, namely suppression of GLI2 transcription factor accumulation in the primary cilium in NIH 3T3 cells. We previously showed that a complete lack of GPR161 from cilia alone is important in suppressing GLI2 levels in the primary cilium[6]. Thus, this assay is a precise read out of GPR161 activity in cilia. Surprisingly, a GPR161 mutant that is unable to couple to G_s and support cAMP production (GPR161-AAA) retains the ability to suppress ciliary GLI2 accumulation in NIH 3T3 cells, suggesting that cAMP production of GPR161 is less

crucial for Hedgehog pathway repression in cilia than current models suggest. Rather, we find that preventing the anchoring of the PKA type I complexes to GPR161 plays a more important role in suppressing GLI2 levels in cilia.

We propose the following model for repression of the Hedgehog pathway by GPR161 in cilia (Fig. 1.5f). In the absence of Hedgehog, GPR161 bound to PKA is localized to primary cilia. PKA within the cilia phosphorylates GLI resulting in processing into its repressor form. The general presence of ciliary cAMP is important for this process but could be generated by other ciliary GPCRs[51], by receptor independent activation[53] of adenylyl cyclases by G_s in cilia, or by G-protein independent activity of adenylyl cyclase[54]. A complete disruption of transport of adenylyl cyclases into cilia from upstream maturation defects during ER-Golgi transit in the secretory pathway, as seen in the *Ankmy2* knockout, prevents GLI-R formation in embryos and results in GLI2 accumulation in cilia[55]. In addition, although basal levels of cAMP in cilia are controversial, with reports ranging from levels comparable to that in the cytoplasm[56] to supraphysiological levels ~4.5 μM[54], PKA-C can be activated with sub-optimal cAMP levels[57]. In the presence of Hedgehog, GPR161 traffics out of the cilia, removing PKA-RI with it[58].

The above model does not exclude the role of extraciliary GPR161, particularly that in the periciliary endosomal compartment[6], in GLI-R processing and thereby regulating tissue-specific morpho-phenotypes. We have recently demonstrated that GPR161 is not only localized to primary cilia but is also located in periciliary endosomes[6]. Both ciliary and extraciliary pools of GPR161 contribute to GLI-R formation and regulate tissue-specific repression of Hedgehog pathway in mice[6]. Although we show that GPR161 AKAP activity is critical for suppression of

GLI2 trafficking in cilia, the AKAP function of GPR161 is not fully necessary for its suppression of Hedgehog pathway phenotypes in zebrafish[57–60]. The most parsimonious model explaining these paradoxical results would be that while ciliary AKAP function of GPR161 is critical for suppression of GLI2 trafficking to cilia, GPR161 functions outside cilia through sterol-mediated cAMP signaling. Production of cAMP has been reported for other GPCRs in the endosomal compartment[59–62]. Further experiments in organismal models will be needed to test the role of cAMP and AKAP signaling by GPR161 in the extraciliary endosomal compartment in tissue-specific Hedgehog signaling.

Our revised model for GPR161 provides a compelling parallel to recent reports that highlight direct interactions between Smoothened and PKA-C in Hedgehog signaling. As a G_i-coupled GPCR, Smoothened suppression of cAMP generation was initially described as critical for Hedgehog signaling in the fly[63]. Subsequent studies in vertebrates have called into question the importance of Smoothened-induced G_i signaling in cilia in vertebrates[62,64]. More recently, the identification of a PKA-inhibitory motif in the Smoothened C-terminus suggests that activated Smoothened directly sequesters the catalytic subunits of PKA (PKA-C) to suppress enzymatic activity[65]. Instead of acting via cAMP on PKA, we propose that two GPCRs important to Hedgehog signaling, GPR161 and Smoothened, predominantly depend on binary interactions with PKA-C or RI subunit complexes to regulate the Hedgehog pathway in cilia (Fig. 1.5f).

Our identification of a conserved sterol binding site in GPR161 raises fundamental questions about the role of sterols in control of GPR161 signaling. We compared relative conservation of the GPR161 sterol binding site to the PKA-binding helix in the C-terminus (Fig. 1.11b).

Although the sterol binding site is conserved in deuterostome genomes, the PKA-binding motif is not clearly identified in echinoderms (e.g., *S. purpuratus*), early branching chordates (e.g., *B. belcheri*), and hemichordates (e.g., *S. kowalevskii*). The strong conservation of the cholesterol binding site, and the importance of this site for GPR161 to couple to G_s and generate cAMP, points to sterol driven cAMP generation of GPR161 having an important biological function.

Our finding that the conserved sterol binding site is not critical for controlling GLI2 levels in the primary cilium suggests several possibilities. First, as noted above, it is possible that extraciliary GPR161 uses a mechanism distinct from its ciliary AKAP function to control the Hedgehog pathway. For example, extraciliary GPR161 may depend on cholesterol or other sterols to promote cAMP formation to control GLI-R formation. Second, it is entirely possible that cAMP production by GPR161 has roles in adult tissues outside of the Hedgehog pathway. Indeed, GPR161 is expressed in many different cell types in adult tissues that are not hedgehog regulated, such as the adult hippocampus CA1 region[66]. GPR161 has also been reported to localize to cilia of hippocampal neurons[8,66,67]. Ciliary peptidergic GPCR signaling in the CA1 pyramidal neurons has been recently shown to regulate chromatin accessibility[68], but the role of cAMP signaling mediated by cilia remains unknown. Third, GPR161 could also have roles in cancers beyond its role in Hedgehog pathway repression in medulloblastoma[10,14]. For example, *GPR161* is overexpressed in triple-negative breast cancer and has been proposed to promote cell proliferation and invasiveness in tumor cells by forming a signaling complex with β-arrestin2 and IQGAP1[11,69]. Our identification of a GPR161 mutant that specifically attenuates cAMP production will enable a careful dissection of these potential roles of GPR161, irrespective of its function in cilia and beyond the role proposed in the Hedgehog pathway.

Our study highlights how GPCR cAMP-PKA signaling establishes precise signaling microdomains in primary cilia. Such restrictive signaling in nanodomains has been an emerging feature of subcellular signaling by cAMP[70,71]. Most broadly, our work highlights that orphan GPCRs may have functions beyond the biological pathway where they are first encountered. Directly observing the stimuli that activate orphan GPCRs will enable precise approaches to dissect the functional relevance of a specific signaling pathway in a biological outcome. The advent of structure-based methods to interrogate orphan GPCRs will therefore broaden views on the possible biology coordinated by this fascinating family of understudied proteins.

1.5 Figures

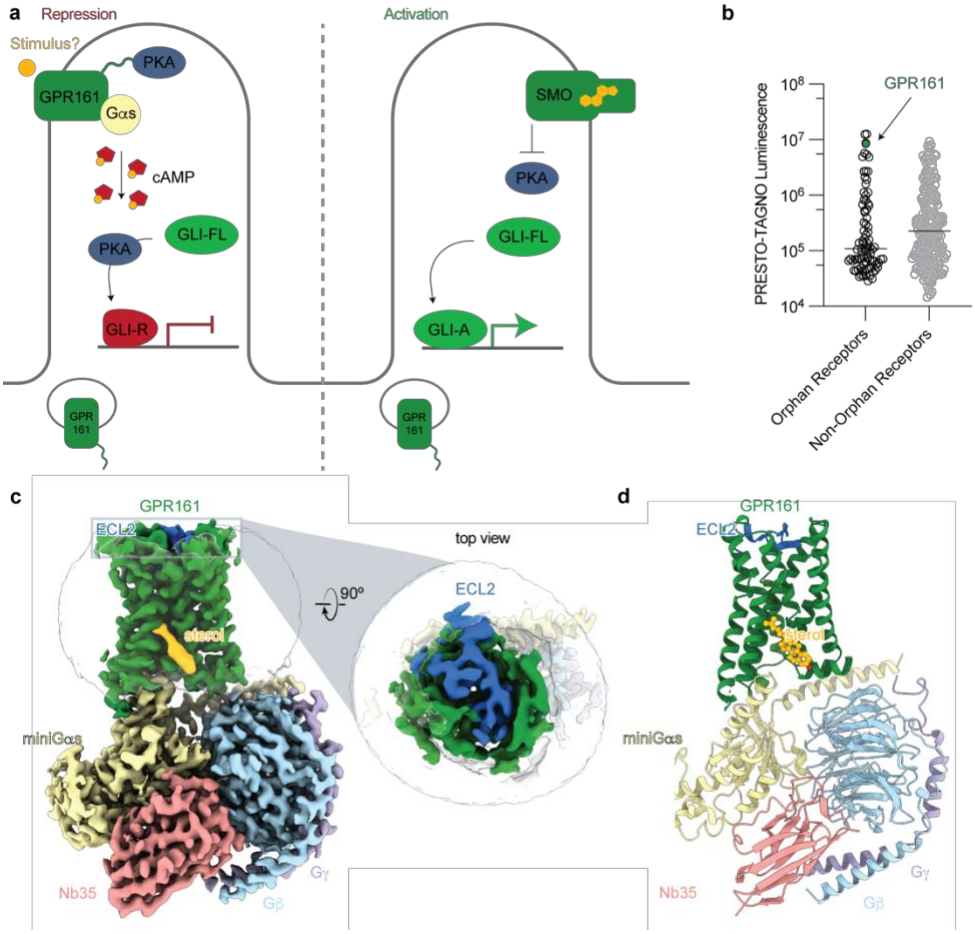


Figure 1.1 Structure-inspired deorphanization of GPR161

a) The Hedgehog pathway is regulated by two key GPCRs, GPR161 and SMO. In the absence of Hedgehog, GPR161 represses the pathway through constitutive cAMP signaling from an unknown stimulus. In the presence of Hedgehog, SMO activates the pathway by entering the cilia, binding cholesterol, and inhibiting PKA, while GPR161 exits cilia. **b)** GPR161 yields an exceptionally strong signal for β-arrestin recruitment in the PRESTO-Tango assay when compared across 314 GPCRs (data replotted from Kroeze WM *et al* 2015[22]). This assay is performed in a modified HEK293 cell line, suggesting that GPR161 is constitutively active under heterologous expression conditions. **c)** Cryo-EM density map of GPR161 in complex with G_s heterotrimer (miniGα_s, Gβ, and Gγ) and stabilizing nanobody 35 (Nb35). The map reveals a density consistent with the shape of a sterol (yellow) and an extracellular loop 2 (ECL2, blue) that is packed within the seven transmembrane core of GPR161. **d)** Ribbon diagram of activated GPR161 heterotrimer complex. Cholesterol is modeled into the sterol density (yellow).

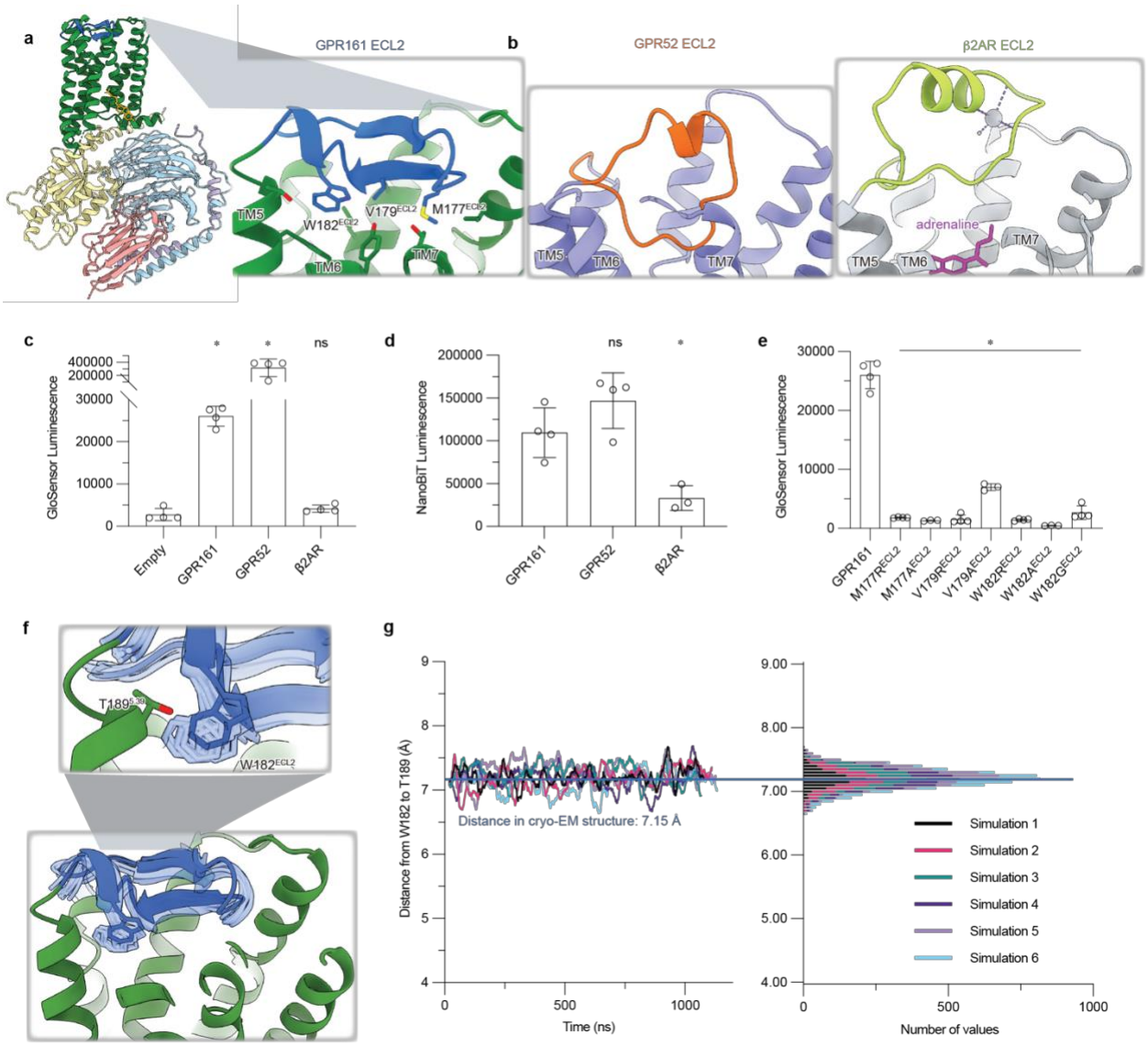


Figure 1.2 Extracellular loop 2 of GPR161 occupies classic GPCR orthosteric site

ECL2 of GPR161 makes hydrophobic contacts with the core of the receptor. **b)** Comparison of ECL2 of the self-activating orphan GPCR GPR52 (PDB ID: 6LI3[37]) and the prototypical agonist-activated GPCR β2-adrenergic receptor (β_2AR) bound to agonist adrenaline (PDB ID: 4LDO[72]). **c)** cAMP production assay for GPR161, GPR52, and β_2AR. GPR161 is constitutively active for cAMP production. Data are mean ± sd, n=4 biologically independent samples (*$P < 0.05$; ns, not significant; one-way ANOVA followed by Dunnett's multiple comparison tests). **d)** Nanoluc complementation assay for receptor recruitment of miniG_s. Both GPR52 and GPR161 constitutively recruit miniG_s. Data are mean ± sd, n=3-4 biologically independent samples (*$P <$ 0.05; ns, not significant; one-way ANOVA followed by Dunnett's multiple comparison tests). **e)** cAMP production assay assessing mutations in ECL2 of GPR161 for residues that make hydrophobic contacts with the transmembrane bundle. Data are mean ± sd, n=4 biologically independent samples (*$P < 0.05$; ns, not significant; one-way ANOVA followed by Dunnett's multiple comparison tests). **f)** Molecular dynamics simulations snapshots of GPR161. **g)** ECL2 stably occupies the canonical ligand binding pocket as represented by rolling average of the

distance between ECL2 W182^{ECL2} and the Cα atom of T189$^{5.39}$ in molecular dynamics simulations (6 independent trajectories).

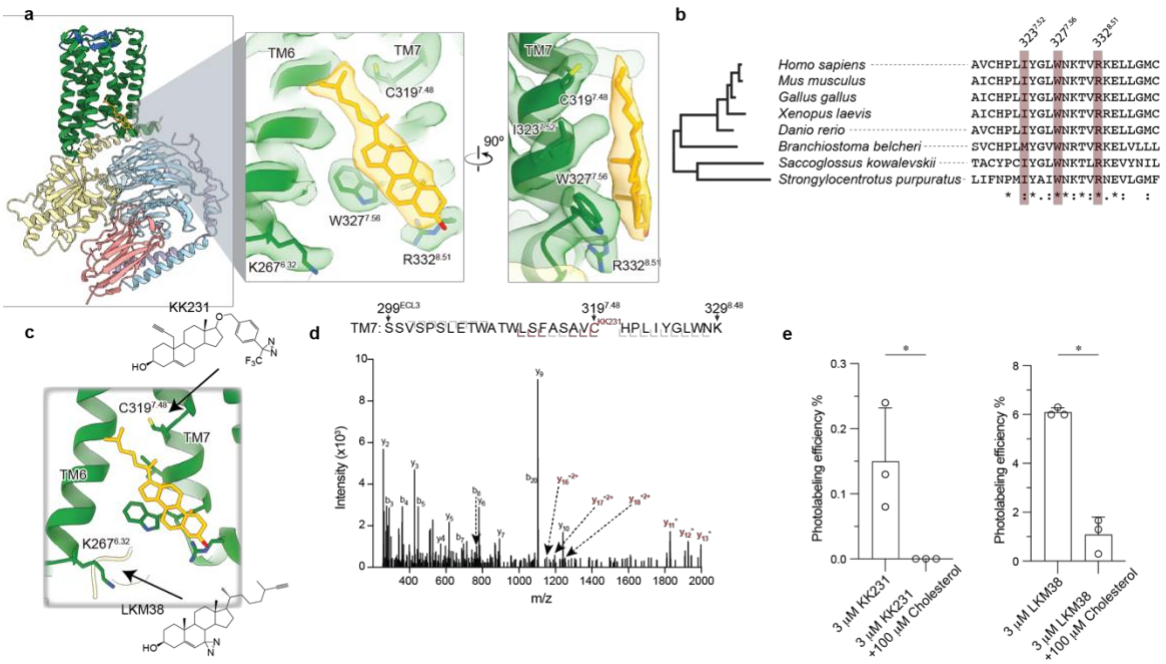


Figure 1.3 GPR161-miniG$_s$ stably and specifically binds cholesterol

a) Close up view of cholesterol bound to the intracellular side of transmembrane helix 6 (TM6) and TM7. Three key interacting residues (I323$^{7.52}$, W327$^{7.56}$, R332$^{8.51}$) are highlighted as sticks. **b)** GPR161 cholesterol binding site residues are conserved from humans to sea urchins (*Strongylocentrotus purpuratus*). A full alignment from these organisms used to define the dendrogram on left is shown in Supplementary Fig. 6. **c)** Detergent solubilized GPR161-miniG$_s$ purified without cholesterol hemisuccinate was incubated with either one of two distinct photoaffinity cholesterol analogs (KK231 or LKM38) and then crosslinked with >320 nm UV light. The resulting photo-labeled preparation was digested with trypsin and analyzed by collision-induced dissociation mass spectrometry, which revealed that KK231 labels position C319$^{7.48}$ while LKM38 labels K267$^{6.32}$. **d)** Product ion spectrum of KK231-labeled GPR161-miniG$_s$ sample with peptides mapped to TM7 and Helix 8. This peptide is modified with a mass consistent with KK231 at position C319$^{7.48}$. Red brackets and peaks indicate product ions that contain the KK231 adduct. **e)** Photolabeling efficiency of GPR161-miniG$_s$ by KK231 and LKM238 in the absence or presence of excess unlabeled cholesterol. Data are mean ± s.d of n=3 technically independent replicates from two independently prepared protein samples (*P < 0.05, Student's paired t-test).

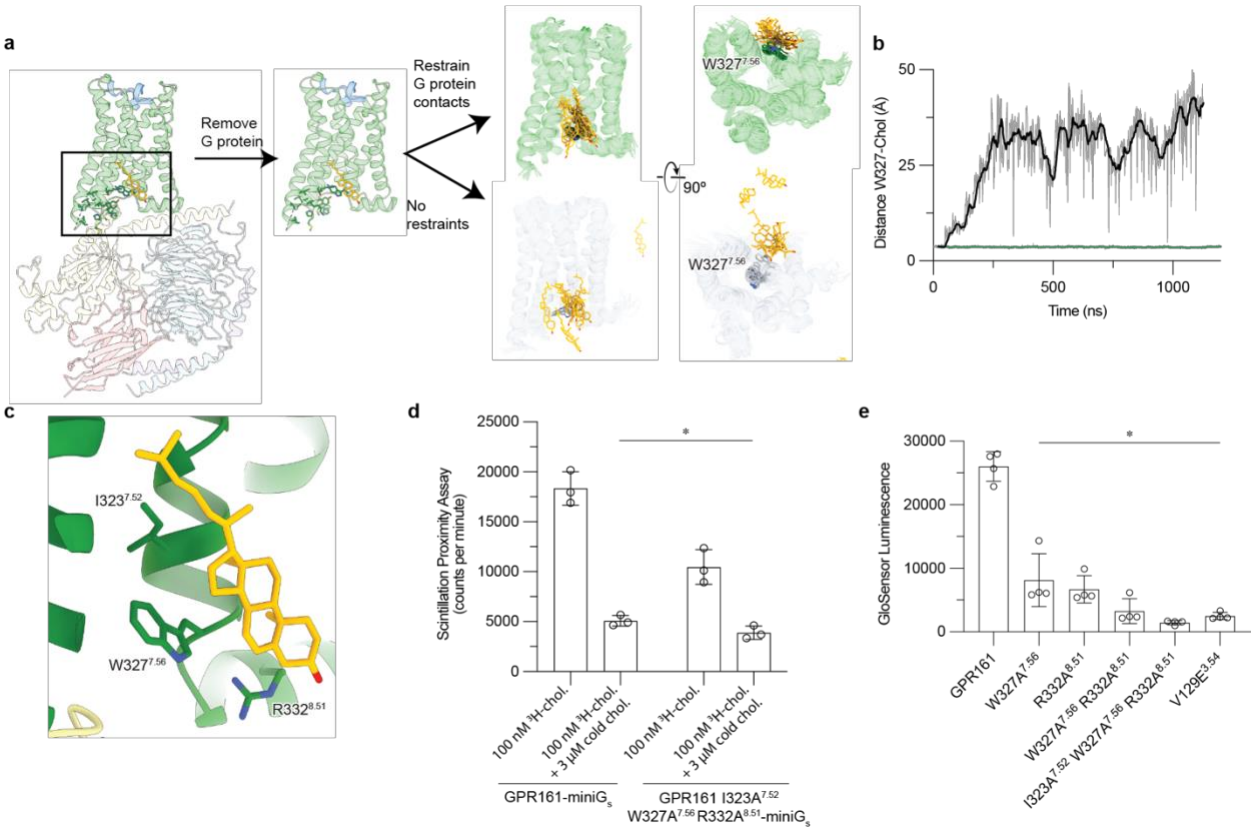


Figure 1.4 Cholesterol binding to GPR161 facilitates G_s coupling

a) Molecular dynamics simulation snapshots of active GPR161 both unrestrained (gray) and restrained to G protein bound conformation (green). When GPR161 is restrained to be in the G protein-bound conformation, cholesterol remains consistently near $W327^{7.56}$ as shown in both simulation snapshots and in a time trajectory from a representative simulation. **b)** By contrast, when GPR161 is not restrained to be in the G protein-bound state, the cholesterol is dynamic and moves away from $W327^{7.56}$. **c)** Close up view of cholesterol bound to the intracellular side of transmembrane helix 6 (TM6) and TM7. Three key interacting residues ($I323^{7.52}$, $W327^{7.56}$, $R332^{8.51}$) are highlighted as sticks. **d)** ^{3}H-cholesterol (^{3}H-chol) binding assay for purified GPR161-miniG_s and GPR161-$AAA^{7.52, 7.56, 8.51}$-miniG_s. Individual technical replicates are shown, bar graphs and error represent mean and s.d. (*$P < 0.05$; two-way ANOVA followed by Dunnett's multiple comparison tests). **e)** GloSensor cAMP production assay assessing mutations in cholesterol binding site of GPR161. Data are mean ± s.d. from n=4 biologically independent replicates (*$P <$ 0.05; one-way ANOVA followed by Dunnett's multiple comparison tests).

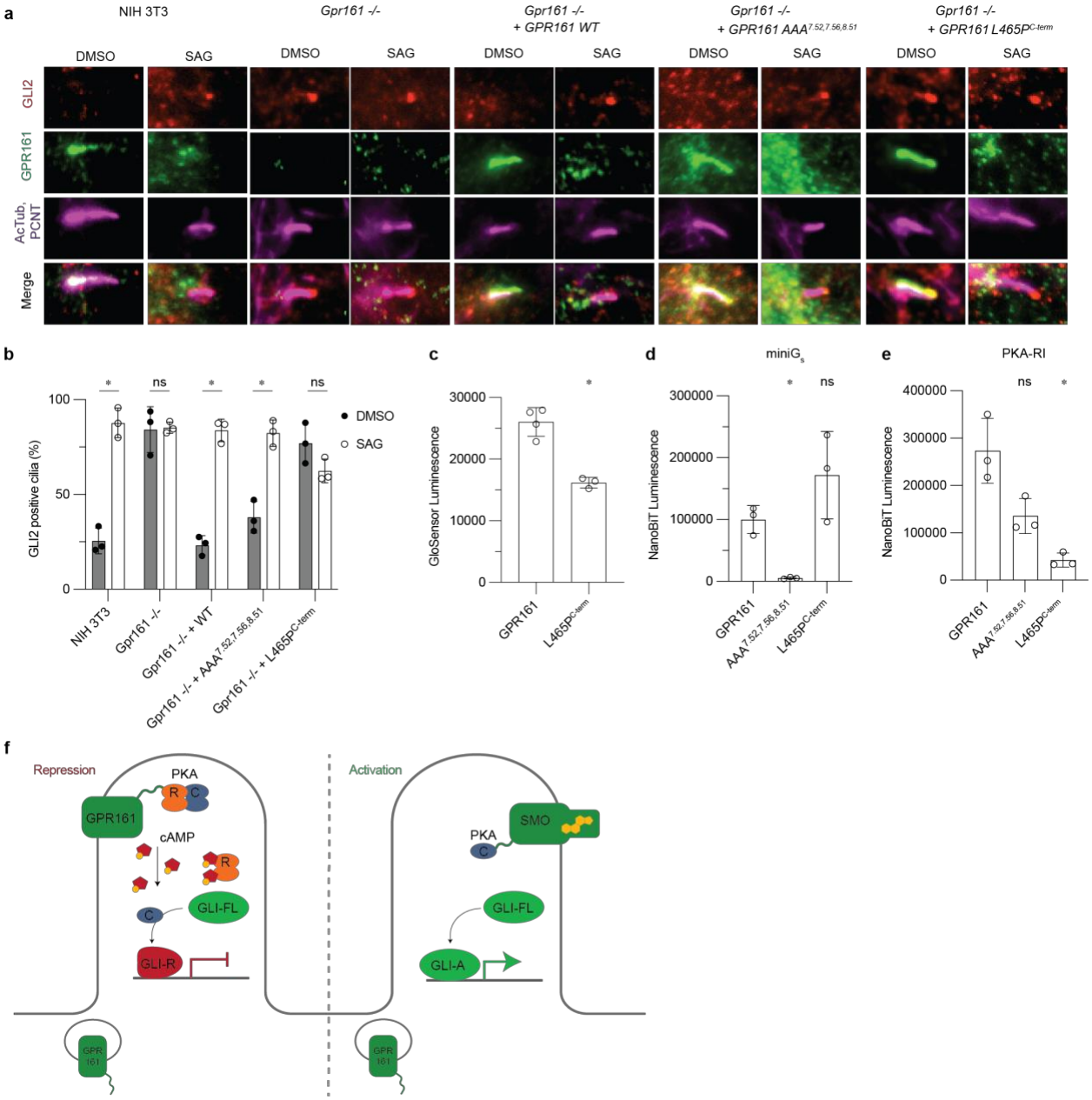


Figure 1.5 GPR161 PKA-RI binding, but not cAMP generation, is necessary to repress ciliary trafficking of GLI2

a) Representative images of the effect of site-directed mutagenesis of GPR161 on ciliary localization and GLI2 repression in ciliary tips in NIH 3T3 cells. NIH 3T3 Flp-In CRISPR based *Gpr161*$^{-/-}$ cells stably expressing untagged mouse wild-type or *GPR161* mutants were starved for 24 hr upon confluence and were treated for further 24 hr ± SAG (500 nM). After fixation, cells were immunostained with anti-GLI2 (red), anti-GPR161 (green), anti-acetylated, and centrosome (AcTub; PCNT purple) antibodies. Scale bar, 5 µm. **b)** Quantification GLI2 positive cilia indicating Hedgehog pathway activation. AAA$^{7.52,\ 7.56,\ 8.51}$ rescues function similar to WT, and L465P$^{C\text{-term}}$ does not, similar to *Gpr161*$^{-/-}$. Quantification of GLI2 positive cilia are shown from 3 biologically independent experiments from images taken from 2-3 different regions/experiment and counting 15-30 cells/region. Data are mean ± s.d. (*P < 0.05; ns, not significant; two-way

ANOVA followed by Šidák's multiple comparison tests). **c)** cAMP production assay assessing L465P$^{C\text{-}term}$ mutation. Data are mean ± sd, n=3-4 biologically independent samples (*$P < 0.05$; one-way ANOVA followed by Dunnett's multiple comparison tests). **d)** Nanoluc complementation assay for receptor recruitment of miniG_s. Both GPR161 and L465P$^{C\text{-}term}$ constitutively recruit miniG_s while AAA$^{7.52,\ 7.56,\ 8.51}$ does not. Data are mean ± sd, n=3-4 biologically independent samples (*$P < 0.05$; ns, not significant; one-way ANOVA followed by Dunnett's multiple comparison tests). **e)** Nanoluc complementation assay for receptor recruitment of PKA-RI. GPR161, AAA$^{7.52,\ 7.56,\ 8.51}$, and L465P$^{C\text{-}term}$ each recruit less PKA-RI, respectively. Data are mean ± sd, n=3-4 biologically independent samples (*$P < 0.05$; ns, not significant; one-way ANOVA followed by Dunnett's multiple comparison tests). **f)** PKA-centric model of Hedgehog pathway repression in cilia. In the absence of Hedgehog, GPR161 represses the pathway in cilia through coupling PKA. GPR161 also functions in periciliary endosomal compartments in regulating GLI-R formation[71]. In the presence of Hedgehog, SMO activates the pathway by entering the cilia, binding cholesterol, and sequestering PKA-C, while GPR161 exits cilia.

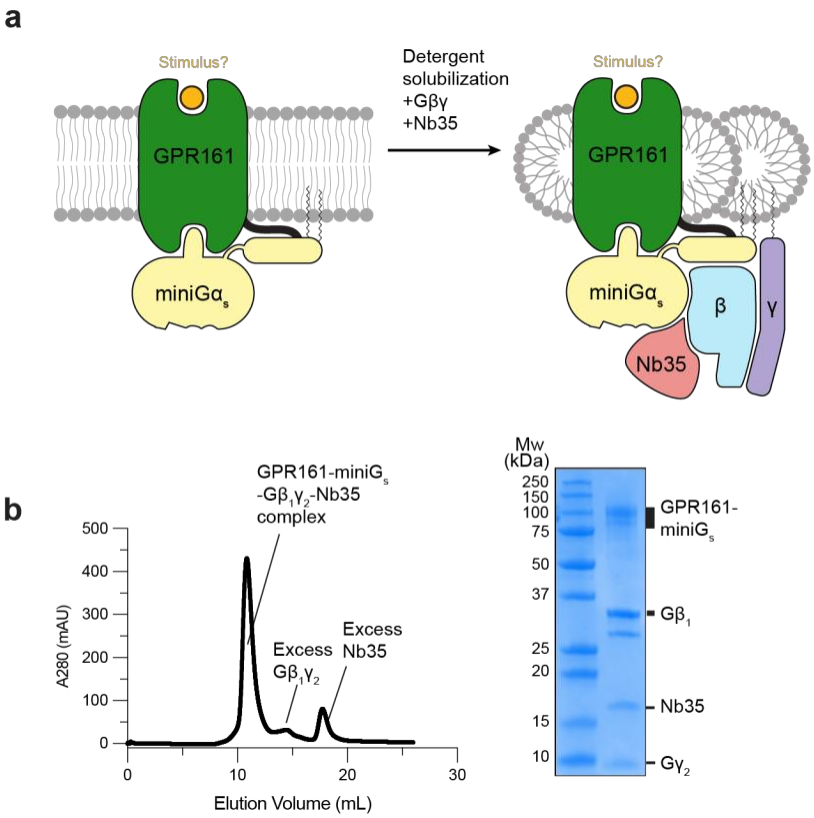


Figure 1.6 Biochemical preparation of GPR161-miniG$_s$ complex
a) Cartoon depiction of GPR161 stabilization, solubilization, and purification. **b)** Size-exclusion chromatogram (left) and SDS-PAGE gel (right) of purified GPR161-G$_s$ complex with Nb35.

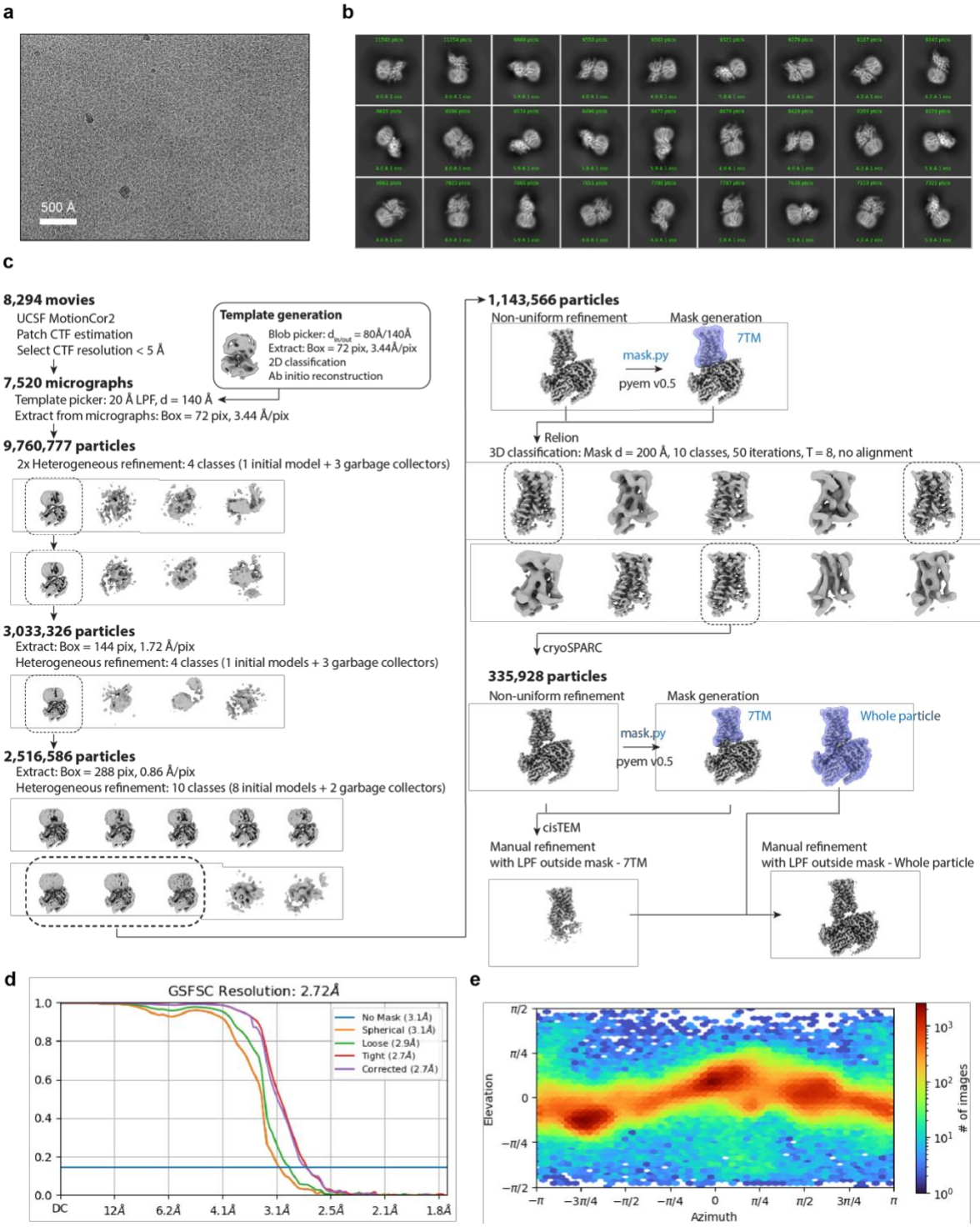


Figure 1.7 Cryogenic electron microscopy processing of GPR161

a) A representative motion-corrected cryogenic electron microscopy (cryo-EM) micrograph obtained from a Titan Krios microscope (n = 8,294)**. b)** A subset of highly populated, reference-free 2D-class averages. **c)** Schematic showing the cryo-EM data processing workflow. Initial processing was performed using UCSF MotionCor2 and cryoSPARC. Particles were transferred using the pyem script package to RELION for alignment-free 3D classification. Finally, particles were processed in cisTEM using the manual refinement job type with a 7TM mask followed by a

full particle mask. Dashed boxes indicated selected classes. **d)** Gold-standard Fourier Shell Correlation (GSFSC) curve for final refined and sharpened map computed in cryoSPARC. **e)** Euler angle distribution of final refined map computed in cryoSPARC.

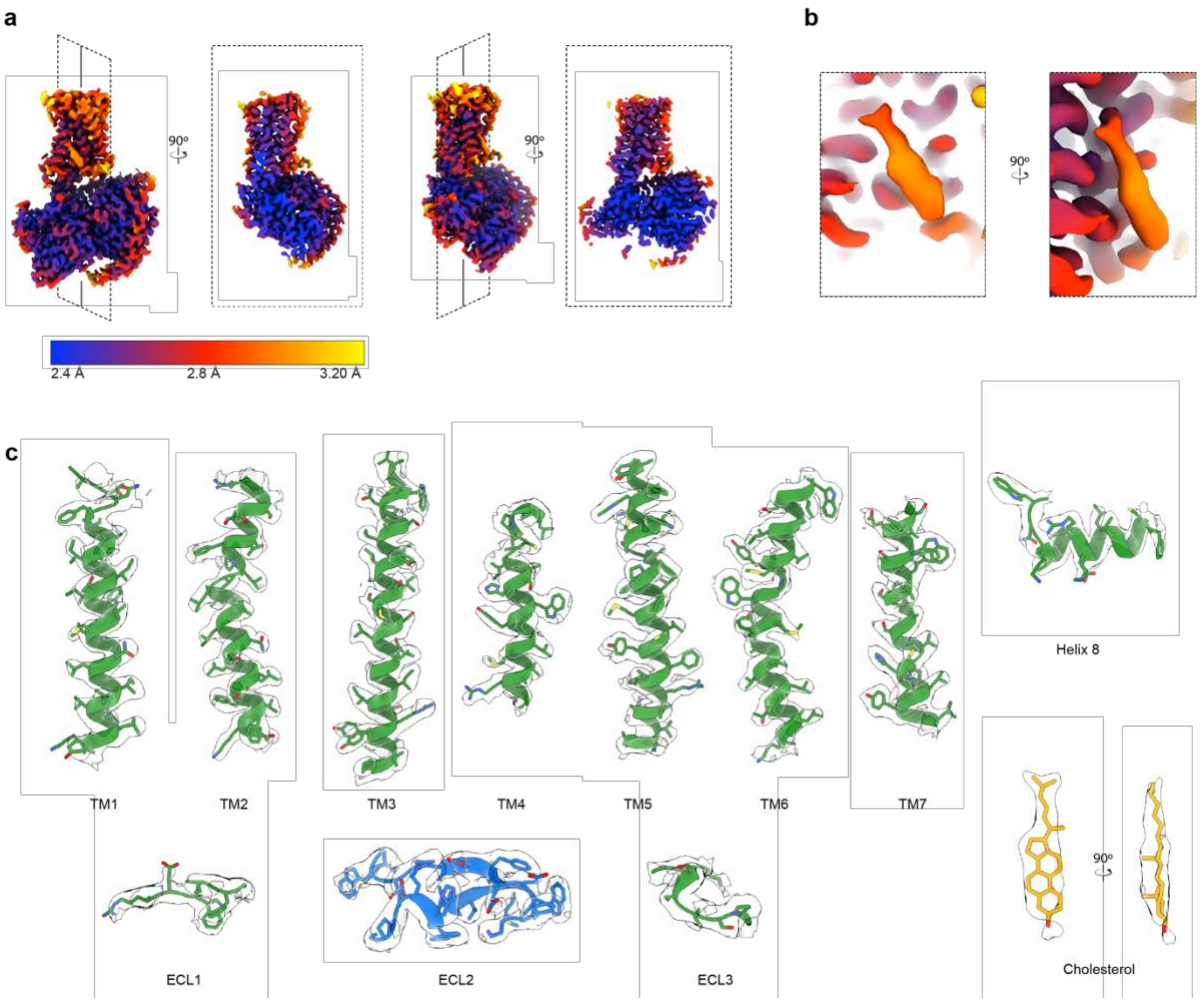


Figure 1.8 Cryo-EM local density

a) Orthogonal views of local resolution for the sharpened, final map of GPR161-G_s complex computed with local resolution in cryoSPARC. **b)** Close up of local resolution for sterol density. **c)** Isolated cryo-EM densities from the unsharpened, final map of GPR161 complex. Shown are the transmembrane (TM) helices, extracellular loops, and cholesterol-like density.

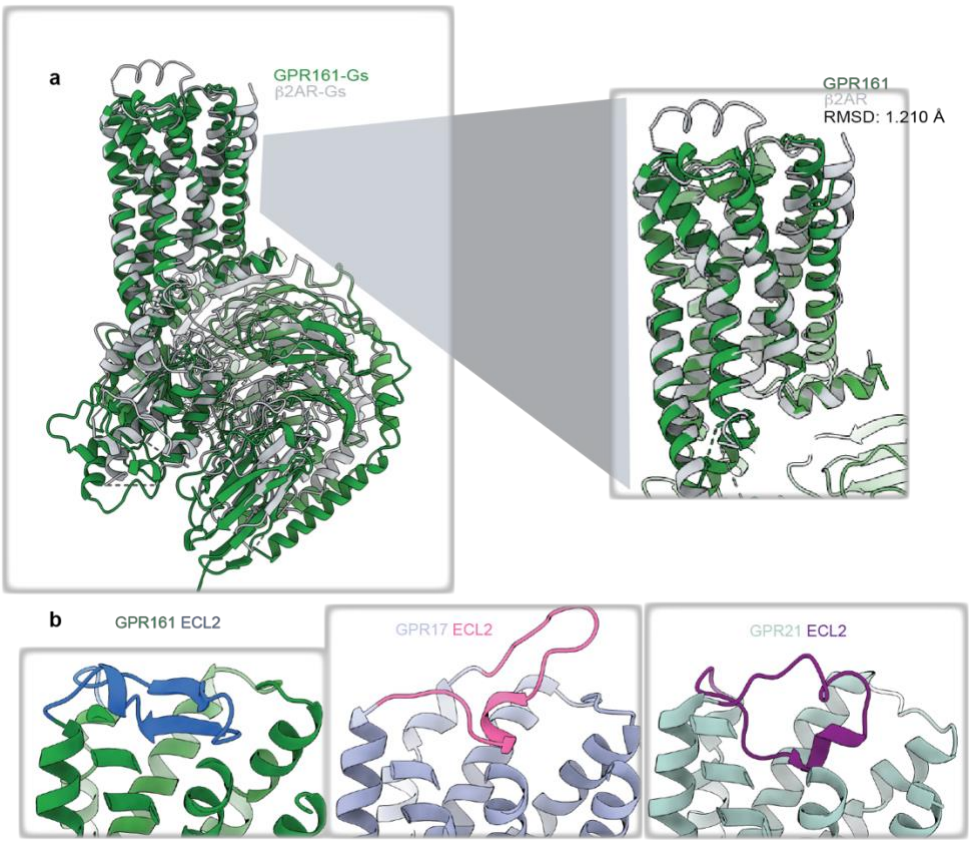


Figure 1.9 Comparison to additional GPCR structures

a) Structural comparison of GPR161 heterotrimer complex and β_2AR heterotrimer complex (PDB ID: 3SN6 [72]). GPR161 has the same hallmarks of GPCR activation as the prototypical receptor, β_2AR **b)** Structural comparison of GPR161 to other orphan GPCRs with self-activating ECL2, including GPR17 (PDB ID: 7Y89) and GPR21 (PDB ID: 8HMV)[38,39]. The cis-interaction of ECL2 with the canonical ligand-binding site is seen across self-activating orphan GPCRs but the precise loop conformation changes between receptors.

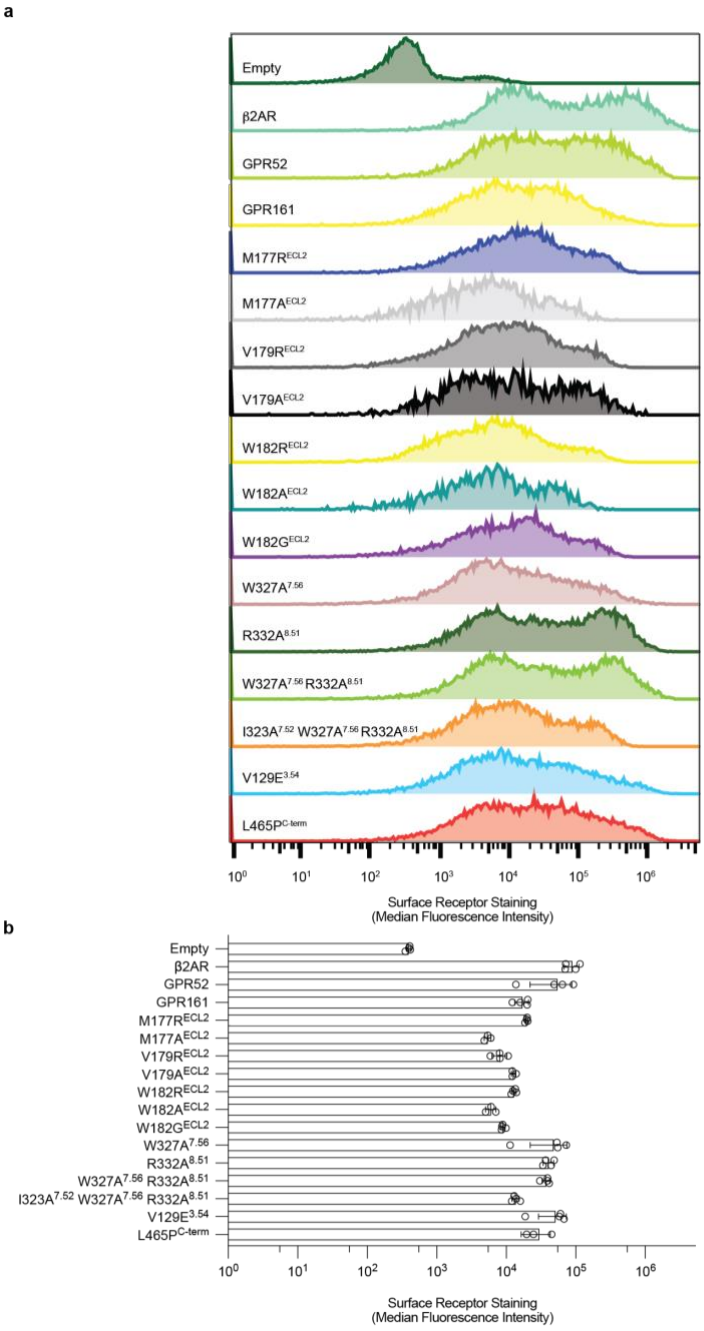


Figure 1.10 Surface expression of GPR161 mutants

a) Representative flow cytometry surface expression histograms for receptors and mutants used in cell-based assays. **b)** Surface expression of receptors and mutants quantified by anti-FLAG-A647 median fluorescence intensity ± sd from n = 3-4 biologically independent samples.

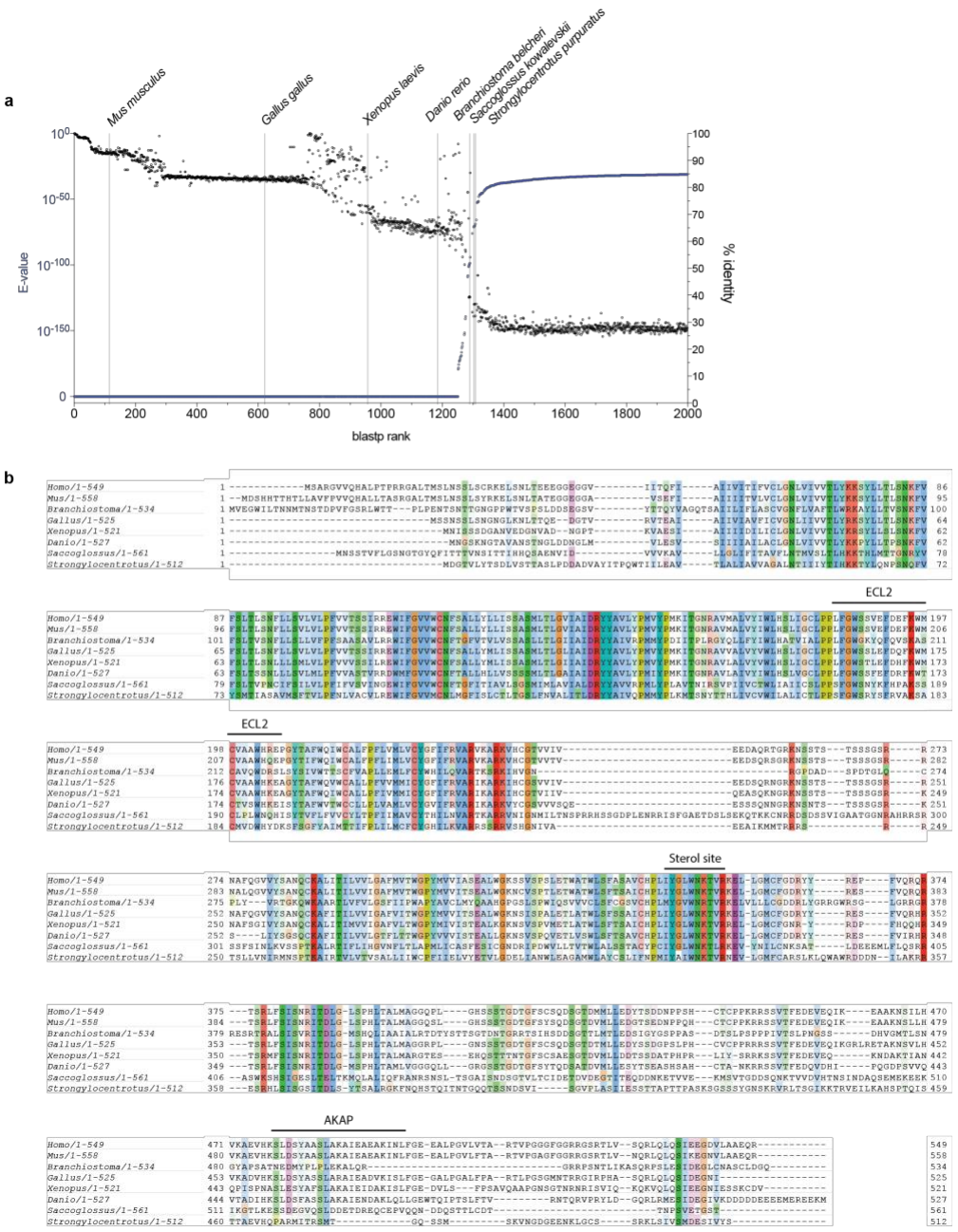


Figure 1.11 Phylogenetic analysis of GPR161

a) BLAST search results for Human *GPR161* (Uniprot: Q8N6U8). Sequences are plotted from highest confidence (E-Value) and highest sequence identity (% identity) to lowest. Representative organisms spanning the full range of homologous GPR161 sequences are listed. **b)** Full sequence alignment of eight *GPR161* model organism sequences identified in BLAST search.

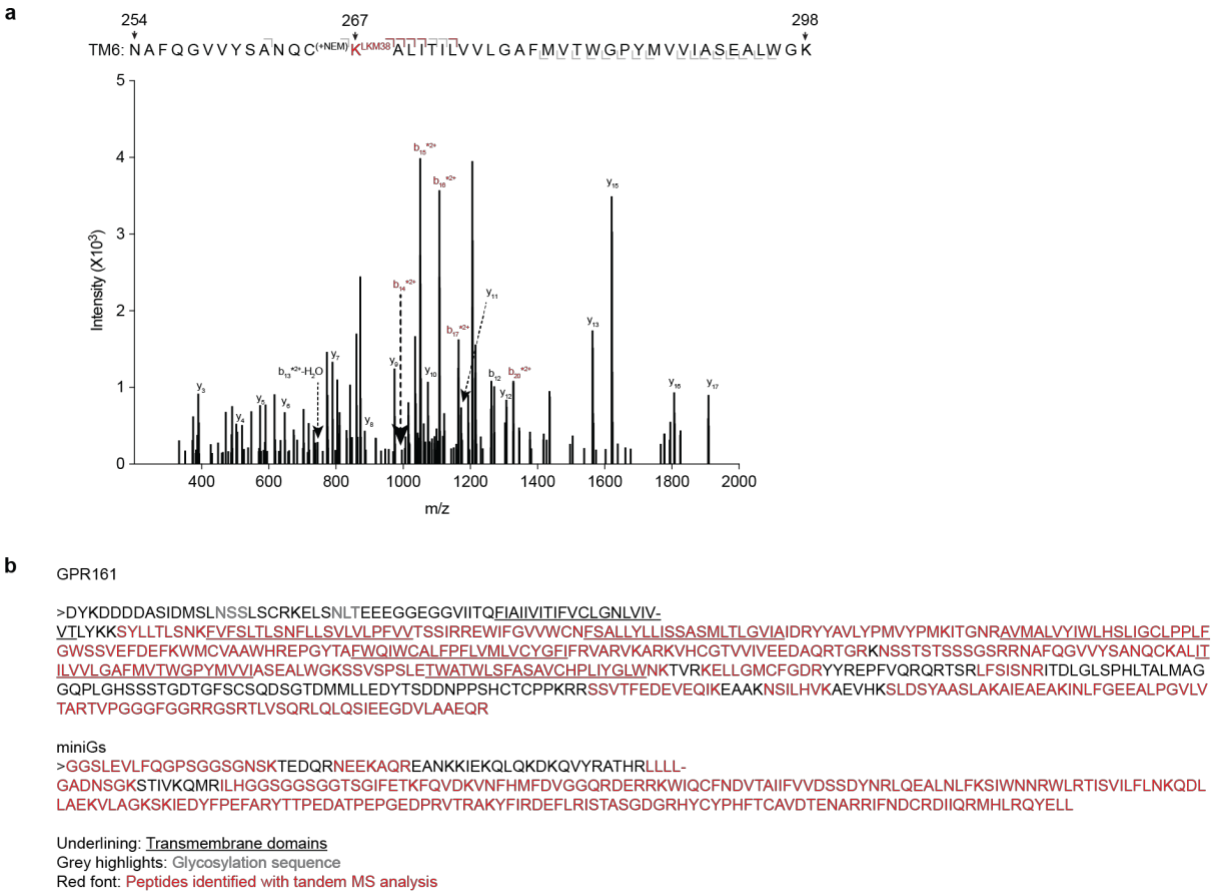


Figure 1.12 Photolabeling with LKM238 and mass spectrometry sequence coverage
a) Product ion spectrum of LKM238-labeled GPR161-miniG_s with peptides mapped to TM6. This peptide is modified with a mass consistent with LKM238 at position K267$^{6.32}$. Red brackets and peaks indicate product ions that contain the LKM238 adduct. **b)** Mass spectrometric sequence coverage of GPR161-miniG_s. Underlined segments indicate transmembrane spanning helices, red font indicates peptides identified by tandem MS analysis and gray font indicates glycosylation sites.

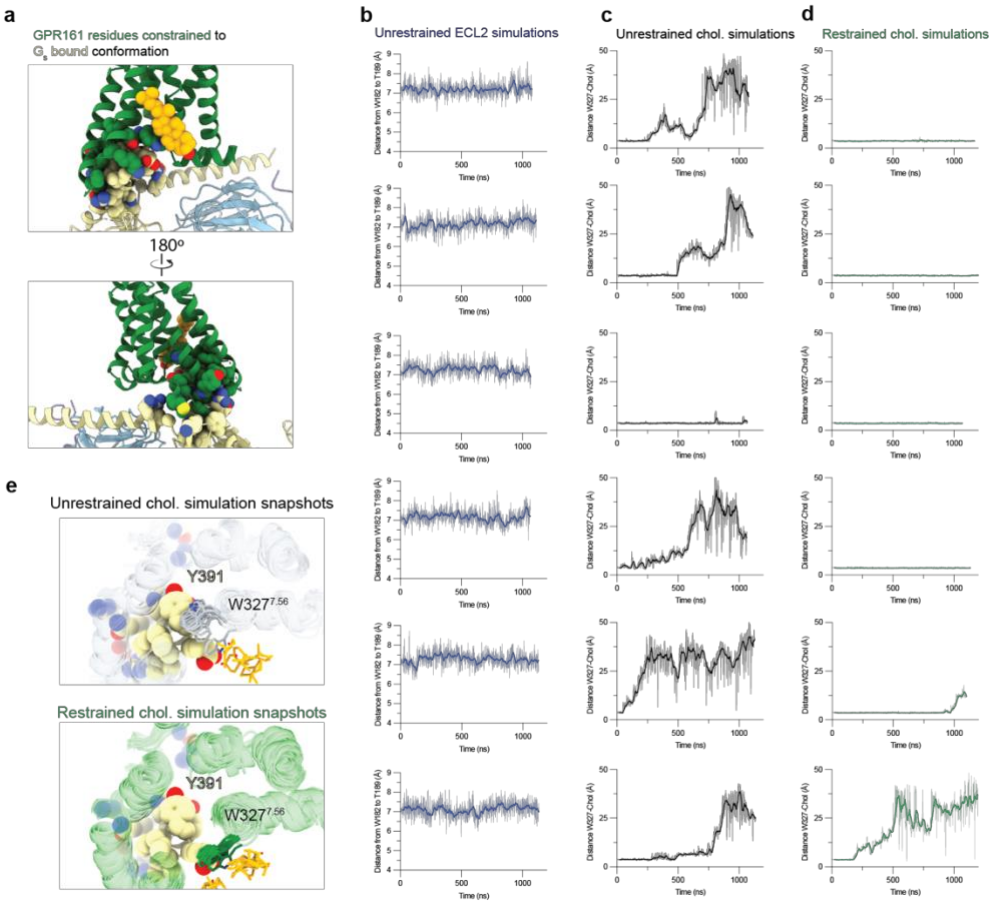


Figure 1.13 GPR161 molecular dynamics simulation trajectories

a) Structural diagrams of the GPR161 residues that contact G_s, which are restrained to their G_s bound conformation during restrained simulations. **b)** Independent trajectories for unrestrained simulations quantifying all atoms distance from W182^{ECL2} to T189$^{5.39}$. **c)** Independent trajectories for cholesterol-bound GPR161 without G protein-contacts restrained. All non-hydrogen atom distance from W327$^{7.56}$ to cholesterol is plotted. **d)** Independent trajectories for cholesterol-bound GPR161 with G protein-contacts restrained. All non-hydrogen atom distance from W327$^{7.56}$ to cholesterol is plotted. **e)** Simulation snapshots of W327$^{7.56}$ inward flip, which removes a key contact for cholesterol and occludes binding of the C-terminal α-helix of G_s.

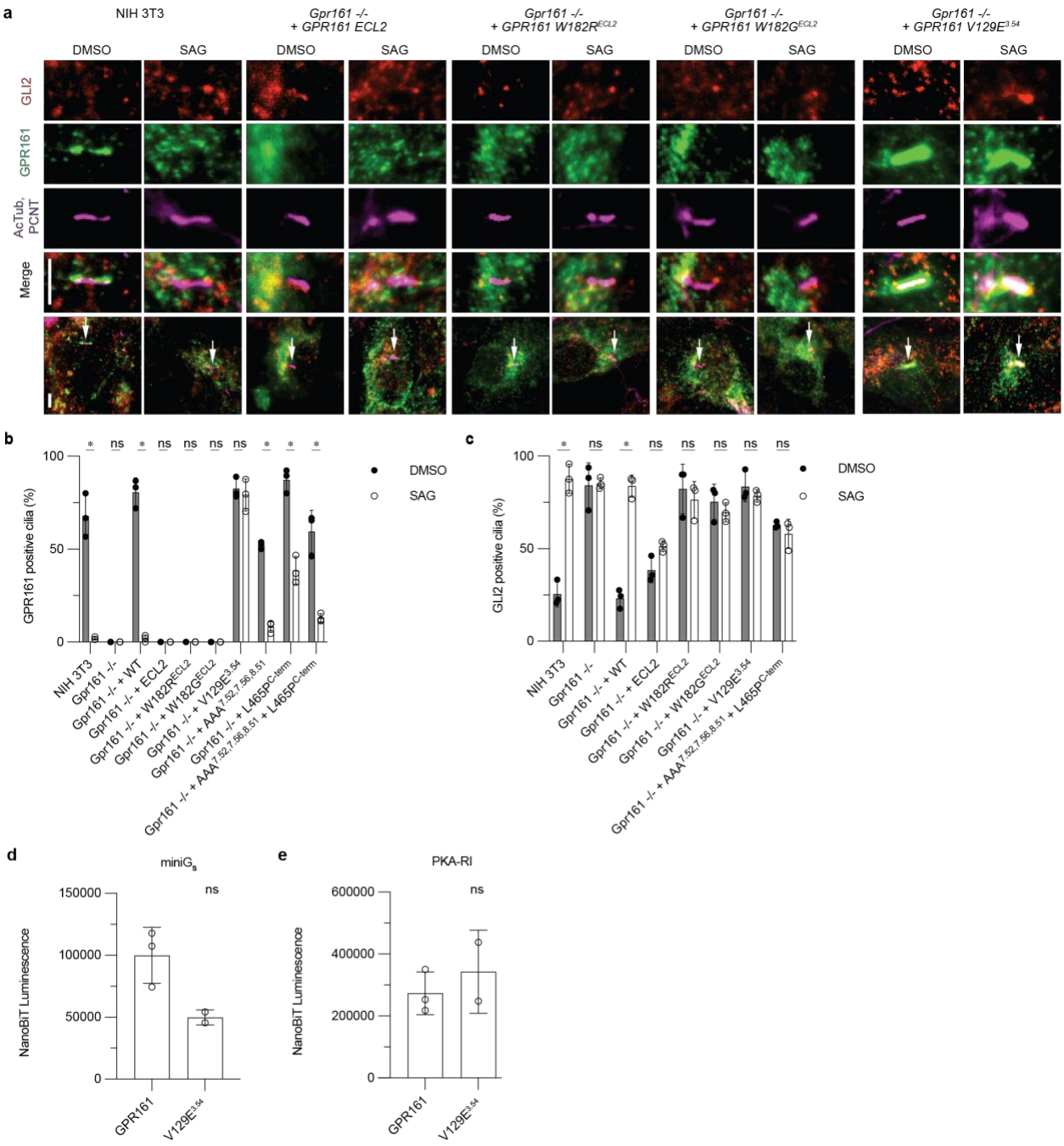


Figure 1.14 GPR161 localization and repression of ciliary trafficking of GLI2 for ECL2 mutants and V129E$^{3.54}$

a) Representative images of the effect of site-directed mutagenesis of GPR161 on ciliary localization and GLI2 repression in ciliary tips in NIH 3T3 cells for ECL2 mutants and V129E$^{3.54}$. NIH 3T3 Flp-In CRISPR based *Gpr161*$^{-/-}$ cells stably expressing untagged mouse wild-type or *GPR161* mutants were starved for 24 hr upon confluence and were treated for further 24 hr ± SAG (500 nM). After fixation, cells were immunostained with anti-GLI2 (red), anti-GPR161 (green), anti-acetylated, and centrosome (AcTub; PCNT purple) antibodies. Whole cell images with an arrow indicating imaged cilia. Scale bar, 5 µm. **b)** Quantification of GPR161 positive cilia indicating trafficking and egress of GPR161 from cilia in the pathway off and on state,

respectively. ECL2 mutants do not traffic to cilia suggesting impaired biogenesis. GPR161-V129E$^{3.54}$ does not egress from cilia following pathway activation and GPR161-L465P$^{C\text{-}term}$ has reduced egress compared to GPR161. All other mutants traffic similar to GPR161. Quantification of GPR161 positive cilia are shown from 3 independent experiments from images taken from 2-3 different regions/experiment and counting 15-30 cells/region. Data are mean ± s.d. (*$P < 0.05$; ns, not significant; two-way ANOVA followed by Šidák's multiple comparison tests). **c)** Quantification GLI2 positive cilia indicating Hedgehog pathway activation. ECL2 mutants and GPR161-V129E$^{3.52}$ do not rescue, similar to *Gpr161*$^{-/-}$. Quantification of GLI2 positive cilia are shown from 3 independent experiments from images taken from 2-3 different regions/experiment and counting 15-30 cells/region. Data are mean ± s.d. (*$P < 0.05$; ns, not significant; two-way ANOVA followed by Šidák's multiple comparison tests). **d)** GPR161-V129E$^{3.54}$ has reduced recruitment of miniG$_s$ compared to WT. Data are mean ± s.d., n=2-3 biologically independent samples (*$P < 0.05$; ns, not significant; one-way ANOVA followed by Dunnett's multiple comparison tests). **e)** GPR161-V129E$^{3.54}$ has similar recruitment of PKA-RI compared to GPR161. Nanoluc complementation assay for receptor recruitment of miniG$_s$. Both GPR161 and GPR161-L465P$^{C\text{-}term}$ constitutively recruit miniG$_s$ while GPR161-AAA$^{7.52, 7.56, 8.51}$ does not. Data are mean ± s.d., n=2-3 biologically independent samples (*$P < 0.05$; ns, not significant; one-way ANOVA followed by Dunnett's multiple comparison tests).

1.6 Tables

Table 1.1 Cryo-EM data collection and model statistics

Extended Data Table 1. Cryo-EM data collection, refinement, and validation statistics

	GPR161-G_s
EMDB: Full map	EMD-40603
RCSB PDB: Model	8SMV
Data collection	
Microscope	Thermo Scientific Krios G3i
Detector	Gatan K3 with Gatan BioQuantum Energy filter
Voltage (kV)	300
Magnification	105,000
Defocus range (μm)	-0.8 to -2.2
Pixel size, physical (Å)	0.86
Total exposure ($e^-/Å^2$)	50.7
Frame exposure ($e^-/Å^2$/frame)	0.845
Images, number of	8,294
Frames/image, number of	60
Initial particles, number of	9,760,777
Final particles, number of	335,928
Symmetry imposed	C1
Map sharpening, *B* factor ($Å^2$)	
Full map	-90
Map resolution, masked (Å)	
Full map	2.72
FSC threshold	0.143
Refinement	
Initial model used (AlphaFold code)	Q8N6U8
Model resolution (Å)	2.72
Model composition	
Chains	6
Non-hydrogen atoms	8,169
Protein residues	1,034
Ligands	1
B factors ($Å^2$)	
Receptor	45.0
Ligand	53.52
G protein with Nb35	21.58
R.m.s. deviations	
Bond length (Å)	0.004
Bond angles (°)	1.033
Validation	
MolProbity score	1.12
Clash score	1.85
EMRinger score	3.47
Rotamer outliers (%)	0.00
Ramachandran plot	
Favored (%)	96.96
Allowed (%)	3.04
Disallowed (%)	0.00

1.7 Materials and Methods

Expression and purification of GPR161 active-state complex

The human *GPR161* gene with an N-terminal influenza hemagglutinin signal sequence and FLAG epitope was cloned into a pcDNA3.1 vector with zeocin resistance and a tetracycline inducible cassette, as described previously[74]. The miniG_s399 protein[75] was fused to the C terminus of GPR161 preceded by a glycine/serine linker and rhinovirus 3C protease recognition site. The resulting fusion construct was transfected into inducible Expi293F-TetR cells (unauthenticated and untested for mycoplasma contamination, Thermo Fisher) using the ExpiFectamine transfection reagent per manufacturer instructions. After 18 h, protein expression was induced with 1 μg/mL doxycycline hyclate for 28 h before collection by centrifugation. Pelleted cells were washed with 50 mL phosphate buffered saline, pH 7.5 before storage at −80 °C.

For complex purification for cryo-EM, frozen cells were hypotonically lysed in a buffer comprised of 20 mM HEPES pH 7.5, 1 mM EDTA, 160 μg/mL benzamidine, 160 μg/mL leupeptin, and 100 μM TCEP for 10 min at 25˚C. The membrane fraction was collected by centrifugation, and the fusion protein was extracted with a buffer comprised of 20 mM HEPES, pH 7.5, 300 mM NaCl, 1% (w/v) lauryl maltose neopentyl glycol (L-MNG, Anatrace), 0.1% (w/v) cholesteryl hemisuccinate (CHS, Steraloids), 2 mM $MgCl_2$, 2 mM $CaCl_2$, 160 μg/mL benzamidine, 2 μg/mL leupeptin, and 100 μM TCEP with dounce homogenization and incubation with stirring for one hour at 4 °C. The soluble fraction was separated from the insoluble fraction by centrifugation and was applied to a column of homemade M1–FLAG antibody-conjugated Sepharose beads at a rate of 1 mL/min. Sepharose resin was then washed with ten column volumes of 20 mM HEPES pH 7.5, 300 mM NaCl, 0.1% (w/v) L-MNG, 0.01% (w/v) CHS, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 100 μM

TCEP and then washed with 10 column volumes of 20 mM HEPES pH 7.5, 150 mM NaCl, 0.0075% (w/v) L-MNG, 0.00075% (w/v) CHS, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 100 μM TCEP prior to elution with 20 mM HEPES pH 7.5, 150 mM NaCl, 0.0075% (w/v) L-MNG, 0.00075% (w/v) CHS, 100 μM TCEP, 5 mM EDTA, and 0.2 mg/mL FLAG peptide. The eluted GPR161-miniG_s fusion protein was concentrated in a 100 kDa MWCO Amicon spin concentrator, and injected onto a Superdex 200 Increase 10/300GL (Cytiva) gel filtration column equilibrated in 20 mM HEPES pH 7.5, 150 mM NaCl, 100 μM TCEP, 0.0075% (w/v) L-MNG, 0.0025% glyco-diosgenin (GDN, Anatrace), and 0.0005% CHS. Monodisperse fractions of GPR161-miniG_s were complexed with $G_{\beta 1\gamma 2}$ heterodimer and Nb35 (purified as described previously [76]) at 2-fold molar excess overnight at 4˚C. The next day, the complex was concentrated with a 100 kDa MWCO spin concentrator and excess $G_{\beta 1\gamma 2}$ and Nb35 was removed via size-exclusion chromatography, using a Superdex200 Increase 10/300 GL column (GE Healthcare) equilibrated in 20 mM HEPES pH 7.5, 150 mM NaCl, 100 μM TCEP, 0.0075% (w/v) L-MNG, 0.0025% (w/v) GDN, and 0.00075% CHS. The resulting complex was concentrated to 2.9 mg/mL with a 100 kDa MWCO spin concentrator for preparation of cryo-EM grids.

Two separate preparations of GPR161-miniG_s were made for biochemical experiments that deviated slightly from the purification protocol for cryoEM. For cholesterol photolabeling experiments, GPR161-miniG_s was expressed and purified using the above protocol except CHS and GDN was excluded from all buffers. For ^{3}H-cholesterol binding experiments, n-Dodecyl-***β***-D-Maltopyranoside (DDM, Anatrace) was used in lieu of LMNG at a final concentration of 0.02% and CHS and GDN were excluded from all buffers. The resulting size exclusion chromatography-purified protein samples were flash frozen in liquid nitrogen for downstream assay use.

Cryo-EM vitrification, data collection, and processing

The GPR161-G_s-Nb35 complex was concentrated to 3 mg/mL and 3 µl was applied onto a glow-discharged 300 mesh 1.2/1.3 gold grid covered in a holey carbon film (Quantifoil). Excess sample was removed with a blotting time of 3 s and a blotting force of 1 at 4 °C prior to plunge freezing into liquid ethane using a Vitrobot Mark IV (Thermo Fisher). A total of 8,294 movies were recorded with a K3 detector (Gatan) on a Titan Krios (Thermo Fisher) microscope operated at 300 keV with a BioQuantum post-column energy filter set to a zero-loss energy selection slit width set of 20 eV. The 60-frame movies were recorded for 2.6 s at a physical pixel size of 0.86 Å per pixel and a defocus range of −0.8 to − 2.2 µm for a total dose of 50.7 e−/Å2. Exposure areas were acquired with an automated image shift collection strategy using EPU (Thermo Fisher).

Movies were motion-corrected and dose-fractionated on-the-fly during data collection using UCSF MotionCor2[75]. Corrected micrographs were imported into cryoSPARC v3.1 (Structura Biotechnology)[76] for CTF estimation via the Patch CTF Estimation job type. Micrographs with a CTF fit resolution of > 5 Å were excluded from further processing. Templates for particle picking were generated using a 20 Å low-pass filtered model generated from an *ab initio* model made from blob-picked and 2D classified particles. Template picking yielded 9,760,777 particles, which were extracted in a 288-pixel box and Fourier cropped to 72 pixels. Particles were classified in 3D with alignment using the 20 Å low-pass filtered *ab initio* model and three "random" reconstructions generated from a prematurely terminated *ab initio* reconstruction job, called "garbage collectors," with the Heterogeneous Refinement job type. Two rounds of Heterogeneous Refinement yielded 3,033,326 particles that were re-extracted in the same box sized cropped to 144 pixels. Additional

Heterogeneous refinement and extraction without binning yielded 1,143,566 particles that were refined using the Non-Uniform Refinement job type. Particles were exported using csparc2star.py from the pyem script package[77], and a mask covering the 7TM domain of GPR161 was generated using the Segger tool in UCSF ChimeraX[78] and the mask.py script in pyem. The particles and mask were imported into Relion v3.0[79] and classified in 3D without alignment. Particles comprising the three highest resolution classes were reimported into cryoSPARC for Non-Uniform Refinement. Finally, particles were exported into cisTEM[80] for two local refinements using the Manual Refinement job type and low-pass filtering outside of masks. In the first local refinement, the previous 7TM mask was used, and the second local refinement used a full-particle mask.

Model building and refinement

Model building and refinement began with the Alphafold2[81] predicted structure as the starting model, which was fitted into the experimental cryoEM map using UCSF ChimeraX. The model was iteratively refined with real space refinement in Phenix[82] and manually in Coot[83]. The cholesterol model and rotamer library were generated with the PRODRG server[84], docked using Coot, and refined in Phenix. Final map-model validations were carried out using Molprobity and EMRinger in Phenix.

cAMP signaling assays

We measured cAMP production to determine activation of G_s signaling by GPR161. For each GPR161 construct (wild-type (WT), M177R, V179R, W182G, W182R, W327A, W327R, R332A, AA, AAA, L465P), a 2 mL suspension culture of Expi293F-TetR cells was co-transfected with a 3:1 ratioof a pcDNA3.1 plasmid expressing GPR161 and a luciferase-based cAMP biosensor,

pGlosensor-22F (Promega). Surface expression levels of constructs were titrated to similar levels with doxycycline and measured by flow cytometry using an Alexa-647 conjugated anti-M1 Flag antibody. Cells were collected 24 h post-induction, resuspended in Expi293 expression media (Gibco) supplemented with 10% DMSO, and gradually frozen to -80 ˚C in a Mr. Frosty Freezing container for future use. To perform the assay, frozen cells were rapidly thawed in a 37˚C water bath and resuspended in fresh Expi293 expression medium. Cells were diluted to a final concentration of 1e6 cells per mL in Expi293 expression medium plus 2% (v/v) Glosensor assay reagent (Promega) and intubated for 75 min at room temperature with gentle rotation. Cells were then plated in a white 384-well plate (Greiner) to a final density of 15,000 cells per well. Immediately after cell addition, luminescence was measured using a CLARIOstar instrument. Statistical analyses were performed using one-way ANOVA followed by Dunnett's multiple comparison tests between all possible pairs using GraphPad Prism 9 (Dotmatics).

^{3}H-Cholesterol binding assay

To measure cholesterol binding to GPR161, we developed a scintillation proximity assay (SPA) using purified receptor and ^{3}H-cholesterol (PerkinElmer). To capture M1-FLAG tagged receptor, we used Protein A coated beads and purified M1-FLAG antibody. Each binding reaction was performed in a final volume of 100 µL in a binding buffer comprised of 0.01% dodecylmaltoside, 20 mM HEPES pH 7.5, 150 mM NaCl, 100 nM purified M1-FLAG antibody, and 2 mM $CaCl_2$ (to enable M1-FLAG tag binding to antibody). SPA beads were added to a final concentration of 0.675 mg/mL, ^{3}H-cholesterol was added to 100 nM, and 100 nM of purified GPR161-miniG_s or GPR161-AAA$^{7.52, 7.56, 8.51}$-miniG_s was added to start the reaction. For competition with cold cholesterol a 3 µM final concentration was used. The reactions were incubated with shaking at room temperature

for 24 hrs, and bound ^{3}H-cholesterol was measured in a scintillation counter (Perkin Elmer). Statistical analyses were performed using two-way ANOVA followed by Dunnett's multiple comparison tests between all possible pairs using GraphPad Prism 9 (Dotmatics).

Photolabeling and MS analysis

Photolabeling reagents were synthesized as previously described[84] and stored in the dark at -20˚C as 10 mM stocks in ethanol. Aliquots of the photolabeling reagents (KK231 or LKM38) were air dried in the dark at room temperature and resolubilized in 20 mM HEPES pH 7.5, 150 mM NaCl, 0.0075% LMNG (no CHS or GDN) containing 20 µg of purified GPR161-miniG_s in a total volume of 50 µL. The protein was incubated with the photolabeling regent for one hour in the dark at 4 ˚C and then irradiated in a quartz cuvette with >320 nm UV light as previously described[85]. For site identification experiments, the photolabeling reagent concentration was 100 µM. For cholesterol competition experiments, aliquots of ethanolic stocks of the photolabeling reagent and cholesterol (10 mM stock) were added into the same tube and air dried prior to solubilization with GPR161-miniG_s. The final concentration of the photolabeling reagents was 3 µM and of cholesterol 100 µM.

For mass spectrometric analysis, the samples were desalted using Biospin 6 columns (BioRad, CA), equilibrated with 50 mM triethylammonium bicarbonate and 0.02%(w/v) n-Dodecyl-β-D-Maltoside. The proteins were sequentially reduced with 5 mM tris(2-carboxyethyl)phosphine (TCEP) for 30 mins, alkylated with 5 mM N-ethylmaleimide (NEM) in the dark for 45 mins, and quenched with 5 mM dithiothreitol (DTT) for 15 mins. These three steps were done at room temperature. The proteins were digested with 8 µg trypsin at 4 ˚C for one week at which time the

digestions were terminated by addition of formic acid (FA) to a final concentration of 1%.

The resultant peptides were analyzed with an OrbiTrap ELITE mass spectrometer (Thermo Fisher Scientific). Briefly, 15 μL samples were injected by an autosampler (UltiMate 3000 UHPLC system; ThermoFisher) onto a home-packed polymeric reverse phase PLRP-S (Agilent, Santa Clara, CA) column (10 cm × 75 μm, 300 Å) at a flow rate of 800 nL/min. A 10%-95% acetonitrile (ACN) gradient was applied for 150 minutes to separate peptides. Solvent A was 0.1% FA/water, and solvent B was 0.1% FA/ACN. The following gradient was applied: isocratic elution at 10% solvent B, 1–60 minutes; 10%–95% solvent B, 60–110 minutes; 95% solvent B, 110–140 minutes; 95%–10% solvent B, 140–145 minutes; 10% solvent B, 145–150 minutes. For the first 60 minutes, a built-in divert valve on the mass spectrometer was used to remove the hydrophilic contaminants from the mass spectrometer. Mass spectra (MS1) were acquired at high resolution (resolution of 60,000) in the range of m/z = 100-2,000. Top 20 ion precursors in MS1 were selected for MS2 using data-dependent acquisition with exclusion of singly charged precursors. Fragmentation was performed with high-energy dissociation (HCD) using a normalized collision energy of 35%. Product ion spectra (MS2) were acquired at a resolution of 15,000.

The data were searched against a database containing the sequence of GPR161-miniG_s using PEAKS Studio X pro (Bioinformatics Solutions Inc, Waterloo, ON, Canada) with the following settings: precursor ions mass accuracy of 20 ppm, fragmentation ion accuracy of 0.1 Da, up to three missed cleavages on either side of peptide with trypsin digestion; methionine oxidation, cysteine alkylation with NEM and DTT, any amino acids with adduct of LKM38 (mass = 396.34) and KK231 (mass = 484.26) were included as variable modifications. The searched results were

filtered with a 1% false discovery rate and the detected peptides were confirmed by manual analysis for monoisotopic mass accuracy and retention time with Xcalibur 2.2 (ThermoFisher). Fragment ions were also manually confirmed and were accepted based on the presence of a monoisotopic mass within 20 ppm mass accuracy. Photolabeling efficiency was estimated by generating selected ion chromatograms (SIC) of both unlabeled and photolabeled peptides, determining the area under the curve and calculating efficiency as: labeled peptide / (unlabeled peptide + labeled peptide). Statistical significance was analyzed with Student's paired t-test using GraphPad Prism 9 (Dotmatics).

NanoBiT recruitment assays

We measured protein recruitment to determine functionality of GPR161 mutants. For each GPR161 construct (wild-type (WT), AAA, L465P, V129E), as well as for β_2AR and GPR52, a 2-mL suspension culture of Expi293F-TetR cells was co-transfected with a 500 ng of a pcDNA3.1 plasmid expressing receptor fused C-terminally to smBiT and 100 ng of transducer fused to lgBiT (miniG_s: N-terminally, PKA-RI: C-terminally). Surface expression levels of constructs were titrated to similar levels with doxycycline and measured by flow cytometry using an Alexa-647 conjugated anti-M1 FLAG antibody. After 24 h of induction, cells were centrifuged at 300x*g*, and resuspended in DPBS at a concentration of ~55,000 cells per 200 µL. 40 µL of 30 µM coelenterazine-h diluted in PBS was added to cells for a final concentration of 5 µM. Cells were incubated for ~30 min at room temperature with gentle shaking. Luminescence was measured using a CLARIOstar instrument. Statistical analyses were performed using one-way ANOVA followed by Dunnett's multiple comparison tests between all possible pairs using GraphPad Prism 9 (Dotmatics).

Molecular dynamics simulations

Simulation setup

We performed simulations of GPR161 with cholesterol under two conditions: A) GPR161 unrestrained (6 independent simulations, roughly 1 μs each) B) GPR161 restrained to its G protein—bound conformation (6 independent simulations, roughly 1 μs each). For all conditions, the initial structures were based on the cryoEM structure reported [MOU1] in this paper and were prepared using Maestro (Schrödinger, LLC). In both conditions, the nanobody and G protein were removed from the structure.

Missing amino acid side chains were modeled using Prime (Schrödinger, LLC). Neutral acetyl and methylamide groups were added to cap the N- and C-termini, respectively, of the protein chains. Titratable residues were kept in their dominant protonation state at pH 7.4. Histidine residues were modeled as neutral, with a hydrogen atom bound to either the delta or epsilon nitrogen depending on which tautomeric state optimized the local hydrogen-bonding network. Dowser [86] was used to add water molecules to protein cavities. GPR161 was aligned on the receptor in the crystal structure of Prostaglandin E2 receptor EP2 subtype (PDB ID: 7CX4) [87] in the Orientation of Proteins in Membranes (OPM) database [88]. The aligned structures were inserted into a pre-equilibrated palmitoyloleoyl-phosphatidylcholine (POPC) membrane bilayer using Dabble [89]. Sodium and chloride ions were added to neutralize each system at a concentration of 150 mM. To simulate the G protein–bound conformation in condition B, 0.5 kcal·mol^{-1} ·Å^{-2} restraints were applied throughout the production simulation on non-hydrogen atoms of GPR161 residues that are within 4 Å of the G protein in the experimentally determined structure. These residues were: 125,

128, 129, 130, 131, 132, 133, 135, 136, 137, 211, 214, 215, 218, 219, 267, 268, 271, 272, and 327. Both final systems consist of 52716 atoms, including 140 lipid molecules and 9810 water molecules (initial system dimensions: 85 Å x 80 Å x 82 Å).

Simulation protocols

For each simulation, initial atom velocities were assigned randomly and independently. We employed the CHARMM36m force field for protein molecules, the CHARMM36 parameter set for lipid molecules and salt ions, and the associated CHARMM TIP3P model for water [90,91]. Simulations were run using the AMBER20 software [92] under periodic boundary conditions with the Compute Unified Device Architecture (CUDA) version of Particle-Mesh Ewald Molecular Dynamics (PMEMD) [93] on one GPU.

The systems were first heated over 12.5 ps from 0 K to 100 K in the NVT ensemble using a Langevin thermostat with harmonic restraints of 10.0 kcal·mol^{-1} ·Å^{-2} on the non-hydrogen atoms of the lipids, protein, and cholesterol. Initial velocities were sampled from a Boltzmann distribution. The systems were then heated to 310 K over 125 ps in the NPT ensemble. Equilibration was performed at 310 K and 1 bar in the NPT ensemble, with harmonic restraints on the protein and cholesterol non-hydrogen atoms tapered off by 1.0 kcal·mol^{-1} ·Å^{-2} starting at 5.0 kcal·mol^{-1} ·Å^{-2} in a stepwise manner every 2 ns for 10 ns, and finally by 0.1 kcal·mol^{-1} ·Å^{-2} every 2 ns for an additional 18 ns. Except for the restrained residues listed above in condition B, all restraints were completely removed during production simulation. Production simulations were performed at 310 K and 1 bar in the NPT ensemble using the Langevin thermostat and Berendsen barostat.

Lengths of bonds to hydrogen atoms were constrained using SHAKE, and the simulations were performed using a timestep of 4.0 fs while using hydrogen mass repartitioning [94]. Non-bonded interactions were cut off at 9.0 Å, and long-range electrostatic interactions were calculated using the particle-mesh Ewald (PME) method with an Ewald coefficient (β) of approximately 0.31 Å and B-spline interpolation of order 4. The PME grid size was chosen such that the width of a grid cell was approximately 1 Å. Snapshots of the trajectory were saved every 200 ps.

Simulation analysis protocols

The AmberTools17 CPPTRAJ package [95] was used to reimage trajectories at 1 ns per frame, Visual Molecular Dynamics (VMD) [96] was used for visualization and analysis, and PyMOL (The PyMOL Molecular Graphics System, Schrödinger, LLC) was used for renderings.

Plots of time traces from representative simulations were generated with Matplotlib [97] and show both original, unsmoothed traces (transparent lines) and traces smoothed with a moving average (thick lines), using an averaging window of 20 ns. All traces include the initial equilibration with harmonic restraints on the protein and cholesterol non-hydrogen atoms.

To monitor ECL2 movement, we measure the minimal distance between all atoms of W182 and T189 (Fig. 1.2f,g and Fig. 1.13b). To capture cholesterol motion, we measure the minimal distance between all non-hydrogen atoms between cholesterol and W327. We show a representative trace of this measurement for each condition (Fig. 1.4b and Fig. 1.13c,d).

Ciliary localization and Hedgehog pathway activation

Cell lines

NIH 3T3-FlpIn cells were authenticated by and purchased from Thermo Fisher Scientific. They have tested negative for Mycoplasma. The *Gpr161*$^{-/-}$ NIH 3T3 Flp-In cell line was a gift from Rajat Rohatgi[96]. The cells were cultured in DMEM-high glucose media (D5796; Sigma) with 10% BCS (Sigma-Aldrich), 0.05 mg/ml penicillin, 0.05 mg/ml streptomycin, and 4.5 mM glutamine. Stable knockout cell lines were generated by retroviral infection with pBABE constructs having untagged wild type or mutant *GPR161* inserts followed by antibiotic selection. Single or multiple amino acid mutations in full-length *GPR161* were generated using Q5 site-directed mutagenesis kit (NEB).

Immunostaining and microscopy

For immunofluorescence experiments in cell lines, cells were cultured on coverslips until confluent and starved for 48 h. To quantify ciliary GLI2 and GPR161 levels, cells were treated with 500 nM SAG or DMSO for 24 h after 24 h of serum starvation. Cells were fixed with 4% PFA for 10 min at room temperature. After blocking with 5% normal donkey serum, the cells were incubated with primary antibody solutions for 1 h at room temperature followed by treatment with secondary antibodies for 30 min along with DAPI. Primary antibodies used were against GPR161 (1:200, custom-made)[97], acetylated α-tubulin (mAb 6-11B-1, Sigma; 1:2000), GLI2 (1:500, gift from Jonathan Eggenschwiler)[98], pericentrin (611814, BD Biosciences; 1:500). Coverslips were mounted with Fluoromount-G and images were acquired with a Zeiss AxioImager.Z1 microscope using a 40x oil immersion objective lens.

Quantification and statistical analysis

Cilia positive for GLI2 or GPR161 in *Gpr161*$^{-/-}$ cells expressing untagged wild-type or mutant *GPR161* were counted. Statistical analyses were performed using two-way ANOVA followed by Šidák's multiple comparison tests between all possible pairs using GraphPad Prism 9 (Dotmatics).

1.8 Acknowledgements

This work was supported by the National Institutes of Health (NIH) grants R01GM127359 (R.O.D.), R01GM108799 (A.S.E.), P50MH122379 (A.S.E.), R35GM144136 (S.M.) and R01GM138992 (R.O.D. and A.M.). Additional support was from NSF Graduate Research Fellowship (M.K.) and Human Frontier Science Program Long-Term Fellowship LT000916/2018-L (C.-M.S.). Cryo-EM equipment at UCSF is partially supported by NIH grants S10OD020054 and S10OD021741. Some of this work was performed at the Stanford-SLAC Cryo-EM Center (S2C2), which is supported by the National Institutes of Health Common Fund Transformative High-Resolution Cryo-Electron Microscopy program (U24 GM129541). The authors would also like to thank the following S2C2 personnel for their invaluable support and assistance: Corey Hecksel. The authors thank Jeremy Reiter for helpful suggestions on GPR161 homologous sequences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. A.M. acknowledges support from the Edward Mallinckrodt, Jr. Foundation and the Vallee Foundation. A.M. is a Chan Zuckerberg Biohub Investigator.

1.9 Author Contributions

N.H., S.H., and I.D. cloned, expressed, and biochemically optimized the purification of GPR161 complex constructs for structural studies. N.H., S.H., and I.D. performed cryo-EM data collection, with help from SLAC Cryo-EM Center, and data processing. N.H., S.H., I.D., and A.M. built and refined models of GPR161. N.H. and S.H. generated receptor constructs and determined expression levels by flow cytometry and performed signaling studies, complementation assays, and analyzed the data. M.K. and C.M.S. performed and analyzed molecular dynamics simulations

under the supervision of R.O.D. N.H. prepared samples for, performed, and analyzed scintillation proximity assay data with A.M. Z.C. performed and analyzed mass spectrometry experiments using reagents provided by D.C. under the supervision of A.E. S.H.H., V.R.P, and S.M. prepared constructs, performed, and analyzed GPR161 localization and Hedgehog pathway repression experiments. S.B. performed phylogenetic analysis under the supervision of D.M. P.T., D.R., and E.S. analyzed GPR161 variants. All authors contributed to figures. N.H., S.H., A.M., and S.M. wrote the manuscript, with edits and approval from all authors. A.M. supervised the overall project.

1.10 References

1. Dixon, R. A. *et al.* Cloning of the gene and cDNA for mammalian beta-adrenergic receptor and homology with rhodopsin. *Nature* **321**, 75–79 (1986).
2. Lodes, M. J., Dillon, D. C., Houghton, R. L. & Skeiky, Y. A. W. Expression cloning. *Methods Mol. Med.* **94**, 91–106 (2004).
3. Libert, F., Vassart, G. & Parmentier, M. Current developments in G-protein-coupled receptors. *Curr. Opin. Cell Biol.* **3**, 218–223 (1991).
4. Thorens, B. Expression cloning of the pancreatic beta cell receptor for the gluco-incretin hormone glucagon-like peptide 1. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 8641–8645 (1992).
5. Parmentier, M. *et al.* Molecular cloning of the thyrotropin receptor. *Science* **246**, 1620–1622 (1989).
6. Sakurai, T. *et al.* Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. *Cell* **92**, 573–585 (1998).
7. Laschet, C., Dupuis, N. & Hanson, J. The G protein-coupled receptors deorphanization landscape. *Biochem. Pharmacol.* **153**, 62–74 (2018).
8. Mukhopadhyay, S. *et al.* The Ciliary G-Protein-Coupled Receptor Gpr161 Negatively Regulates the Sonic Hedgehog Pathway via cAMP Signaling. *Cell* **152**, 210–223 (2013).
9. Kim, S.-E. *et al.* Dominant negative GPR161 rare variants are risk factors of human spina bifida. *Hum. Mol. Genet.* **28**, 200–208 (2019).
10. Begemann, M. *et al.* Germline GPR161 Mutations Predispose to Pediatric Medulloblastoma. *J. Clin. Oncol.* **38**, 43–50 (2020).
11. Feigin, M. E., Xue, B., Hammell, M. C. & Muthuswamy, S. K. G-protein–coupled receptor GPR161 is overexpressed in breast cancer and is a promoter of cell proliferation and invasion.

Proceedings of the National Academy of Sciences **111**, 4191–4196 (2014).

12. Civelli, O., Saito, Y., Wang, Z., Nothacker, H.-P. & Reinscheid, R. K. Orphan GPCRs and their ligands. *Pharmacol. Ther.* **110**, 525–532 (2006).
13. Hwang, S.-H., Somatilaka, B. N., White, K. & Mukhopadhyay, S. Ciliary and extraciliary Gpr161 pools repress hedgehog signaling in a tissue-specific manner. *Elife* **10**, (2021).
14. Shimada, I. S. *et al.* Basal Suppression of the Sonic Hedgehog Pathway by the G-Protein-Coupled Receptor Gpr161 Restricts Medulloblastoma Pathogenesis. *Cell Rep.* **22**, 1169–1184 (2018).
15. Hwang, S.-H. *et al.* The G protein-coupled receptor Gpr161 regulates forelimb formation, limb patterning and skeletal morphogenesis in a primary cilium-dependent manner. *Development* **145**, (2018).
16. Kim, S.-E. *et al.* Wnt1 Lineage Specific Deletion of Gpr161 Results in Embryonic Midbrain Malformation and Failure of Craniofacial Skeletal Development. *Front. Genet.* **12**, 761418 (2021).
17. Shimada, I. S. *et al.* Derepression of sonic hedgehog signaling upon Gpr161 deletion unravels forebrain and ventricular abnormalities. *Dev. Biol.* **450**, 47–62 (2019).
18. Li, B. I. *et al.* The orphan GPCR, Gpr161, regulates the retinoic acid and canonical Wnt pathways during neurulation. *Dev. Biol.* **402**, 17–31 (2015).
19. Matteson, P. G. *et al.* The orphan G protein-coupled receptor, Gpr161, encodes the vacuolated lens locus and controls neurulation and lens development. *Proceedings of the National Academy of Sciences* **105**, 2088–2093 (2008).
20. Karaca, E. *et al.* Whole-exome sequencing identifies homozygous GPR161 mutation in a family with pituitary stalk interruption syndrome. *J. Clin. Endocrinol. Metab.* **100**, (2015).

21. Website. https://bpspubs.onlinelibrary.wiley.com/doi/10.1111/bph.16053.

22. McMahon, A. P., Ingham, P. W. & Tabin, C. J. 1 Developmental roles and clinical significance of Hedgehog signaling. in *Current Topics in Developmental Biology* vol. 53 1–114 (Academic Press, 2003).

23. Kopinke, D., Norris, A. M. & Mukhopadhyay, S. Developmental and regenerative paradigms of cilia regulated hedgehog signaling. *Semin. Cell Dev. Biol.* **110**, 89–103 (2021).

24. Truong, M. E. *et al.* Vertebrate cells differentially interpret ciliary and extraciliary cAMP. *Cell* **184**, 2911-2926.e18 (2021).

25. Hilgendorf, K. I., Johnson, C. T. & Jackson, P. K. The primary cilium as a cellular receiver: organizing ciliary GPCR signaling. *Curr. Opin. Cell Biol.* **39**, 84–92 (2016).

26. Mukhopadhyay, S. & Rohatgi, R. G-protein-coupled receptors, Hedgehog signaling and primary cilia. *Semin. Cell Dev. Biol.* **33**, (2014).

27. Tschaikner, P. M. *et al.* Feedback control of the Gpr161-Gαs-PKA axis contributes to basal Hedgehog repression in zebrafish. *Development* **148**, dev192443 (2021).

28. Pal, K. *et al.* Smoothened determines β-arrestin–mediated removal of the G protein–coupled receptor Gpr161 from the primary cilium. *J. Cell Biol.* **212**, 861–875 (2016).

29. Kroeze, W. K. *et al.* PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nat. Struct. Mol. Biol.* **22**, 362–369 (2015).

30. Pusapati, G. V. *et al.* G protein–coupled receptors control the sensitivity of cells to the morphogen Sonic Hedgehog. *Sci. Signal.* **11**, eaao5749 (2018).

31. Foster, S. R. *et al.* Discovery of Human Signaling Systems: Pairing Peptides to G Protein-Coupled Receptors. *Cell* **179**, (2019).

32. Bachmann, V. A. *et al.* Gpr161 anchoring of PKA consolidates GPCR and cAMP signaling.

Proc. Natl. Acad. Sci. U. S. A. **113**, 7786–7791 (2016).

33. Nehmé, R. *et al.* Mini-G proteins: Novel tools for studying GPCRs in their active conformation. *PLoS One* **12**, e0175642 (2017).
34. 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. Sel.* **29**, 583–594 (2016).
35. Rasmussen, S. G. F. *et al.* Crystal structure of the β2 adrenergic receptor–Gs protein complex. *Nature* **477**, 549–555 (2011).
36. Zhou, Q. *et al.* Common activation mechanism of class A GPCRs. *Elife* **8**, (2019).
37. Lin, X. *et al.* Structural basis of ligand recognition and self-activation of orphan GPR52. *Nature* **579**, 152–157 (2020).
38. Ye, F. *et al.* Cryo-EM structure of G-protein-coupled receptor GPR17 in complex with inhibitory G protein. *MedComm (2020)* **3**, e159 (2022).
39. Wong, T.-S. *et al.* Cryo-EM structure of orphan G protein-coupled receptor GPR21. *MedComm (2020)* **4**, e205 (2023).
40. 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).
41. Dixon, A. S. *et al.* NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. *ACS Chem. Biol.* **11**, 400–408 (2016).
42. Copp, A. J. *et al.* Spina bifida. *Nat Rev Dis Primers* **1**, 1–18 (2015).
43. Eaton, S. Multiple roles for lipids in the Hedgehog signalling pathway. *Nat. Rev. Mol. Cell Biol.* **9**, 437–445 (2008).
44. Luchetti, G. *et al.* Cholesterol activates the G-protein coupled receptor Smoothened to

promote Hedgehog signaling. *Elife* **5**, (2016).

45. Cooper, M. K. *et al.* A defective response to Hedgehog signaling in disorders of cholesterol biosynthesis. *Nat. Genet.* **33**, 508–513 (2003).
46. Budelier, M. M. *et al.* Photoaffinity labeling with cholesterol analogues precisely maps a cholesterol-binding site in voltage-dependent anion channel-1. *J. Biol. Chem.* **292**, 9294–9304 (2017).
47. Krishnan, K. *et al.* Validation of Trifluoromethylphenyl Diazirine Cholesterol Analogues As Cholesterol Mimetics and Photolabeling Reagents. *ACS Chem. Biol.* **16**, 1493–1507 (2021).
48. Castellano, B. M. *et al.* Lysosomal cholesterol activates mTORC1 via an SLC38A9-Niemann-Pick C1 signaling complex. *Science* **355**, 1306–1311 (2017).
49. Shin, H. R. *et al.* Lysosomal GPCR-like protein LYCHOS signals cholesterol sufficiency to mTORC1. *Science* **377**, 1290–1298 (2022).
50. Chen, M.-H. *et al.* Cilium-independent regulation of Gli protein function by Sufu in Hedgehog signaling is evolutionarily conserved. *Genes Dev.* **23**, 1910–1928 (2009).
51. Haycraft, C. J. *et al.* Gli2 and Gli3 localize to cilia and require the intraflagellar transport protein polaris for processing and function. *PLoS Genet.* **1**, e53 (2005).
52. Kim, J., Kato, M. & Beachy, P. A. Gli2 trafficking links Hedgehog-dependent activation of Smoothened in the primary cilium to transcriptional activation in the nucleus. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 21666–21671 (2009).
53. Dessauer, C. W. Adenylyl cyclase--A-kinase anchoring protein complexes: the next dimension in cAMP signaling. *Mol. Pharmacol.* **76**, (2009).
54. Moore, B. S. *et al.* Cilia have high cAMP levels that are inhibited by Sonic Hedgehog-regulated calcium dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 13069–13074 (2016).

55. Somatilaka, B. N. *et al.* Ankmy2 Prevents Smoothened-Independent Hyperactivation of the Hedgehog Pathway via Cilia-Regulated Adenylyl Cyclase Signaling. *Dev. Cell* **54**, 710-726.e8 (2020).

56. Jiang, J. Y., Falcone, J. L., Curci, S. & Hofer, A. M. Direct visualization of cAMP signaling in primary cilia reveals up-regulation of ciliary GPCR activity following Hedgehog activation. *Proceedings of the National Academy of Sciences* **116**, 12066–12071 (2019).

57. Smith, F. D. *et al.* Local protein kinase A action proceeds through intact holoenzymes. *Science* **356**, 1288–1293 (2017).

58. May, E. A. *et al.* Time-resolved proteomics profiling of the ciliary Hedgehog response. *J. Cell Biol.* **220**, e202007207 (2021).

59. Calebiro, D. *et al.* Persistent cAMP-signals triggered by internalized G-protein-coupled receptors. *PLoS Biol.* **7**, (2009).

60. Crilly, S. E. & Puthenveedu, M. A. Compartmentalized GPCR Signaling from Intracellular Membranes. *J. Membr. Biol.* **254**, (2021).

61. Irannejad, R. *et al.* Conformational biosensors reveal GPCR signalling from endosomes. *Nature* **495**, (2013).

62. Vilardaga, J.-P., Jean-Alphonse, F. G. & Gardella, T. J. Endosomal generation of cAMP in GPCR signaling. *Nat. Chem. Biol.* **10**, 700–706 (2014).

63. Ogden, S. K. *et al.* G protein Gαi functions immediately downstream of Smoothened in Hedgehog signalling. *Nature* **456**, 967–970 (2008).

64. Ayers, K. L. & Thérond, P. P. Evaluating Smoothened as a G-protein-coupled receptor for Hedgehog signalling. *Trends Cell Biol.* **20**, 287–298 (2010).

65. Happ, J. T. *et al.* A PKA inhibitor motif within SMOOTHENED controls Hedgehog signal

transduction. *Nat. Struct. Mol. Biol.* **29**, 990–999 (2022).

66. Stubbs, T., Bingman, J. I., Besse, J. & Mykytyn, K. Ciliary signaling proteins are mislocalized in the brains of Bardet-Biedl syndrome 1-null mice. *Frontiers in cell and developmental biology* **10**, (2023).
67. Badgandi, H. B., Hwang, S. H., Shimada, I. S., Loriot, E. & Mukhopadhyay, S. Tubby family proteins are adapters for ciliary trafficking of integral membrane proteins. *J. Cell Biol.* **216**, (2017).
68. Sheu, S. H. *et al.* A serotonergic axon-cilium synapse drives nuclear signaling to alter chromatin accessibility. *Cell* **185**, (2022).
69. Chou, C.-H. *et al.* Bisdemethoxycurcumin Promotes Apoptosis and Inhibits the Epithelial-Mesenchymal Transition through the Inhibition of the G-Protein-Coupled Receptor 161/Mammalian Target of Rapamycin Signaling Pathway in Triple Negative Breast Cancer Cells. *J. Agric. Food Chem.* **69**, 14557–14567 (2021).
70. Bock, A. *et al.* Optical Mapping of cAMP Signaling at the Nanometer Scale. *Cell* **182**, (2020).
71. Zhang, J. Z. *et al.* Phase Separation of a PKA Regulatory Subunit Controls cAMP Compartmentation and Oncogenic Signaling. *Cell* **182**, (2020).
72. Ring, A. M. *et al.* Adrenaline-activated structure of β2-adrenoceptor stabilized by an engineered nanobody. *Nature* **502**, 575–579 (2013).
73. Staus, D. P. *et al.* Sortase ligation enables homogeneous GPCR phosphorylation to reveal diversity in β-arrestin coupling. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3834–3839 (2018).
74. Faust, B. *et al.* Autoantibody mimicry of hormone action at the thyrotropin receptor. *Nature* **609**, 846–853 (2022).
75. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved

cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

76. 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).
77. Asarnow, D., Palovcak, E. & Cheng, Y. *asarnow/pyem: UCSF pyem v0.5*. (2019). doi:10.5281/zenodo.3576630.
78. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
79. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, (2018).
80. Grant, T., Rohou, A. & Grigorieff, N. cisTEM, user-friendly software for single-particle image processing. *Elife* **7**, (2018).
81. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
82. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
83. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
84. Schüttelkopf, A. W. & van Aalten, D. M. F. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 1355–1363 (2004).
85. Darbandi-Tonkabon, R. *et al.* Photoaffinity labeling with a neuroactive steroid analogue. 6-azi-pregnanolone labels voltage-dependent anion channel-1 in rat brain. *J. Biol. Chem.* **278**, 13196–13206 (2003).

86. Zhang, L. & Hermans, J. Hydrophilicity of cavities in proteins. *Proteins* **24**, 433–438 (1996).

87. Qu, C. *et al.* Ligand recognition, unconventional activation, and G protein coupling of the prostaglandin E_2 receptor EP2 subtype. *Sci Adv* **7**, (2021).

88. Lomize, M. A., Lomize, A. L., Pogozheva, I. D. & Mosberg, H. I. OPM: orientations of proteins in membranes database. *Bioinformatics* **22**, 623–625 (2006).

89. Betz, R. *Dabble*. (2017). doi:10.5281/zenodo.836914.

90. Huang, J. *et al.* CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).

91. Klauda, J. B. *et al.* Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).

92. D.A. Case, H.M. Aktulga, K. Belfon, I.Y. Ben-Shalom, J.T. Berryman, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, G.A. Cisneros, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, K. Kasavajhala, M.C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, A. Shajan, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, J. Wang, H. Wei, R.M. Wolf, X. Wu, Y. Xiong, Y. Xue, D.M. York, S. Zhao, and P.A. Kollman. Amber21. *University of California, San Francisco* (2022).

93. Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S. & Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **9**, 3878–3888 (2013).

94. Hopkins, C. W., Le Grand, S., Walker, R. C. & Roitberg, A. E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **11**, 1864–1874 (2015).

95. Roe, D. R. & Cheatham, T. E., 3rd. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).

96. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–8, 27–8 (1996).

97. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).

98. Norman, R. X. *et al.* Tubby-like protein 3 (TULP3) regulates patterning in the mouse embryo through inhibition of Hedgehog signaling. *Hum. Mol. Genet.* **18**, 1740–1754 (2009).

99. Reeh, P. W. & Steen, K. H. Chapter 8. Tissue acidosis in nociception and pain. in *Progress in Brain Research* (eds. Kumazawa, T., Kruger, L. & Mizumura, K.) vol. 113 143–151 (Elsevier, 1996).

100. Jordt, S. E., Tominaga, M. & Julius, D. Acid potentiation of the capsaicin receptor determined by a key extracellular site. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 8134–8139 (2000).

101. Jasti, J., Furukawa, H., Gonzales, E. B. & Gouaux, E. Structure of acid-sensing ion channel 1 at 1.9 A resolution and low pH. *Nature* **449**, 316–323 (2007).

102. Yang, J. *et al.* PAC, an evolutionarily conserved membrane protein, is a proton-activated chloride channel. *Science* **364**, 395–399 (2019).

103. Rahman, M. F., Askwith, C. & Govindarajan, R. Molecular determinants of acidic pH-dependent transport of human equilibrative nucleoside transporter 3. *J. Biol. Chem.* **292**, 14775–14785 (2017).

104. Li, B., Rietmeijer, R. A. & Brohawn, S. G. Structural basis for pH gating of the two-pore domain K+ channel TASK2. *Nature* **586**, 457–462 (2020).

105. Wang, L., Hall, C., Li, J., Choi, E. & Bai, X.-C. Structural basis of the alkaline pH-dependent activation of insulin receptor-related receptor. *Nat. Struct. Mol. Biol.* (2023) doi:10.1038/s41594-023-00974-0.

106. Choi, J. W., Lee, S. Y. & Choi, Y. Identification of a putative G protein-coupled receptor induced during activation-induced apoptosis of T cells. *Cell. Immunol.* **168**, 78–84 (1996).

107. Ludwig, M.-G. *et al.* Proton-sensing G-protein-coupled receptors. *Nature* **425**, 93–98 (2003).

108. Kumar, N. N. *et al.* PHYSIOLOGY. Regulation of breathing by CO_2 requires the proton-activated receptor GPR4 in retrotrapezoid nucleus neurons. *Science* **348**, 1255–1260 (2015).

109. Lassen, K. G. *et al.* Genetic Coding Variant in GPR65 Alters Lysosomal pH and Links Lysosomal Dysfunction with Colitis Risk. *Immunity* **44**, 1392–1405 (2016).

110. Vullo, S. *et al.* Conformational dynamics and role of the acidic pocket in ASIC pH-dependent gating. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 3768–3773 (2017).

111. Ruan, Z., Osei-Owusu, J., Du, J., Qiu, Z. & Lü, W. Structures and pH-sensing mechanism of the proton-activated chloride channel. *Nature* **588**, 350–354 (2020).

112. Zhang, K., Julius, D. & Cheng, Y. Structural snapshots of TRPV1 reveal mechanism of polymodal functionality. *Cell* **184**, 5138-5150.e12 (2021).

113. Wang, S., Yan, R., Zhang, X., Chu, Q. & Shi, Y. Molecular mechanism of pH-dependent substrate transport by an arginine-agmatine antiporter. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 12734–12739 (2014).

114. Thomaston, J. L. *et al.* High-resolution structures of the M2 channel from influenza A virus reveal dynamic pathways for proton stabilization and transduction. *Proc. Natl. Acad. Sci. U.*

S. A. **112**, 14260–14265 (2015).

115. Liang, R. *et al.* Acid activation mechanism of the influenza A M2 proton channel. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E6955–E6964 (2016).

116. Robert, R. & Mackay, C. R. Gαs-coupled GPCRs GPR65 and GPR174. Downers for immune responses. *Immunology and cell biology* vol. 96 341–343 (2018).

117. Wirasinha, R. C. *et al.* GPR65 inhibits experimental autoimmune encephalomyelitis through CD4+ T cell independent mechanisms that include effects on iNKT cells. *Immunol. Cell Biol.* **96**, 128–136 (2018).

118. Xu, J. *et al.* GPR68 Senses Flow and Is Essential for Vascular Physiology. *Cell* **173**, 762-775.e16 (2018).

119. Wang, J.-Q. *et al.* TDAG8 Is a Proton-sensing and Psychosine-sensitive G-protein-coupled Receptor*. *J. Biol. Chem.* **279**, 45626–45633 (2004).

120. Huang, X.-P., Kenakin, T. P., Gu, S., Shoichet, B. K. & Roth, B. L. Differential Roles of Extracellular Histidine Residues of GPR68 for Proton-Sensing and Allosteric Modulation by Divalent Metal Ions. *Biochemistry* **59**, 3594–3614 (2020).

121. Rowe, J. B., Kapolka, N. J., Taghon, G. J., Morgan, W. M. & Isom, D. G. The evolution and mechanism of GPCR proton sensing. *J. Biol. Chem.* **296**, 100167 (2020).

122. Fowler, D. M. & Fields, S. Deep mutational scanning: a new style of protein science. *Nat. Methods* **11**, 801–807 (2014).

123. Jones, E. M. *et al.* Structural and functional characterization of G protein-coupled receptors with deep mutational scanning. *Elife* **9**, (2020).

124. Macdonald, C. B. *et al.* DIMPLE: deep insertion, deletion, and missense mutation libraries for exploring protein variation in evolution, disease, and biology. *Genome Biol.* **24**, 36 (2023).

125. Coyote-Maestas, W. *et al.* Probing ion channel functional architecture and domain recombination compatibility by massively parallel domain insertion profiling. *Nat. Commun.* **12**, 7114 (2021).

126. Coyote-Maestas, W., Nedrud, D., He, Y. & Schmidt, D. Determinants of trafficking, conduction, and disease within a K+ channel revealed through multiparametric deep mutational scanning. *Elife* **11**, (2022).

127. Huang, X.-P. *et al.* Allosteric ligands for the pharmacologically dark receptors GPR68 and GPR65. *Nature* **527**, 477–483 (2015).

128. Yu, X. *et al.* Design, Synthesis, and Characterization of Ogerin-Based Positive Allosteric Modulators for G Protein-Coupled Receptor 68 (GPR68). *J. Med. Chem.* **62**, 7557–7574 (2019).

129. Harris, J. A. *et al.* Selective G protein signaling driven by substance P-neurokinin receptor dynamics. *Nat. Chem. Biol.* **18**, 109–115 (2022).

130. Koehl, A. *et al.* Structure of the μ-opioid receptor-Gi protein complex. *Nature* **558**, 547–552 (2018).

131. Hong, C. *et al.* Structures of active-state orexin receptor 2 rationalize peptide and small-molecule agonist recognition and receptor activation. *Nat. Commun.* **12**, 815 (2021).

132. Hoppe, N. *et al.* GPR161 structure uncovers the redundant role of sterol-regulated ciliary cAMP signaling in the Hedgehog pathway. *bioRxiv* 2023.05.23.540554 (2023) doi:10.1101/2023.05.23.540554.

133. Thal, D. M., Glukhova, A., Sexton, P. M. & Christopoulos, A. Structural insights into G-protein-coupled receptor allostery. *Nature* **559**, 45–53 (2018).

134. Honig, B. H. & Hubbell, W. L. Stability of "salt bridges" in membrane proteins. *Proc. Natl.*

Acad. Sci. U. S. A. **81**, 5412–5416 (1984).

135. Gao, J., Bosco, D. A., Powers, E. T. & Kelly, J. W. Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins. *Nat. Struct. Mol. Biol.* **16**, 684–690 (2009).

136. Zhuang, Y. *et al.* Molecular recognition of morphine and fentanyl by the human μ-opioid receptor. *Cell* **185**, 4361-4375.e19 (2022).

137. Robertson, M. J., Meyerowitz, J. G., Panova, O., Borrelli, K. & Skiniotis, G. Plasticity in ligand recognition at somatostatin receptors. *Nat. Struct. Mol. Biol.* **29**, 210–217 (2022).

Chapter 2

Structural basis of distributed proton sensing by GPCRs

Contributing Authors

Nicholas Hoppe[1,2]*, Matthew Howard[1,2]*, Xi-Ping Huang[4], Willow Coyote-Maestas[5,6#], Aashish Manglik[1,7,8#]

1. Department of Pharmaceutical Chemistry, University of California, San Francisco, CA, USA
2. Biophysics Graduate Program, University of California, San Francisco, CA, USA
3. Tetrad Graduate Program, University of California, San Francisco, CA, USA
4. Department of Pharmacology and the National Institute of Mental Health Psychoactive Drug Screening Program (NIMH PDSP), The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
5. Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, USA
6. Quantitative Biosciences Institute, University of California, San Francisco, USA
7. Department of Anesthesia and Perioperative Care, University of California, San Francisco, CA, USA
8. Chan Zuckerberg Biohub, San Francisco, CA, USA

*These authors contributed equally

#Correspondance to: Willow Coyote-Maestas (willow.coyote-maestas@ucsf.edu) or Aashish Manglik (aashish.manglik@ucsf.edu)

2.1 Abstract

Proton-sensing G protein-coupled receptors (GPCRs) respond to changes in extracellular pH to regulate diverse physiological processes. The three proton-sensing GPCRs, GPR4, GPR65, and GPR68, regulate breathing, inflammation, and vasodilation, respectively, through an unknown mechanism. In this study, we provide a detailed understanding of GPCR proton sensing, which differs from the mechanism described for other proton-sensing membrane proteins. Using a combination of deep mutational scanning and cryogenic electron microscopy, we identify a distributed network of residues across the entire extracellular half of the proton-sensing GPCRs that contribute to activation. This network is composed of a variety of residues other than histidines, including aspartates, glutamates, arginines, and tyrosines, which were previously underappreciated. Our findings demonstrate that proton sensing in GPCRs is not localized to a single site but is distributed across a network of residues. This provides a new framework for understanding proton-sensing membrane proteins and enables the development of novel modulators targeting these receptors.

2.2 Introduction

The precise regulation of pH is a fundamental aspect of human physiological systems. Different classes of membrane proteins, including ion channels, transporters, and G protein-coupled receptors (GPCRs), have evolved to sense and respond to changes in extracellular pH. These membrane proteins have critical roles in the homeostasis of our nervous, cardiovascular, and metabolic systems and specifically sense pain, hypoxia, acidosis, and inflammation [99–109]. Although these proteins control vital physiological processes, we are just beginning to understand the structural details underpinning their function.

Structures of proton-sensing membrane proteins have revealed a potential general mechanism of proton sensing. Prior studies have demonstrated that membrane proteins sense protons at highly localized, singular sites that mediate the response to changes in pH. We call this the hotspot model of proton sensing. This has been convincingly seen and demonstrated in channels, transporters, and receptor tyrosine kinases, including acid sensing ion channels, two-pore domain K^+ channels, proton-active chloride channels, TRP channels, equilibrative nucleoside transporters, and the insulin receptor-related receptor [100,101,103–105,110–115]. However, we do not know if the hotspot model extends to proton-sensing GPCRs.

There are three proton-sensing GPCRs, GPR4, GPR65, and GPR68. GPR4 regulates CO_2-dependent breathing in the retrotrapezoid nucleus [108]. GPR65 suppresses inflammation and autoimmune diseases [109,116,117]. GPR68 modulates small-diameter artery dilation and remodeling [118]. Since their initial characterization, the proton-sensing GPCRs have been proposed to sense protons through a handful of histidines across their extracellular surface [107,119]. This "distributed" model presents an alternative to the hotspot model whereby protons act at multiple, spatially distinct sites to give rise to activation (Fig. 2.1a).

Although conceptually simple, evidence for a distributed proton-sensing network of histidines has been elusive. Previous mutational studies have shown that these extracellular histidines are not necessary for pH sensing [120,121]. For example, these can be mutated out of GPR68 and the receptor is still activated by pH, albeit at a lower pH_{50} [120]. Also, a case has been made that proton-sensing GPCRs do act via the hotspot model due to a triad of glutamines [121]. Despite the extensive

mutagenesis published to date, the molecular mechanism of proton sensing by these three GPCRs remains unresolved. There could be an undiscovered hotspot controlling activation or there could be a distributed sensing network that point mutations cannot capture.

To understand the mechanism of GPCR proton sensing, we need a comprehensive and unbiased examination of how proton sensing GPCRs respond to pH. Here we combine deep mutational scanning (DMS) and cryogenic electron microscopy (cryo-EM) to determine the proton sensing mechanism for these GPCRs. Our findings reveal a distributed network in which nearly all extracellular titratable residues contribute to the emergent property of proton sensing, sometimes in unexpected ways. This provides a new framework for understanding proton-sensing membrane proteins.

2.3 Results

Proton-sensing GPCRs sense protons through a distributed mechanism

Our initial investigations sought to understand if there is a specific hotspot that confers proton sensing for these GPCRs, similar to other proton-sensing membrane proteins. We reasoned that if there was a lynchpin residue or site that determines the pH response for these GPCRs, it could be identified through exchanging linear segments of one receptor with portions of another. We chose GPR4 and GPR68 for this chimeric receptor experiment as they have the highest sequence identity (45%) but the largest difference in pH_{50} - pH 8 for GPR4 and pH 6.8 for GPR68 when expressed heterologously in HEK293 cells (Fig. 2.1b). If the proton sensors act via the hotspot mechanism, we expected to find that swapping one segment of GPR68 in place of the homologous GPR4 section would be sufficient to convert the pH_{50} of GPR4 into that of GPR68. Conversely, if sensing

is distributed throughout the receptor, we would expect that multiple linear segments of GPR68 would be required to convert the pH_{50}.

We created all possible combinations of chimeric swaps between GPR68 and GPR4. We chose linear segments that spanned both extracellular loops (ECL) and portions of transmembrane (TM) helices containing titratable residues. We chose grafting points by matching the final Ballesteros-Weinstein position where GPR4 and GPR68 shared residue identity before diverging. This resulted in 14 chimeras of GPR4 containing the following portions of GPR68: N terminus and top of TM1, TM2 through TM3 including ECL1, TM4 through TM5 including ECL2, and TM6 through TM7 including ECL3. Surprisingly, we found that any one linear segment had almost no effect on pH_{50}, and even combinations of three segments only shifted the pH_{50} by a half a pH unit (Fig. 2.1c, Fig. 2.7). It was not until all extracellular segments of GPR68 were swapped onto GPR4 that the pH_{50} approached wild-type GPR68. The remaining difference in pH_{50}s between this chimera and wild-type GPR68 suggests that there are residues outside the swapped regions contributing to proton sensing.

This simple chimera experiment yielded three key insights into the mechanism of proton-sensing GPCRs. First, there is no single site of amino acids that accounts for the pH response. Second, shifting the pH_{50} between receptors requires nearly the entire extracellular half of the receptor. Third, there are residues outside the extracellular loops and upper TMs that contribute to the pH response. These findings provide clear support for a distributed model of activation. However, the large number of residues involved presents an enormous challenge to study and conceptualize a detailed molecular model of how these GPCRs sense protons. To probe over 100 extracellular

residues, of which over 40 are titratable, we needed a high-throughput assay that tested the pH response of each site across a variety of amino acid identities. For this, we turned to DMS.

Deep mutational scan provides unbiased and comprehensive view of distributed proton sensing mechanism

DMS provides high-throughput sequence to function analysis. Each position in a protein is mutated to all 20 amino acids and the mutational effect on protein function, or fitness, is quantified by a downstream signal [122]. This signal is often a fluorescence output that allows for fluorescence-activated cell sorting (FACS) of the pooled mutational library. The sorted library is then sequenced and the enrichment of different genotypes in sorted gates is calculated. This enrichment provides a relative metric for how variants increase or decrease fitness compared to wild-type (Fig. 2.8). DMS has been applied to understand functional outcomes of membrane protein mutations, including the prototypical GPCR, β2AR [123–126]. DMS on GPCRs enables detailed understanding of how nuanced changes to each and every amino acid in a receptor changes signaling output.

To apply DMS to a proton-sensing GPCR, we had to overcome two challenges. First, basic and acidic pH is toxic for cells. Second, basal production of cAMP, and consequently, basal production of the fluorescent reporter attenuates the dynamic range of signal produced by pH stimulus. We designed a system to overcome these challenges (Fig. 2.2a). To circumvent the pH range, we tested different pH points spanning different portions of the GPR68 dose-response curve and we used a previously characterized GPR68 PAM, ogerin [127] (Fig. 2.2b, c). Additionally, we minimized the background cAMP signal by fusing a DHFR degron onto eGFP so that it is continuously degraded until the library is stimulated by pH.

We performed triplicate DMS on GPR68 at three different pHs, pH 6.5, pH 6, and pH 5.5. These pH values were chosen to assess function at a fully inactivating, an intermediary, and a fully activating pH, respectively. We measured the effects of all mutations on both cAMP signaling and surface expression relative to wild-type GPR68 to capture the effects of all mutations at all positions (Fig. 2.2d, Fig. 2.9). Measuring surface expression for all variants allows us to distinguish mutants that have lower cAMP production because they do not reach the plasma membrane from those that reach the plasma membrane and have altered signaling. From these data we can see mutations that decrease cAMP signaling, indicated by negative fitness scores, and we are able to see mutations that increase cAMP signaling, indicated by positive fitness scores.

The GPR68 DMS data provides a global view of pH activation determinants that aligns with known constraints on GPCR folding and signaling. Most notable among these global patterns is the mutational tolerance of each TM helix. Mutations in each TM helix produce large decreases in surface expression and signaling fitness, especially when hydrophobic residues are substituted for a charged or proline residue (Fig 2.2d, Fig 2.9). Additionally, mutations to conserved GPCR folding motifs, such the disulfide C94$^{3.25}$ and C172^{ECL2}, are intolerant to mutation. In addition to capturing key GPCR folding determinants, the GPR68 DMS data also captured the functional importance of conserved GPCR signaling determinants, including the DRY motif (D118$^{3.49}$, R119$^{3.50}$, and Y120$^{3.51}$) and DPxxY motif (D282$^{7.49}$, P283$^{7.50}$, Y286$^{7.53}$). Together these motifs are necessary for stabilizing the active state of GPCRs and opening the cytoplasmic binding cavity of G proteins[128]. These trends illustrate that the GPR68 DMS captured known features of GPCR folding and signaling.

To go beyond known GPCR fitness determinants, we next sought to understand specific GPR68 positions affecting pH sensing. Here we observed a range of effects. First, we observed positions that are intolerant to mutations and decrease fitness irrespective of mutation amino acid identity. For example, at both pH 5.5 and pH 6, mutations to R251$^{6.58}$ decrease fitness (Fig 2.2, Fig 2.9). Other examples of this include E174^{ECL2} and H245$^{6.52}$. Second, we observed positions that are permissive to mutations and have increased fitness irrespective of mutation amino acid identity. For example, at all three pHs, D91$^{3.22}$ and E160$^{4.64}$ have increased fitness when mutated with largest increases from non-charged amino acids (Fig 2.2, Fig 2.9). Intriguingly, there are also positions that display a charge-dependent effect on fitness, such as H20$^{1.31}$ and H269$^{7.36}$. These amino acids have decreased fitness when mutated to negatively charged residues and increased fitness when mutated to positively charged residues. This result suggests these histidines are protonation sites in activated GPR68, and it illustrates the additional insights that can be gained from DMS as opposed to alanine scanning or site-directed mutagenesis. The GPR68 DMS provided a rich dataset to understand the relationship between amino acid sequence and function, yet we struggled to understand how this multitude of mutational effects comprised a proton sensing network. To learn the full mapping of sequence, structure, and function, we turned to cryo-EM.

Structures of proton receptors show organization of the distributed proton sensing network

To understand the spatial organization of this proton sensing network, we determined cryo-EM structures of each proton sensor at pH 6 (Fig. 2.3a, Fig. 2.10, Fig. 2.11, Fig 2.12). We had to overcome two challenges to determine these structures. First, we overcame low receptor expression by generating constructs with each proton receptor C-terminally fused to a miniG

protein based on their G protein coupling specificities [120,128]. This enabled measurable quantities of each receptor to be purified. Second, we overcame receptor instability by screening different pH buffers for structural determination. Although each receptor has a different pH_{50} when expressed heterologously in HEK293 cells, each required pH 6 biochemically for high resolution structure determination. At pHs higher than this, reconstructions were low resolution. To overcome instability challenges, we additionally attempted determination of GPR68 structures with various PAMs, including Co^{2+} and MS48107, a more potent ogerin analogue [120,129]. With 10 μM Co^{2+}, we were able to determine high resolution GPR68 structures at pH 6 and pH 7.5, but we were unable to resolve the small molecule, MS48017, bound to GPR68 any pH (Fig. 2.13, Fig. 2.14, Fig 2.15).

From the resulting full particle and focused GPCR refinements, we modeled each proton-sensing GPCR and G protein heterotrimer (Fig. 2.3b). Each proton receptor has a highly similar active-state confirmation (RMSD < 1.6 Å). The three proton receptors share conserved titratable residues in TM4, TM5, TM6, and ECL2. Elsewhere the identity and side chain positioning of titratable residues diverges (Fig. 3c). Of particular note is the conformation of ECL2 and canonical "ligand binding pocket" for each receptor. Previous studies have predicted the proton receptors to have peptide ligands [129–131]. Indeed the ECL2 of each is a beta hairpin similar to peptide-binding GPCRs (Fig. 2.16a, b) [130–132]. However unlike these peptide receptors, the ECL2 of the proton sensors occupies the canonical ligand binding pocket, much like recently described orphan GPCRs (Fig. 2.16a, c) [37,133]. Both ECL2 and the neighboring TM sites are abundant with titratable residues in contrast to self-activating orphans and GPCRs that bind small molecule and peptide ligands[133]. However, there are minimal pockets present in these structures (GPR4: 114 $Å^3$, GPR65: 110 $Å^3$, GPR68: 188 $Å^3$) comparable to about one third the size of aminergic GPCRs like β2AR (370 $Å^3$)

(Fig. 2.16d). These are likely occupied by water molecules, but the presence of these pockets distinctly leaves open the possibility of an undiscovered endogenous stimulus other than protons for these receptors.

GPR68 activation network is distributed across the extracellular side of the receptor and proceeds through canonical GPCR activation motifs

The combination of our DMS and structural data provides a complete and detailed picture of the mechanism of proton-sensing GPCR activation. To visualize this, we focused on variants that retained surface expression (Fig. 2.4a). Retaining surface expression was defined as having a surface expression fitness value of greater than -0.5. This removes variants that have defective folding and biogenesis, which most often are caused by introducing charged or aromatic amino acids into the center of TM helices. With the variants that retained surface expression, we calculated each position's "importance" for pH activated cAMP signaling (Fig. 2.4b). We calculated this per sequence position by summing the absolute value of the fitness score for each surface expressed variant. We then used this importance score to color and display the top 15% most important amino acids on the structure of GPR68 (Fig. 2.4c). This allowed us to visualize positions that have the largest effect on receptor activation.

From the combination of our DMS and structural data, the activation mechanism of GPR68 immediately stands out. There is an expansive network of residues primarily composed of aspartates, glutamates, histidines, arginines, and tyrosines, that are important for proton activation (Fig. 2.4c-e). These residues are found spanning what would be the canonical ligand binding and ligand entry pathway in other Class A GPCRs. Notably these residues are partially to completely

protected from solvent, and titratable residues that are completely solvent exposed have less importance. This network begins with a cluster of titratable residues near TM1 and TM7, including H20$^{1.31}$ and H269$^{7.36}$ (Fig. 2.4d). In the center of the receptor, the network is bifurcated by a glutamate E174^{ECL2} which interacts with R251$^{6.58}$ and Y102$^{3.33}$ at the base of the distributed network (Fig. 4e). Below Y102$^{3.33}$, the network narrows to a conserved histidine H245$^{6.53}$, hallmark GPCR activation motifs like the DPxxY motif and DRY motif, and G protein contacts (Fig. 2.4g,f) [134].

GPR68 DMS reveals position-specific mutational effects on proton sensing

We next sought to understand the details of how specific perturbations to this distributed network tunes proton sensing. To do this we returned to the most important acidic and basic residues (Fig. 2.5a-c). First looking at histidines, we noticed charge-dependent trends (Fig. 2.5d). H20$^{1.31}$ mutations increase fitness when changed to a positively charged residue and decrease fitness when changed to a neutral or negatively charged residue. This trend holds for neutral and negatively charged residues when mutations are assayed in the GloSensor cAMP production assay (Fig. 5g). The same trend is true for H269$^{7.36}$. This suggests that both are protonated during receptor activation. In contrast, H84$^{2.67}$ shows the opposite trend. Negative residues increase fitness, while others decrease. This suggests this residue is not protonated during activation and that the pH_{50} of GPR68 could be made more potent by the substitution of an acidic residue at this position. Second, glutamate mutations can result in loss-of-function (E255^{ECL3}) and gain-of-function (E164^{ECL2}) in addition to charge-dependent (E160$^{4.64}$) changes like those seen with histidine (Fig. 2.5e,h). Third, arginines have a wide range of activation importance. Despite being deeply buried, R189$^{5.42}$ can tolerate mutations to asparagine and histidine and other mutations have minor decreases in fitness.

In contrast, any mutation to R251$^{6.58}$ strongly decreases activation (Fig 2.5f,i). Mapping the effects of change in charge of residues with their neighboring residues allows inference of protonation giving rise to activation and provides informative mechanistic detail.

All three proton sensors activate through a disturbed mechanism with homologous tuning sites

Using our GPR68 network and learned key sites, we sought to extrapolate the mechanism to the other family members, GPR4 and GPR65. Similar to GPR68, most mutations made to titratable residues at homologous positions to GPR4 and GPR65 cause a decrease in pH_{50} (Table 2.1, Table 2.2, Table 2.3). This supports a distributed mechanism of activation for the entire family.

We also probed three important sites in GPR68 in the other two receptors. Having learned the charge dependence of GPR68 H269$^{7.36}$, we tested the homologous set of mutations for GPR4 and GPR65. Indeed, we see that at this site, positive charge increases the potency of proton sensing and neutral or negative charge decreases potency despite deviations in neighboring residues between the receptors (Fig. 2.6 a,b). Next, we examined the acidic ECL2 residue that bifurcates the distributed proton sensing network. This site appears to be critical for proton activation for both GPR65 and GPR68, and mutations here are among the most deleterious to pH_{50} and Emax (Fig. 2.6 c,d) although they have less of an effect on GPR4. Finally, we examined the most buried charged residue in these receptors, E$^{4.53}$. This buried glutamate is in the middle of TM4 and is far from both extracellular solvent and the intracellular G protein binding pocket. This is a highly unusual and energetically unfavorable location. When mutated to glutamine, we observed different effects between the receptors (Fig. 2.6 e,f). GPR4 E145Q$^{4.53}$ has almost no change. GPR65

E142Q$^{4.53}$ increases pH_{50} by 0.1 and GPR68 E149Q$^{4.53}$ increases pH_{50} by a full pH unit. This highlights both the subtle effects amino acid differences between these receptors have on the pH_{50} difference between them and that the unique proton sensing ability of these GPCRs emerges from their distributed electrostatic network.

2.4 Discussion

The proton-sensing GPCRs, GPR4, GPR65, and GPR68 play critical roles regulating diverse biology in the nervous, immune, and cardiovascular systems in response to acidic pH. Our study provides a comprehensive understanding of the proton-sensing mechanism of these GPCRs, revealing a distributed network of residues that contribute to the emergent property of proton sensing. This finding provides a new framework for understanding proton-sensing membrane proteins beyond the traditional hotspot model.

Our data suggests that the proton-sensing mechanism of these GPCRs is not localized to a single site, but rather is distributed across the entire extracellular half to the receptors. This distributed proton-sensing network is primarily composed of aspartates, glutamates, histidines, arginines, and tyrosines, which are found spanning what would be the canonical ligand binding pocket and ligand entry pathway in other Class A GPCRs. Notably, these residues are partially to completely protected from solvent creating increased energetic differences between the protonated and unprotonated states [135,136]. This network of residues is structurally contiguous with hallmark GPCR activation motifs and G protein contacts providing a visualization of residues important for proton sensor activation from extracellular loops to intracellular G protein contacts.

Through the combination of DMS and cryo-EM, our study provides a detailed picture of the distributed activation mechanism of proton-sensing GPCRs. This mechanistic understanding could have significant implications for the development of novel modulators targeting these receptors and provides a new model to apply to future studies of proton-sensing membrane proteins.

2.5 Figures

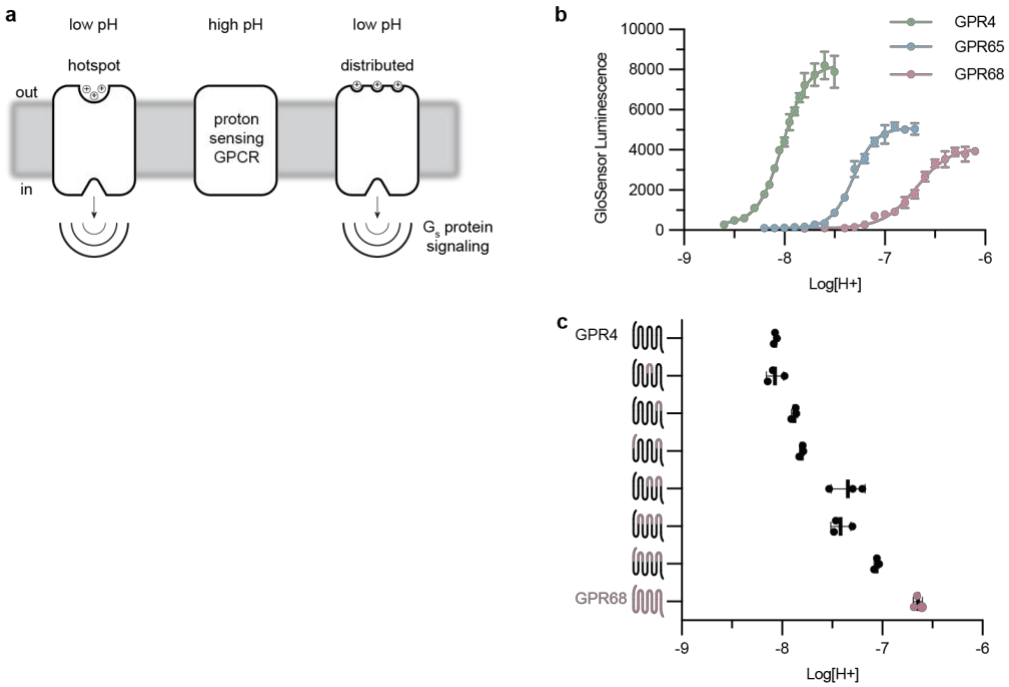


Figure 2.1 Proton-sensing GPCRs sense protons through a distributed mechanism
a) Proton-sensing GPCRs have been hypothesized to initiate G_s protein signaling in response to decreases in pH based on two potential mechanisms, "hotspot" sensing or "distributed" sensing. The hotspot hypothesis posits protonation of amino acids at a single site causes activation, analogous to traditional ligand binding models. The distributed hypothesis posits that protonation of amino acids at several sites causes activation. **b)** cAMP production assay showing the proton-sensing GPCRs, GPR4, GPR65, and GPR68, respond to decreasing pH. dose response curves showing proton activation for each proton sensor, GPR4, GPR65, and GPR68. **c)** Extracellular segments of GPR68 exchanged onto GPR4 are insufficient to convert the pH_{50} of GPR4 to the pH_{50} of GPR68. This result refutes the hotspot mechanism and supports the distributed mechanism of proton sensing. Data are pH_{50} from three independent biological replicates ± s.d.

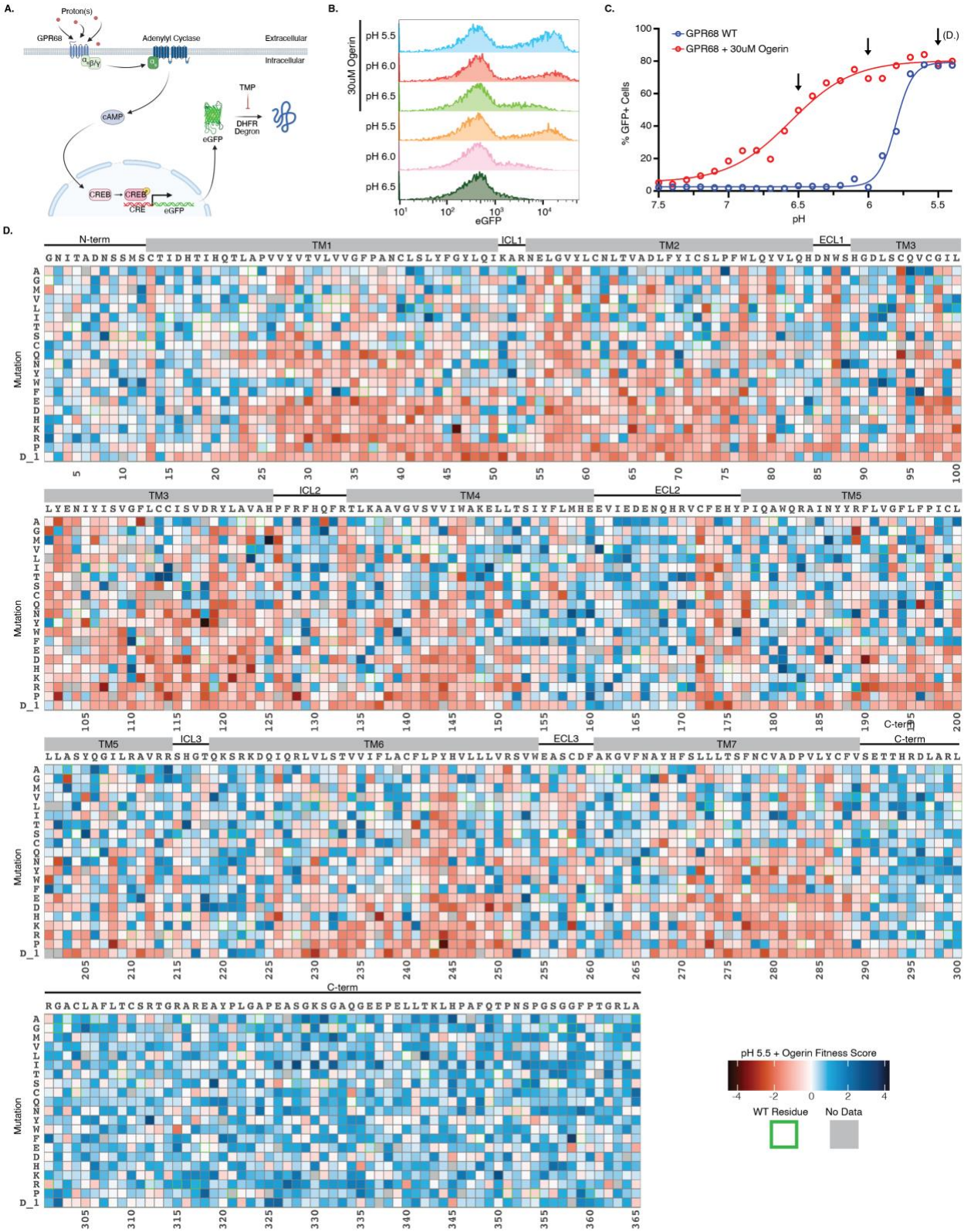


Figure 2.2 Deep Mutational Scan of GPR68

a) Schematic of cAMP reporter system used for DMS. The GPR68 DMS library was integrated into stable cells using DIMPLE. Receptor stimulation is detected as a fluorescent output by

transcription of eGFP downstream of cAMP response element-binding protein (CREB) and the addition of trimethoprim (TMP) to inhibit the basal degradation of eGFP. **b)** Representative flow traces of GPR68 cAMP reporter fluorescence at screened pH conditions. eGFP +/- gate is shown as a dotted line which was used to generate the dose-response curve in c. **c)** pH dose-response curves for GPR68 WT +/- 30uM Ogerin. Arrows indicate conditions screened, individual traces are shown in b. **d)** Heatmap of DMS fitness values for GPR68 at pH 5.5. WT sequence is shown above each section of heatmap, mutations are indicated on the left axis of each section, and the position along the WT sequence is indicated by the numbers below each section. Positions and mutations with no data are shown as gray, and the WT amino acid at each position has a green border. Cutoffs for transmembrane helices, loops obtained from GPCRdb. Fitness scores are relative to WT and were calculated using Enrich2. Increased fitness (more blue) indicates increased cAMP signaling relative to WT and decreased fitness (more red) indicates decreased cAMP signaling relative to WT. Scores above/below the 1st/99th percentile are shown as the value of the 1st/99th percentile. Data are fitness values from three biologically independent deep mutational scans.

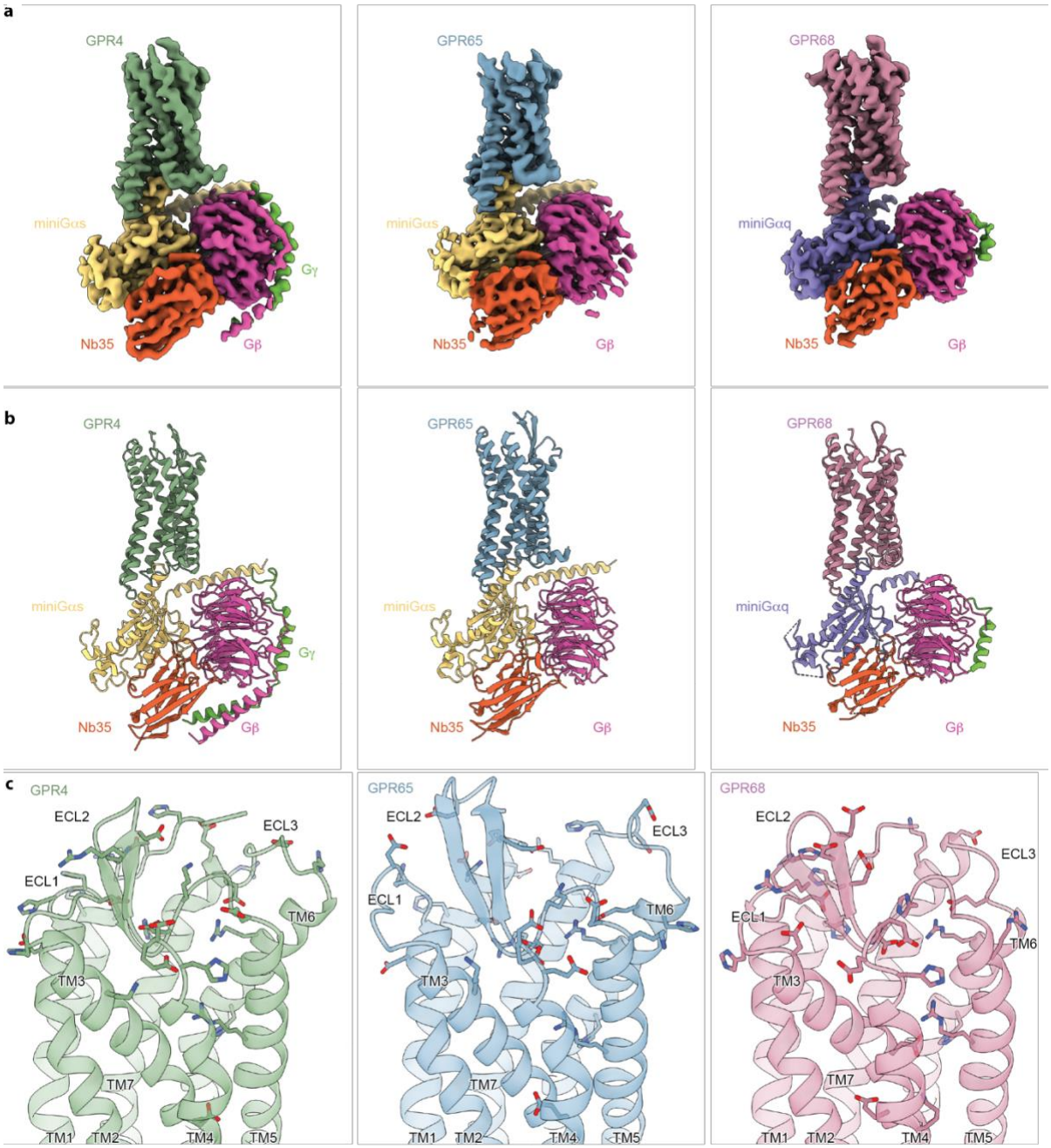


Figure 2.3 Structures of active proton-sensing GPCRs at pH 6

a) Cryo-EM density maps of GPR4, GPR65, and GPR68. GPR4 and GPR65 are fused to mini$G\alpha_s$, and GPR68 is fused to mini$G\alpha_q$. All three are bound to $G\beta\gamma$ and the stabilizing nanobody Nb35. **b)** Ribbon model of GPR4, GPR65, and GPR68 active G protein complexes. **c)** View of the extracellular halves of GPR4, GPR65, and GPR68 with transmembrane (TM) helices and extracellular loops (ECLs) labeled and titratable residues shown as sticks. The three proton receptors have nearly identical active-state backbone conformations. The three proton receptors share conserved titratable residues in TM4, TM5, TM6, and ECL2 but diverge elsewhere.

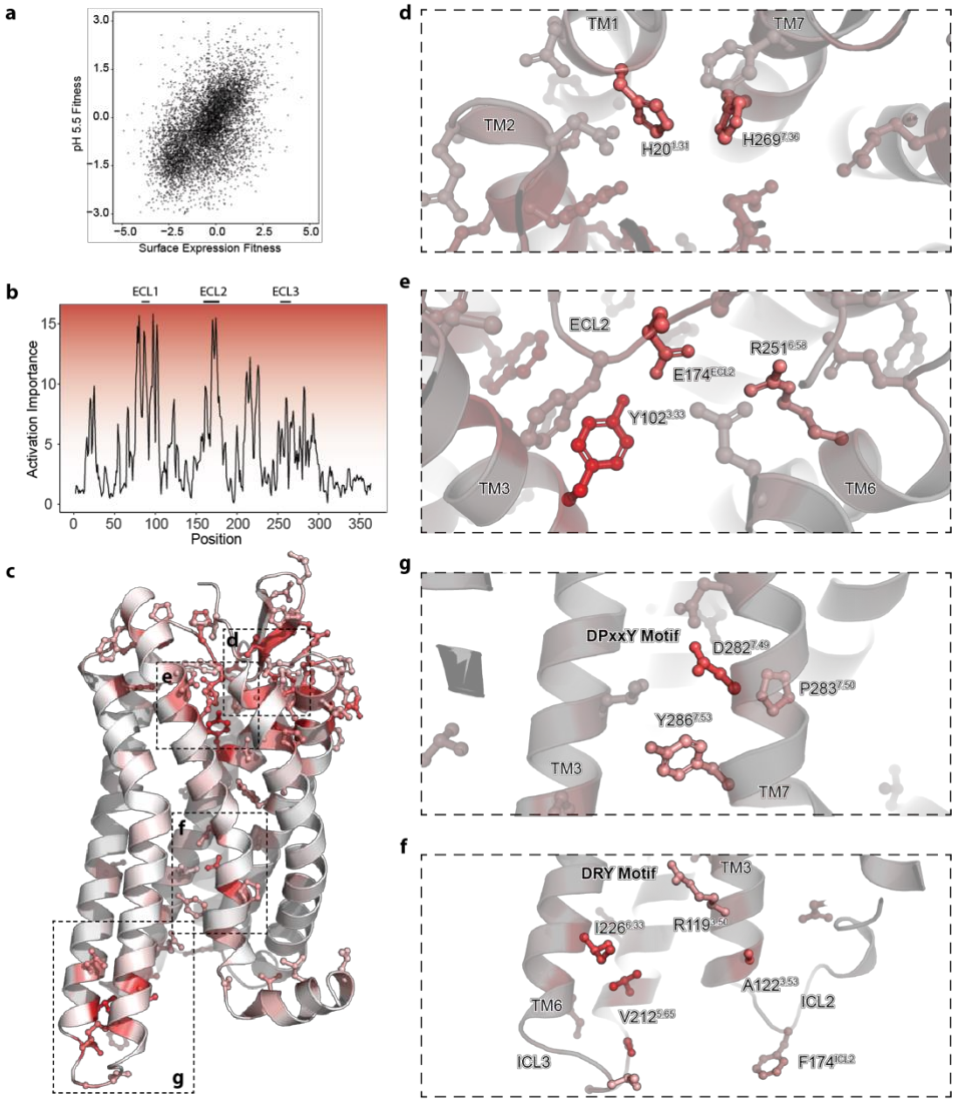


Figure 2.4 GPR68 activation network is distributed across the extracellular side of the receptor and proceeds through canonical GPCR activation motifs

a) Scatter plot of surface expression vs pH 5.5 fitness. Data points with greater than zero surface expression fitness and less than zero pH 5.5 fitness are expressed similar to or better than wild-type GPR68 but have lower cAMP signaling. **b)** Activation importance (summed absolute value of fitness per position) from A colored white to red. Red indicates relatively more important positions and white is less important. **c)** Structure of GPR68 with the top 15% importance (absolute fitness) scores from the pH 5.5 DMS. Positions with the largest importance are distributed across the extracellular side of the receptor (**d, e)** indicating proton-sensing is distributed amongst these residues. The canonical GPCR activation motifs - DPxxY (**g**), DRY (**f**) - and G protein contacts (**f**) are also among the most important sites for cAMP signaling.

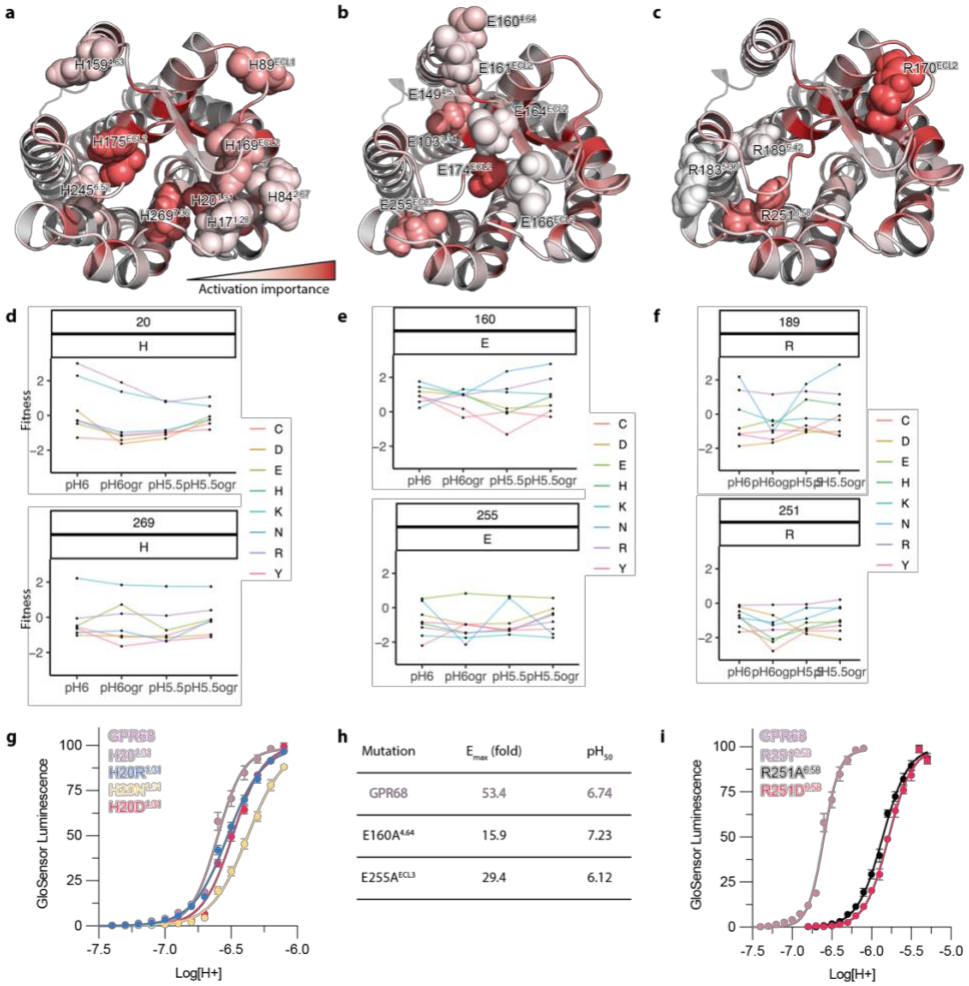


Figure 2.5 DMS reveals position-specific mutational effects on proton sensing
a-c) Top views of GPR68 structure with amino acids colored by activation importance value and different residue types show as spheres - **a)** histidines **b)** glutamate **c)** arginine. **d-f)** Fitness values for specific residues across DMS assay conditions and mutations **d)** H20$^{1.31}$ has increased fitness when mutated to basic residues while others are neutral or decrease fitness. H269$^{7.36}$ has increased fitness when mutated to H269K$^{7.36}$ while others are neutral or decrease fitness. **e)** E160$^{4.64}$ is permissive to most mutations, except alanine. E255^{ECL3} coordinates with R251$^{6.58}$ and mutations decrease activation. **f)** R189$^{5.42}$ is deeply buried and intolerant to mutation expect for histidine and asparagine, which are seen in other GPCRs at this position. R251$^{6.58}$ mutations decrease fitness regardless of amino acid identity. **g)** cAMP production assay showing the pH response of H20$^{1.31}$ mutants. H20R$^{1.31}$, H20N$^{1.31}$, and H20D$^{1.31}$ decrease pH50. **h)** cAMP production E_{max} and pH_{50} values for E160A$^{4.64}$ and E255A^{ECL3}. **i)** cAMP production assay showing the pH response of R251A$^{6.58}$ and R251D$^{6.58}$.

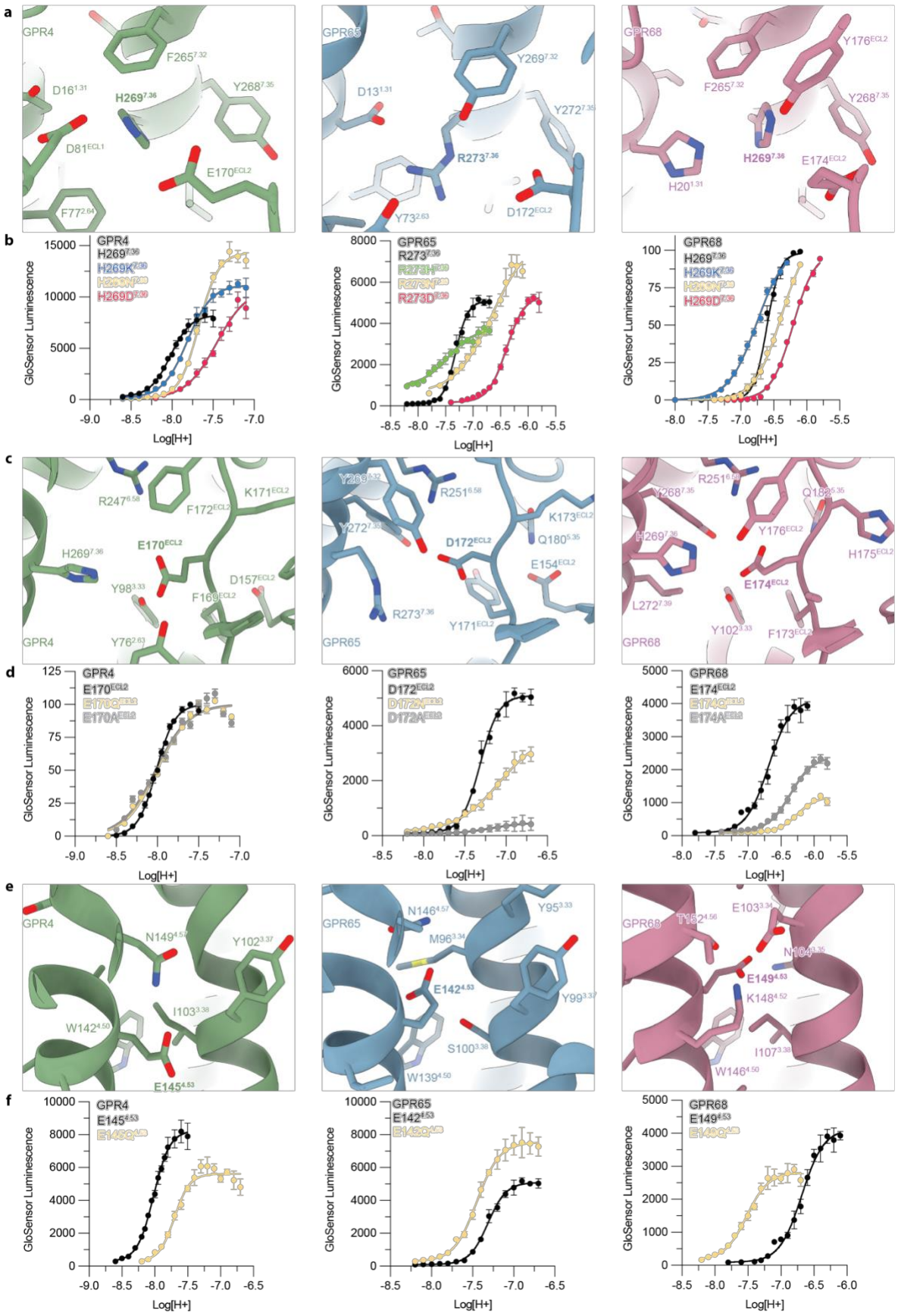

a
GPR4
F265 7.32
D16 1.31
H269 7.36
Y268 7.35
D81 ECL1
E170 ECL2
F77 2.64
GPR65
Y269 7.32
D13 1.31
Y272 7.35
R273 7.36
Y73 2.63
D172 ECL2
GPR68
Y176 ECL2
F265 7.32
Y268 7.35
H269 7.36
E174 ECL2
H20 1.31
b
GPR4
H269 7.36
H269K 7.36
H269N 7.36
H269D 7.36
GloSensor Luminescence
Log[H+]
GPR65
R273 7.36
R273H 7.36
R273N 7.36
R273D 7.36
GloSensor Luminescence
Log[H+]
GPR68
H269 7.36
H269K 7.36
H269N 7.36
H269D 7.36
GloSensor Luminescence
Log[H+]
c
R247 6.58
F172 ECL2
K171 ECL2
H269 7.36
E170 ECL2
Y98 3.33
D157 ECL2
F169 ECL2
GPR4
Y76 2.63
Y269 7.32
R251 6.58
K173 ECL2
Y272 7.35
D172 ECL2
Q180 5.35
E154 ECL2
R273 7.36
Y171 ECL2
GPR65
Y268 7.35
R251 6.58
Q182 5.35
H269 7.36
Y176 ECL2
H175 ECL2
E174 ECL2
L272 7.39
Y102 3.33
F173 ECL2
GPR68
d
GPR4
E170 ECL2
E170Q ECL2
E170A ECL2
GloSensor Luminescence
Log[H+]
GPR65
D172 ECL2
D172N ECL2
D172A ECL2
GloSensor Luminescence
Log[H+]
GPR68
E174 ECL2
E174Q ECL2
E174A ECL2
GloSensor Luminescence
Log[H+]
e
GPR4
N149 4.57
Y102 3.37
I103 3.38
W142 4.50
E145 4.53
GPR65
N146 4.57
Y95 3.33
M96 3.34
E142 4.53
Y99 3.37
S100 3.38
W139 4.50
GPR68
T152 4.56
E103 3.34
N104 3.35
E149 4.53
K148 4.52
I107 3.38
W146 4.50
f
GPR4
E145 4.53
E145Q 4.53
GloSensor Luminescence
Log[H+]
GPR65
E142 4.53
E142Q 4.53
GloSensor Luminescence
Log[H+]
GPR68
E149 4.53
E149Q 4.53
GloSensor Luminescence
Log[H+]

Figure 2.6 All three proton sensors activate through a distributed mechanism with homologous tuning sites

a) Close up structural view of position 7.36 for GPR4, GPR65, and GPR68. Residues within 5 Å of the amino acid at 7.36 are shown. GPR4 contains $H269^{7.36}$ with three neighboring acidic residues. GPR65 contains $R273^{7.36}$ with two neighboring acidic residues. GPR68 contains $H269^{7.36}$ with one neighboring acidic and basic residue. **b)** cAMP production assay showing the pH response of mutations at position 7.36 for GPR4, GPR65, and GPR68. Generally, basic residues increase, neutral residues slightly decrease, and acidic residues most strongly decrease pH_{50}. **c)** Close up structural view of the shared acidic ECL2 at the center of the canonical ligand pocket for GPR4, GPR65, and GPR68, which is two positions past the conserved ECL2 disulfide in Class A GPCRs. Residues within 5 Å of the ECL2 acidic residue are shown. GPR4 contains $E170^{ECL2}$ with two neighboring basic residues and two tyrosines. GPR65 contains $D172^{ECL2}$ with two neighboring basic residues and one tyrosine. GPR68 contains $E174^{ECL2}$ with two neighboring basic residues and three tyrosines. **d)** cAMP production assay showing the pH response of mutations at the ECL2 position for GPR4, GPR65, and GPR68. Mutations to an amide or alanine decrease pH_{50} for GPR65 and GPR68. **e)** Close up structural view of position 4.53 for GPR4, GPR65, and GPR68. Residues within 5 Å of the amino acid at 4.53 are shown. GPR4 contains $E145^{4.53}$ with one neighboring asparagine. GPR65 contains $E142^{4.53}$ with one neighboring asparagine and serine. GPR68 contains $E149^{4.53}$ with two neighboring neutral, one basic, and one acidic residues **f)** cAMP production assay showing the pH response of mutations at the ECL2 position for GPR4, GPR65, and GPR68. Mutations to glutamine decrease pH_{50} for GPR4 while increasing pH_{50} for GPR65 and GPR68. This highlights the differential effects of the homologous tuning sites between these receptors.

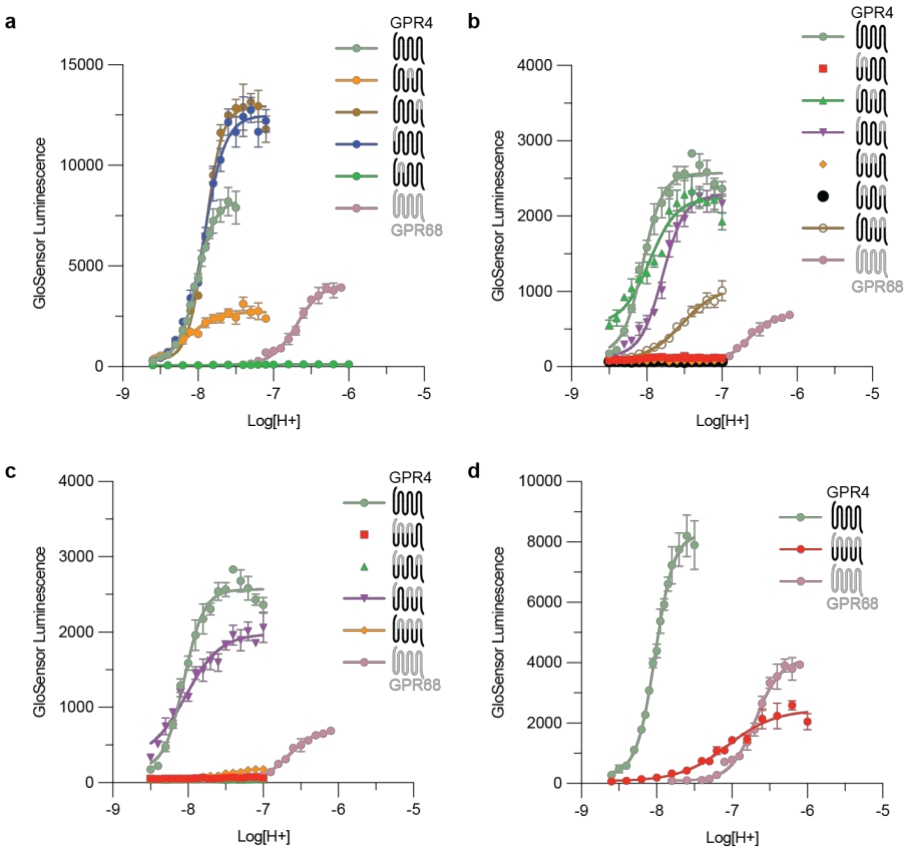


Figure 2.7 cAMP production of GPR4-GPR68 chimeras

a-d) cAMP production assay showing proton activation for GPR4-GPR68 chimeras. All extracellular segments of GPR68 exchanged onto GPR4 are necessary to convert the proton response of GPR4 to the proton response of GPR68. Three or less segments is insufficient. Data are representative technical replicates from three independent biological replicates ± s.d.

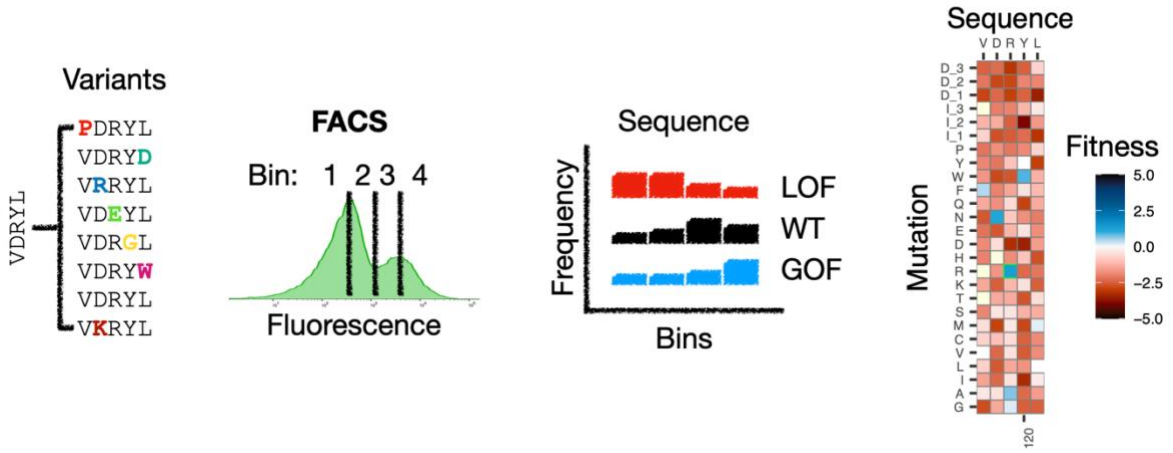
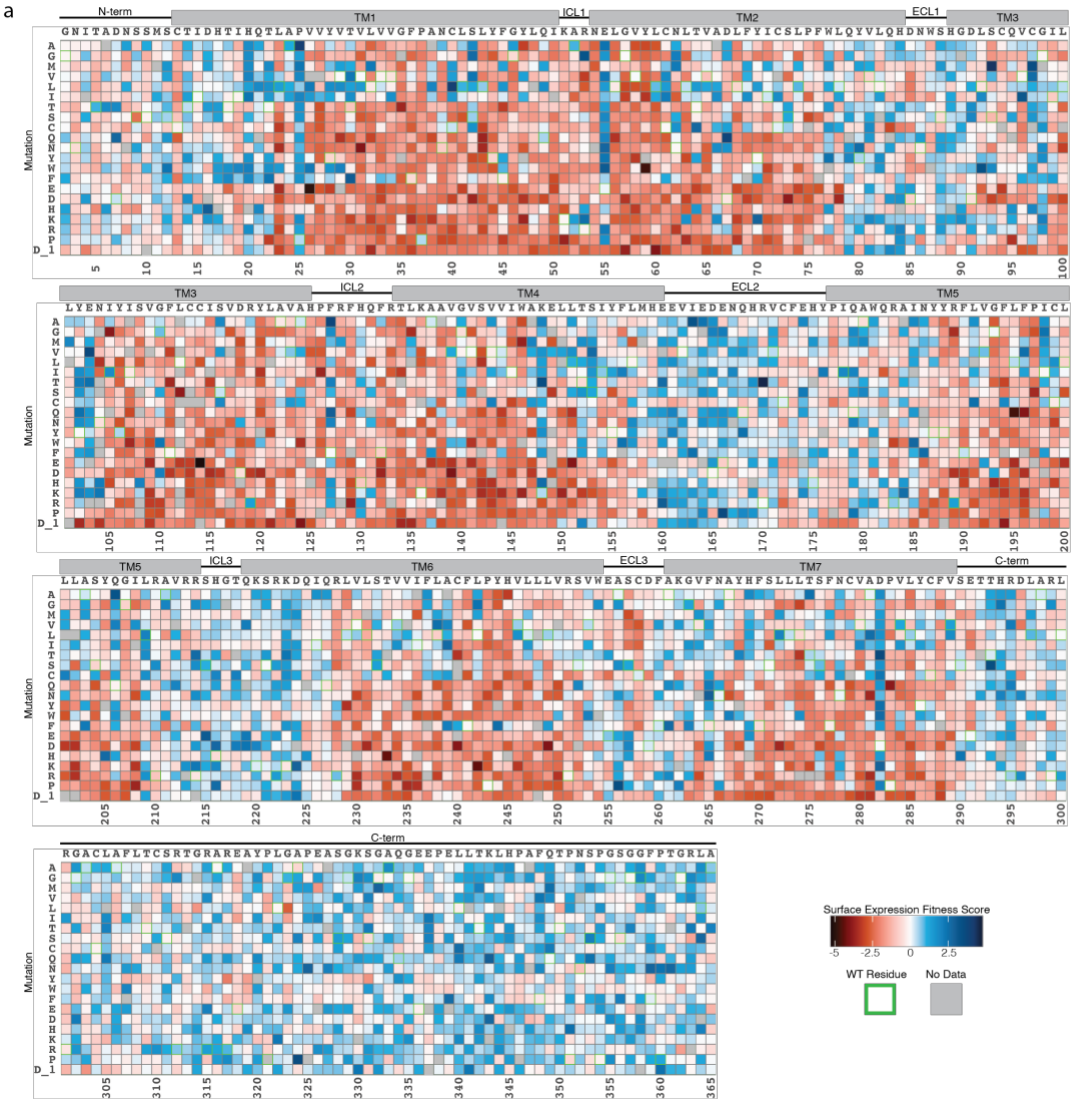
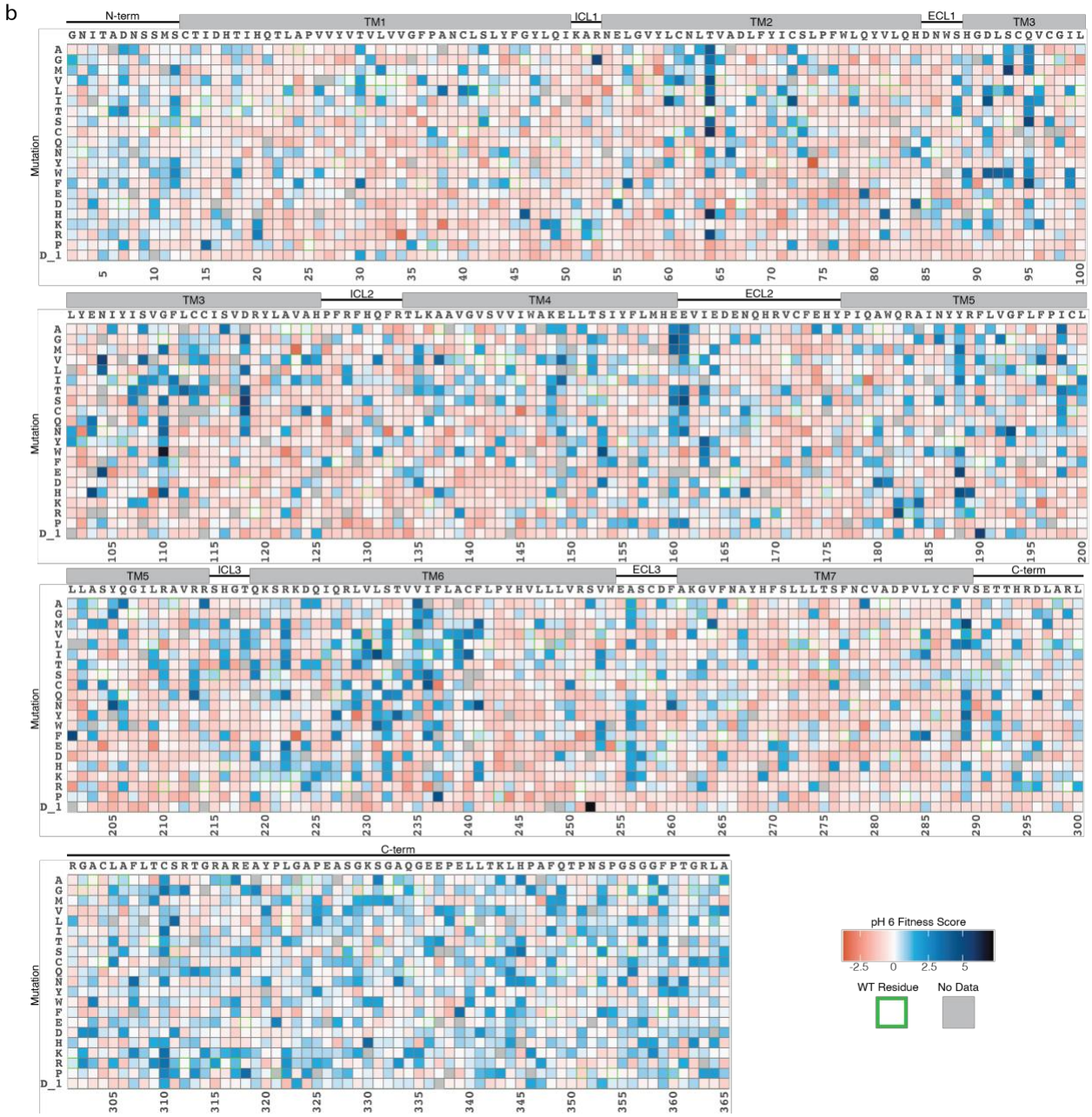


Figure 2.8 DMS workflow schematic
Briefly, we use DIMPLE to generate a library including insertion, missense, synonymous, and deletion variants at all positions of GPR68, we generate stable HEK293 cell lines, sort these cells based on surface-expression using FACS, perform deep sequencing of these subpopulations, and calculate surface expression fitness scores using Enrich2.


a
N-term
TM1
ICL1
TM2
ECL1
TM3
ICL2
TM4
ECL2
TM5
ICL3
TM6
ECL3
TM7
C-term
Mutation
Surface Expression Fitness Score
-5
-2.5
0
2.5
WT Residue
No Data


b
N-term
TM1
ICL1
TM2
ECL1
TM3
TM3
ICL2
TM4
ECL2
TM5
TM5
ICL3
TM6
ECL3
TM7
C-term
C-term
Mutation
pH 6 Fitness Score
-2.5
0
2.5
5
WT Residue
No Data

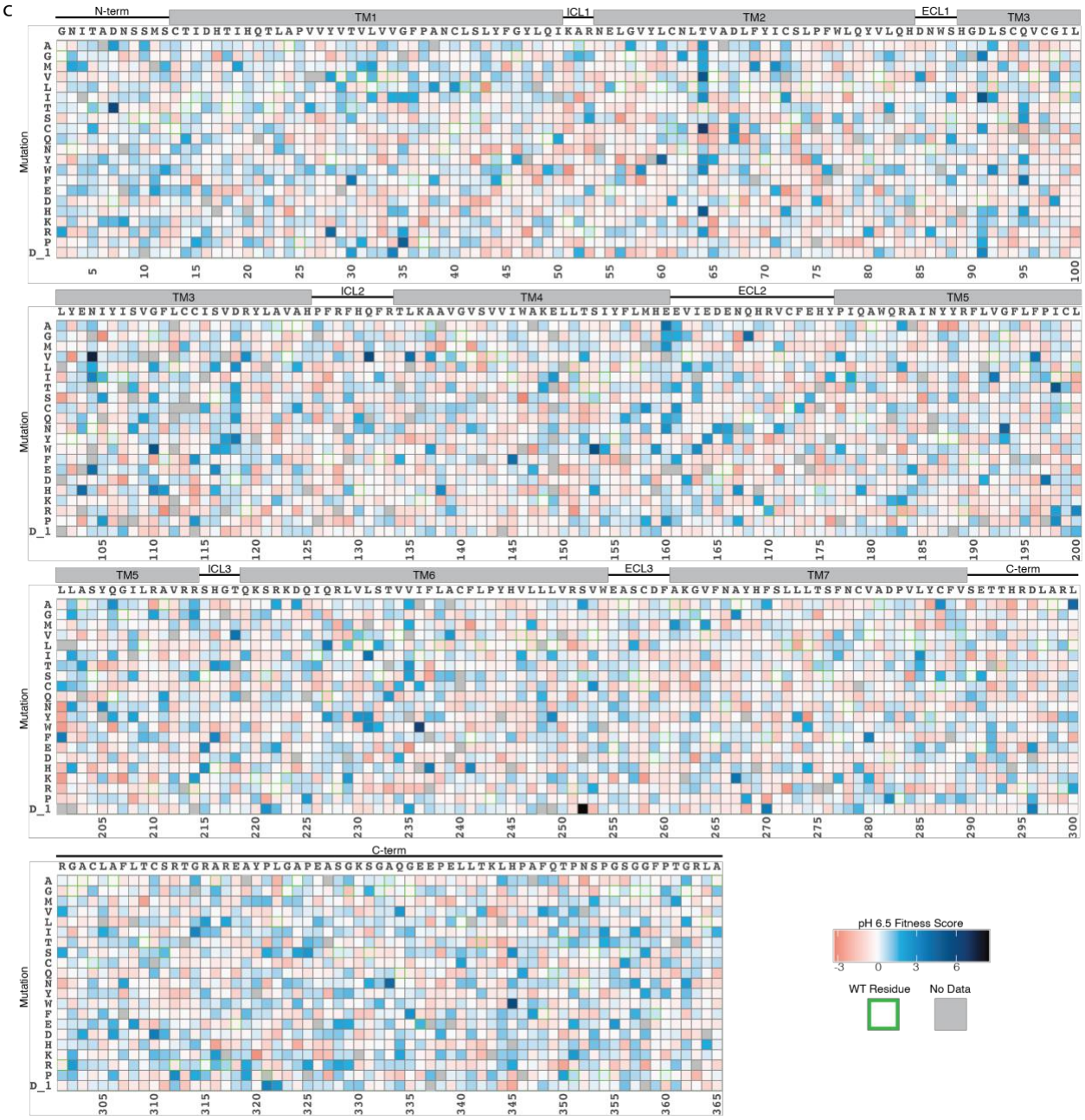


Figure 2.9 DMS heatmap of surface expression

Heatmap of DMS fitness values for GPR68 **a)** surface expression, **b)** pH 6, **c)** pH 6.5. WT sequence is shown above each section of heatmap, mutations are indicated on the left axis of each section, and the position along the WT sequence is indicated by the numbers below each section. Positions and mutations with no data are shown as gray, and the WT amino acid at each position has a green border. Fitness scores are relative to WT and were calculated using Enrich2. Increased fitness (more blue) indicates increased surface expression or signaling relative to WT and decreased fitness (more red) indicates decreased surface expression or signaling relative to WT. Scores above/below the 1st/99th percentile are shown as the value of the 1st/99th percentile. Data are fitness values from three biologically independent deep mutational scans.

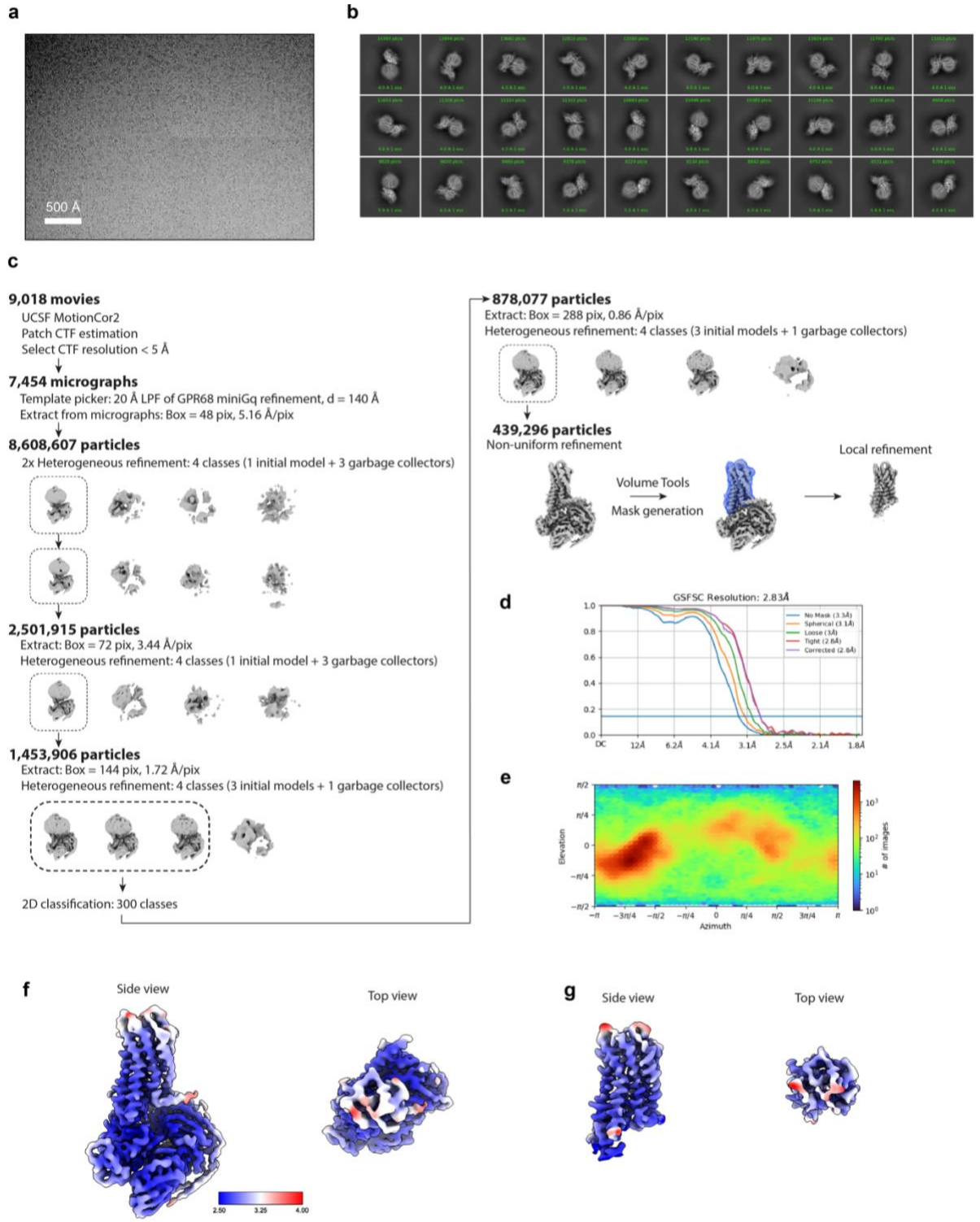


Figure 2.10 Cryogenic electron microscopy processing of GPR4 G_s pH 6
a) A representative motion-corrected cryogenic electron microscopy (cryo-EM) micrograph obtained from a Titan Krios microscope**. b)** A subset of highly populated, reference-free 2D-class averages. **c)** Schematic showing the cryo-EM data processing workflow. Initial processing was performed using UCSF MotionCor2 and cryoSPARC. Particles were selected using iterative Heterogeneous refinement jobs followed by 2D classification. Finally, particles were processed

using the local refinement job type with a 7TM mask. Dashed boxes indicated selected classes. **d)** Gold-standard Fourier Shell Correlation (GSFSC) curve for final full-particle map computed in cryoSPARC. **e)** Euler angle distribution of final full-particle map computed in cryoSPARC. **f)** Side view and top view of local resolution for the final full-particle map of GPR4-G_s pH 6 complex computed with local resolution in cryoSPARC. **g)** Side view and top view of local resolution for the focused 7TM map of GPR4-G_s pH 6 complex computed with local resolution in cryoSPARC.

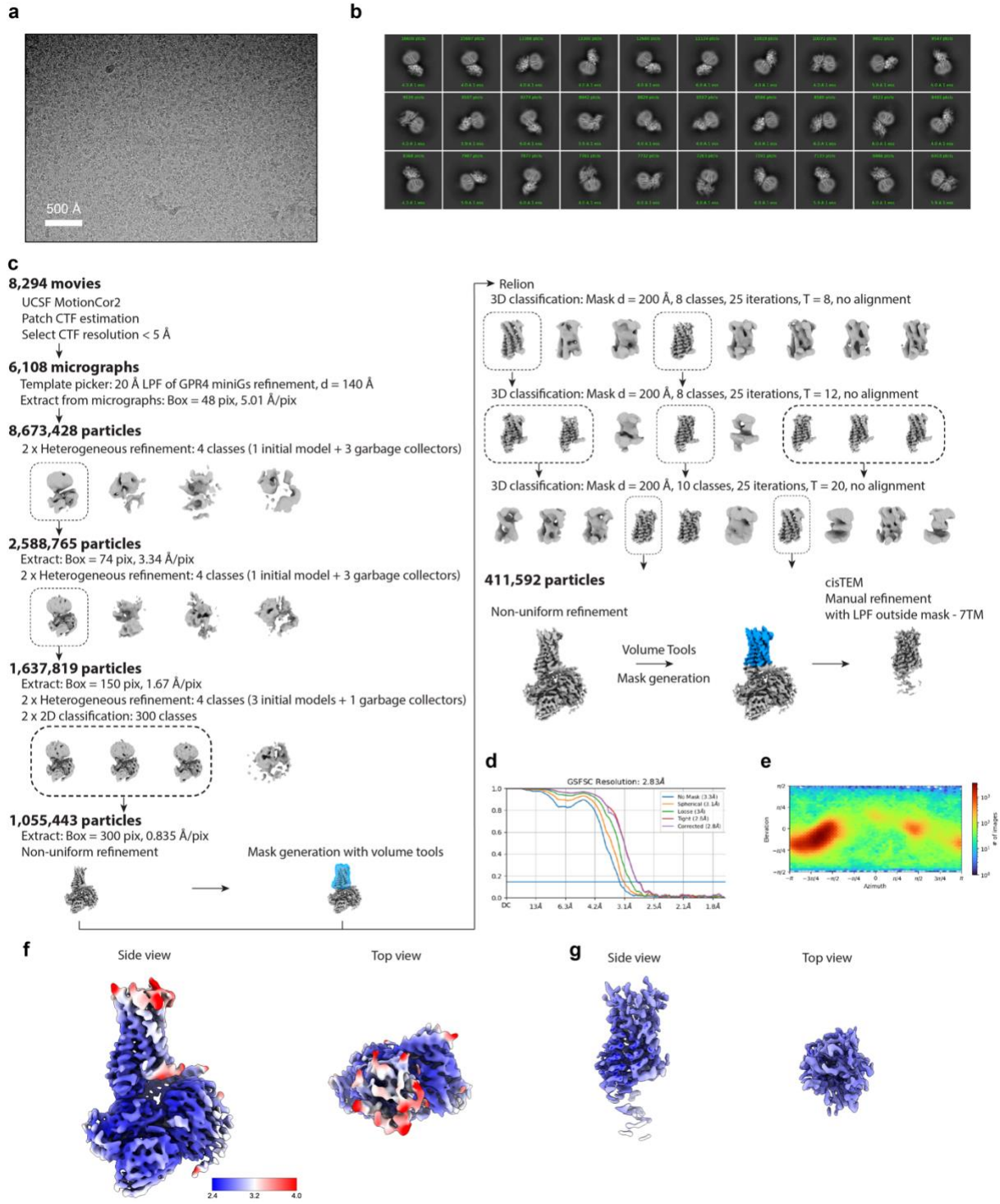


Figure 2.11 Cryogenic electron microscopy processing of GPR65 G_s pH 6

a) A representative motion-corrected cryogenic electron microscopy (cryo-EM) micrograph obtained from a Titan Krios microscope**. b)** A subset of highly populated, reference-free 2D-class averages. **c)** Schematic showing the cryo-EM data processing workflow. Initial processing was performed using UCSF MotionCor2 and cryoSPARC. Particles were transferred using the pyem script package to RELION for alignment-free 3D classification. Finally, particles were processed in cisTEM using the manual refinement job type with a 7TM mask followed by a full particle

mask. Dashed boxes indicated selected classes. **d)** Gold-standard Fourier Shell Correlation (GSFSC) curve for final full-particle map computed in cryoSPARC. **e)** Euler angle distribution of final full-particle map computed in cryoSPARC. **f)** Side view and top view of local resolution for the final full-particle map of GPR65-G_s pH 6 complex computed with local resolution in cryoSPARC. **g)** Side view and top view of local resolution for the focused 7TM map of GPR65-G_s pH 6 complex computed with local resolution in cryoSPARC.

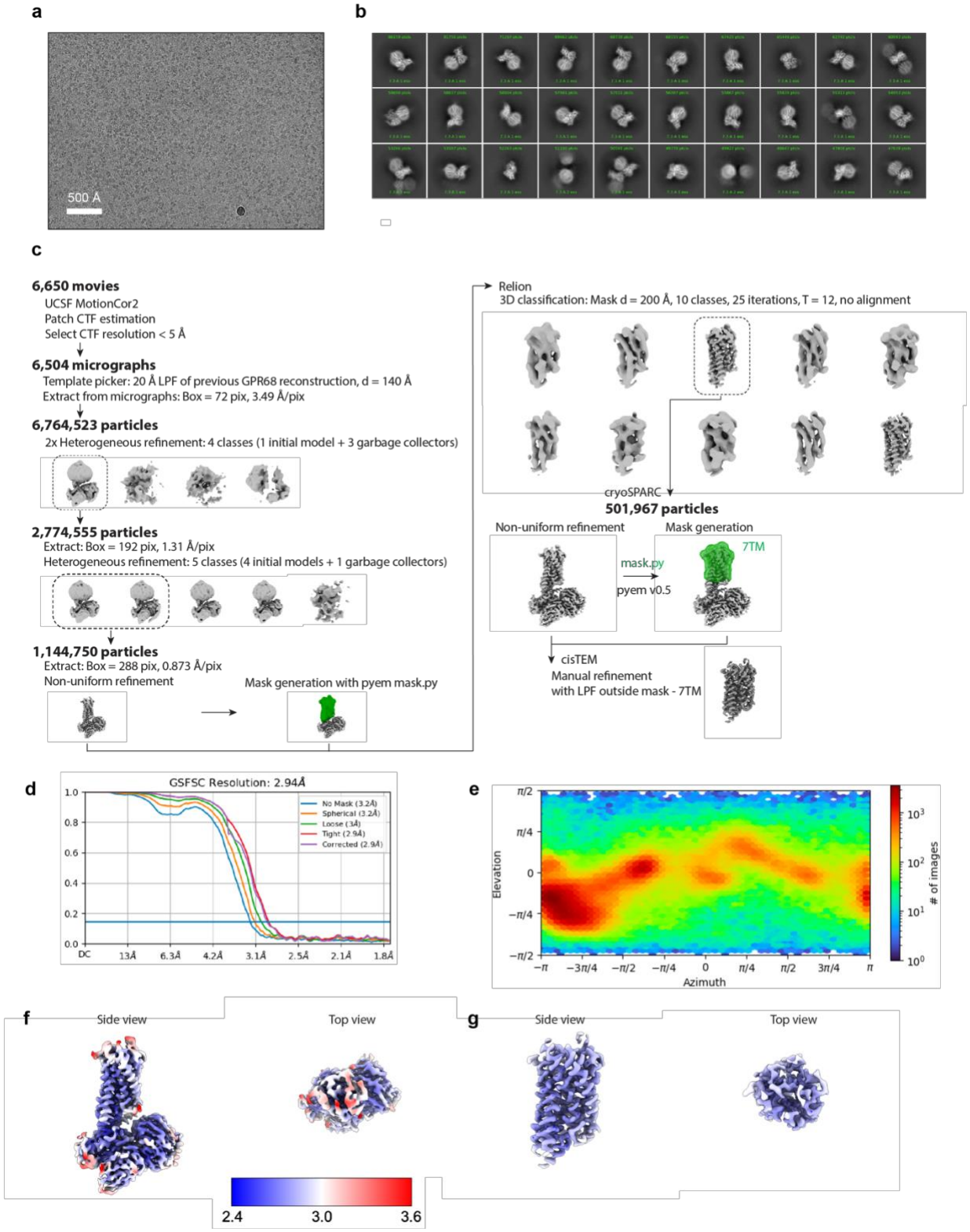


Figure 2.12 Cryogenic electron microscopy processing of GPR68 G_q pH 6

a) A representative motion-corrected cryogenic electron microscopy (cryo-EM) micrograph obtained from a Titan Krios microscope**. b)** A subset of highly populated, reference-free 2D-class averages. **c)** Schematic showing the cryo-EM data processing workflow. Initial processing was

performed using UCSF MotionCor2 and cryoSPARC. Particles were transferred using the pyem script package to RELION for alignment-free 3D classification. Finally, particles were processed in cisTEM using the manual refinement job type with a 7TM mask. Dashed boxes indicated selected classes. **d)** Gold-standard Fourier Shell Correlation (GSFSC) curve for final full-particle map computed in cryoSPARC. **e)** Euler angle distribution of final full-particle map computed in cryoSPARC. **f)** Side view and top view of local resolution for the final full-particle map of GPR68-G_q pH 6 complex computed with local resolution in cryoSPARC. **g)** Side view and top view of local resolution for the focused 7TM map of GPR68-G_q pH 6 complex computed with local resolution in cryoSPARC.

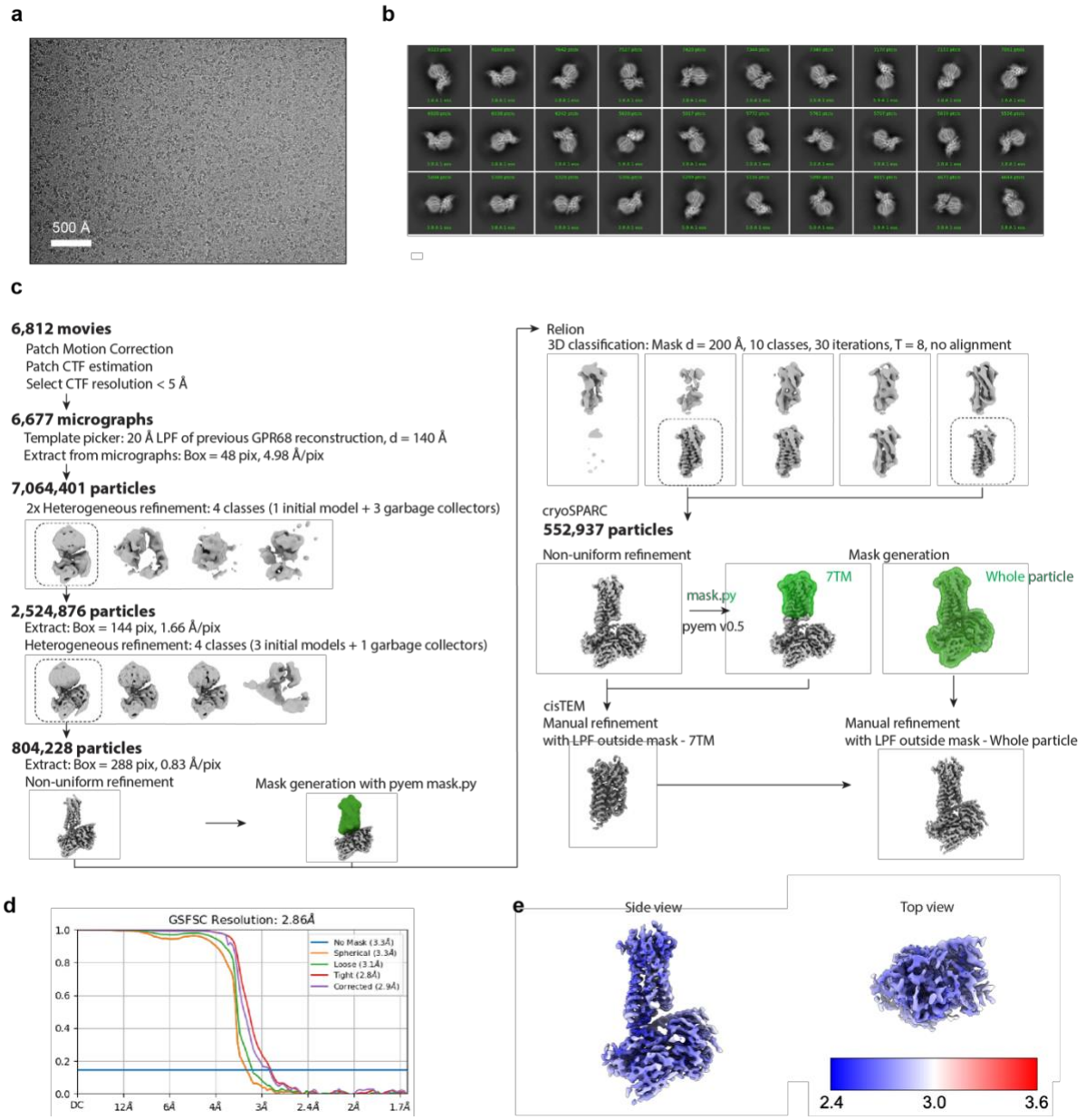


Figure 2.13 Cryogenic electron microscopy processing of GPR68 G_s pH 6

a) A representative motion-corrected cryogenic electron microscopy (cryo-EM) micrograph obtained from a Titan Krios microscope**. b)** A subset of highly populated, reference-free 2D-class averages. **c)** Schematic showing the cryo-EM data processing workflow. Initial processing was performed using UCSF MotionCor2 and cryoSPARC. Particles were transferred using the pyem script package to RELION for alignment-free 3D classification. Finally, particles were processed in cisTEM using the manual refinement job type with a 7TM mask followed by a full particle mask. Dashed boxes indicated selected classes. **d)** Gold-standard Fourier Shell Correlation (GSFSC) curve for final full-particle map computed in cryoSPARC. **e)** Side view and top view of local resolution for the final full-particle map of GPR68-G_s pH 6 complex computed with local resolution in cryoSPARC.

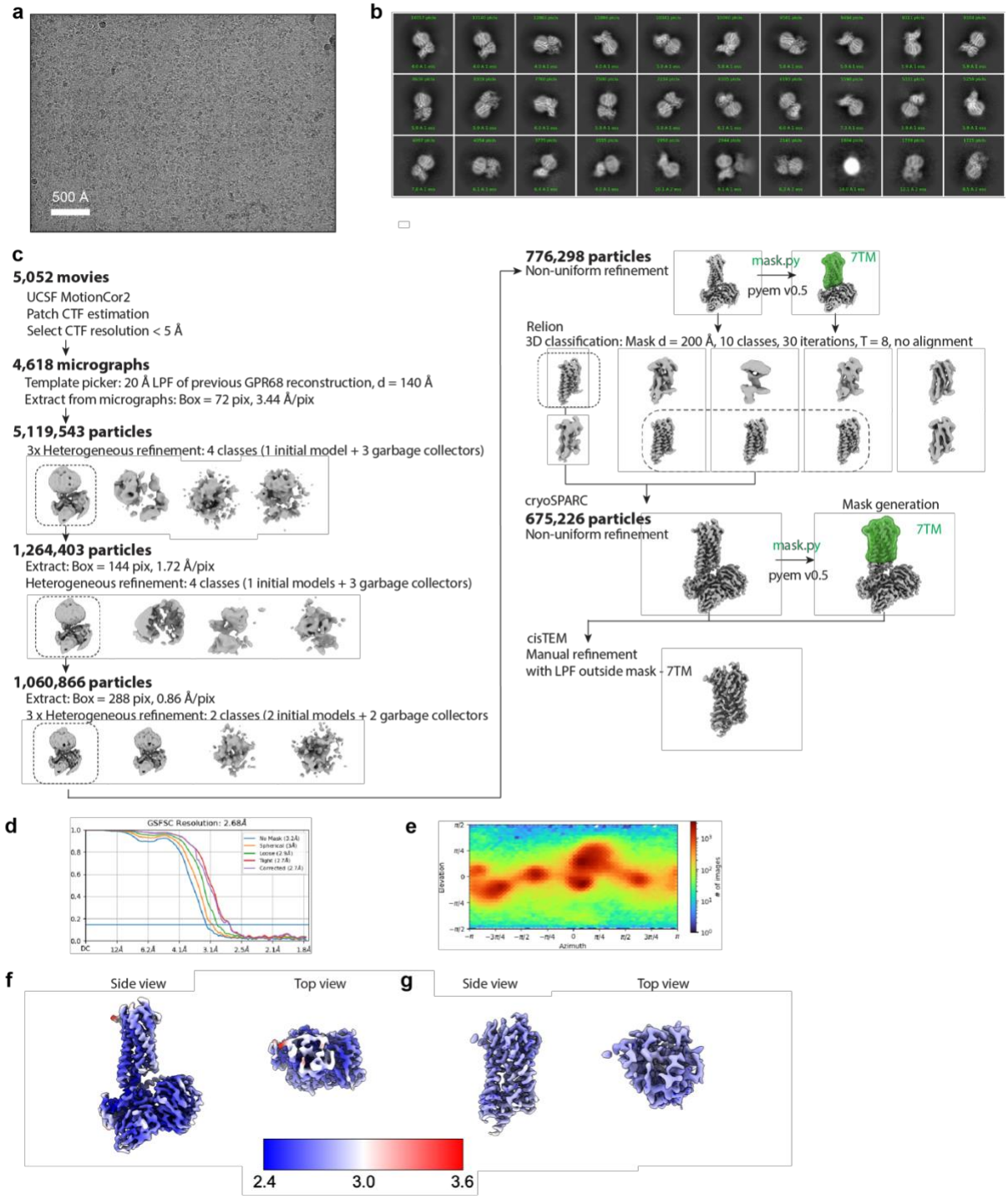


Figure 2.14 Cryogenic electron microscopy processing of GPR68 G_q pH 6

a) A representative motion-corrected cryogenic electron microscopy (cryo-EM) micrograph obtained from a Titan Krios microscope**. b)** A subset of highly populated, reference-free 2D-class averages. **c)** Schematic showing the cryo-EM data processing workflow. Initial processing was performed using UCSF MotionCor2 and cryoSPARC. Particles were transferred using the pyem script package to RELION for alignment-free 3D classification. Finally, particles were processed in cisTEM using the manual refinement job type with a 7TM mask. Dashed boxes indicated selected classes. **d)** Gold-standard Fourier Shell Correlation (GSFSC) curve for final full-particle

map computed in cryoSPARC. **e)** Euler angle distribution of final full-particle map computed in cryoSPARC. **f)** Side view and top view of local resolution for the final full-particle map of GPR68-G_q pH 6 with 10 μM Co^{2+} complex computed with local resolution in cryoSPARC. **g)** Side view and top view of local resolution for the focused 7TM map of GPR68-G_q pH 6 with 10 μM Co^{2+} complex computed with local resolution in cryoSPARC.

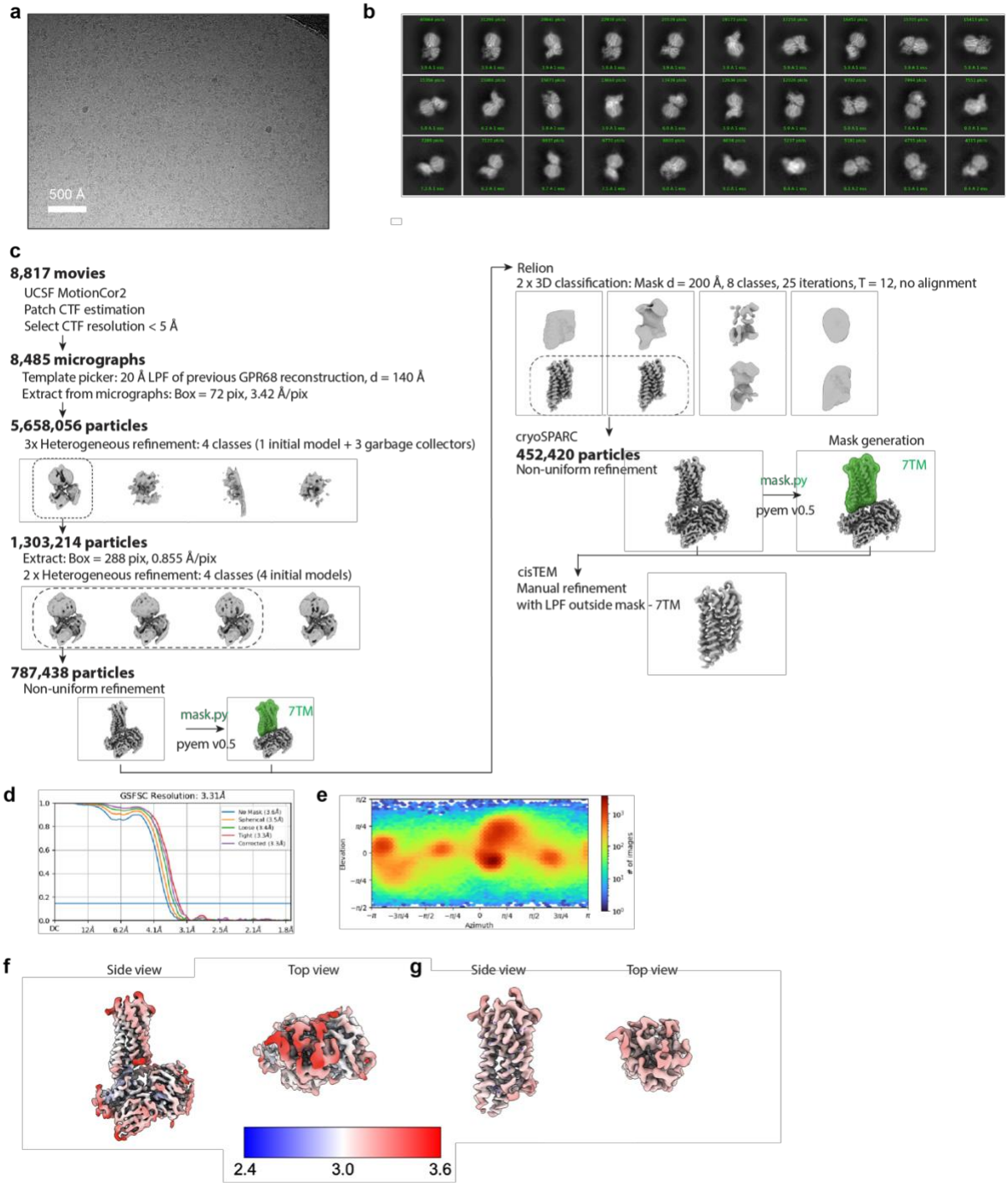


Figure 2.15 Cryogenic electron microscopy processing of GPR68 G_q pH 7.5 with 10 µM Co^{2+}

a) A representative motion-corrected cryogenic electron microscopy (cryo-EM) micrograph obtained from a Titan Krios microscope**. b)** A subset of highly populated, reference-free 2D-class averages. **c)** Schematic showing the cryo-EM data processing workflow. Initial processing was performed using UCSF MotionCor2 and cryoSPARC. Particles were transferred using the pyem script package to RELION for alignment-free 3D classification. Finally, particles were processed in cisTEM using the manual refinement job type with a 7TM mask. Dashed boxes indicated

selected classes. **d)** Gold-standard Fourier Shell Correlation (GSFSC) curve for final full-particle map computed in cryoSPARC. **e)** Euler angle distribution of final full-particle map computed in cryoSPARC. **f)** Side view and top view of local resolution for the final full-particle map of GPR68-G_q pH 7.5 with 10 μM Co^{2+} complex computed with local resolution in cryoSPARC. **g)** Side view and top view of local resolution for the focused 7TM map of GPR68-G_q pH 7.5 with 10 μM Co^{2+} complex computed with local resolution in cryoSPARC.

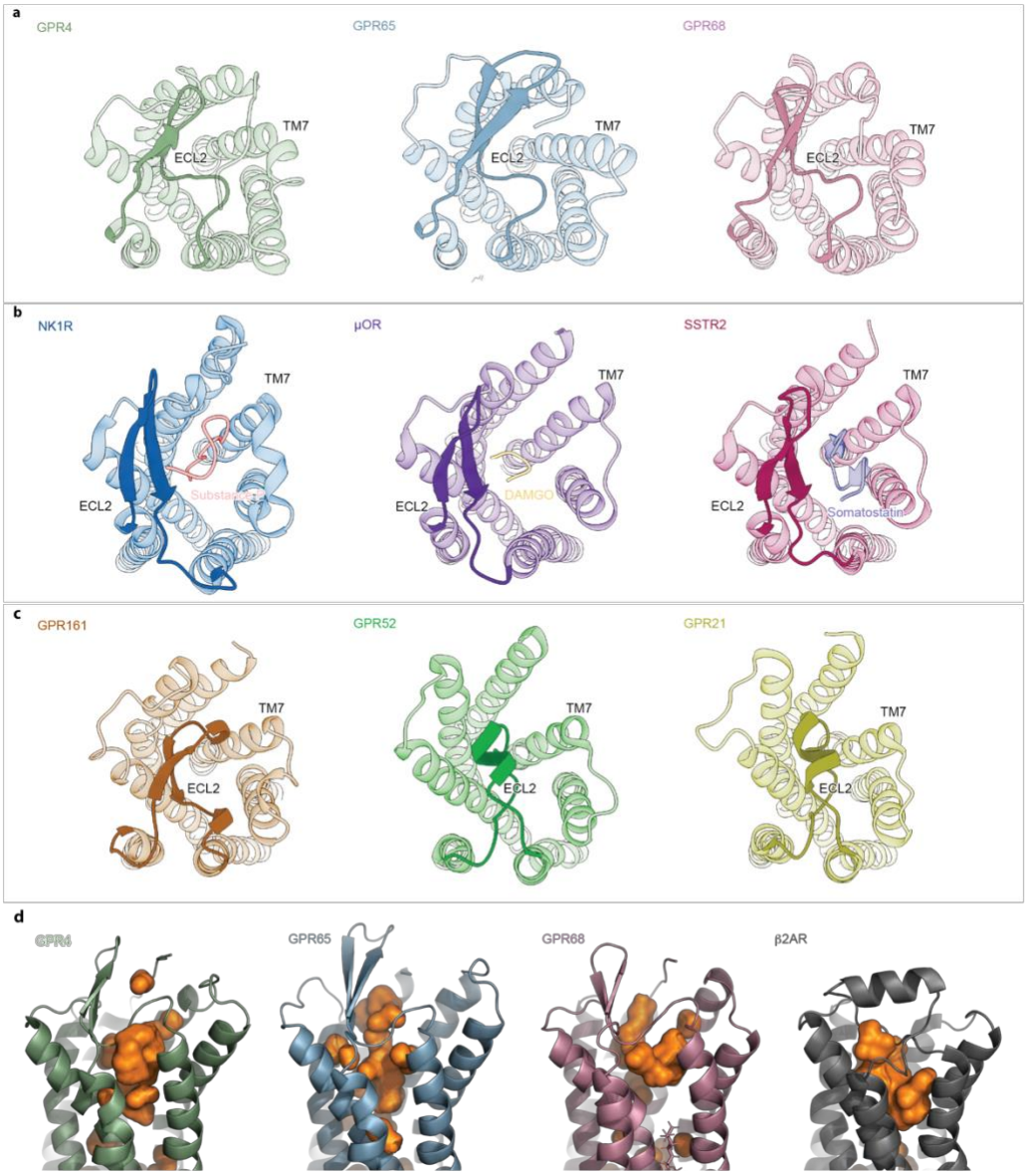


Figure 2.16 Comparison of ECL2 between proton-sensing GPCRs, prototypical peptide-binding GPCRs, and orphan GPCRs

a-c) Top view of GPCR ECL2s with TMs show as transparent helices and ECL2s shown as opaque. **a)** Proton sensor ECL2s have β-hairpin fold that partially occupies canonical ligand binding site, which is a hybrid between prototypical peptide GPCRs and self-activating orphan GPCRs. **b)** Top view of NK1R bound to Substance P (PDB: 7RMG[130]), μOR bound to DAMGO (PDB: 8EFQ[137]), and SSTR2 bound to somatostatin (PDB 7T10[138]). Prototypical Class A peptide-activated GPCRs often have an ECL2 with a β-hairpin and a large open binding pocket with outward displacement of TM7 to accommodate the peptide ligand [132]. **c)** Top view of the self-activating orphan GPCRs, GPR161 (PDB: 8SMV [133]), GPR52 (PDB: 6LI3[126]), and GPR21 (PDB: [127]). The ECL2 of prototypical self-activating orphan GPCRs occlude canonical ligand binding sites, have cis-interactions with TM helices, and no outward deflection of TM7. **d)** The proton-sensing GPCRs have cavities that could be the site of undiscovered endogenous ligand binding. These pockets are highly charged and are comparable in size to aminergic GPCRs like β2AR

(PDB: 4LDO [128]). Cavities calculated with PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

2.6 Tables

Table 2.1 GPR4 mutation parameters

Receptor	Basal RLU	Emax (Fold)	pEC50	Hill
GPR4 WT	213.3 ± 20.5	39.1 ± 3.7	8.01 ± 0.01	4.20 ± 0.24
GPR4 H16D	364.3 ± 114.9	40.4 ± 5.6	7.80 ± 0.01	3.32 ± 0.23
GPR4 H16N	303.1 ± 113.1	51.7 ± 6.6	7.85 ± 0.01	3.88 ± 0.20
GPR4 H16R	204.1 ± 47.5	65.8 ± 10.4	7.75 ± 0.01	3.34 ± 0.15
GPR4 E145Q	943.8 ± 83.6	5.7 ± 0.3	8.27 ± 0.02	3.65 ± 0.35
GPR4 E170A	582.7 ± 78.5	8.2 ± 1.0	8.02 ± 0.04	2.24 ± 0.08
GPR4 E170Q	522.3 ± 66.7	16.4 ± 1.4	7.98 ± 0.00	2.28 ± 0.16
GPR4 R247A	246.1 ± 29.4	21.9 ± 1.9	7.81 ± 0.03	2.53 ± 0.07
GPR4 R247D	241.6 ± 66.9	19.7 ± 7.3	7.49 ± 0.01	1.81 ± 0.17
GPR4 H269D	127.0 ± 20.6	117.4 ± 21.4	7.55 ± 0.03	3.01 ± 0.17
GPR4 H269N	122.6 ± 12.5	111.2 ± 11.0	7.72 ± 0.01	4.21 ± 0.17
GPR4 H269K	181.2 ± 9.1	60.2 ± 3.4	7.85 ± 0.01	3.35 ± 0.12

Table 2.2 GPR65 mutation parameters

Receptor	Basal RLU	Emax (fold)	pEC50	Hill
GPR65	109.3 ± 11.2	79.9 ± 10.5	7.39 ± 0.03	4.31 ± 0.28
GPR65 H13D	180.2 ± 20.0	71.4 ± 4.7	7.12 ± 0.03	2.10 ± 0.08
GPR65 H13N	198.8 ± 17.4	40.6 ± 3.7	7.11 ± 0.03	1.94 ± 0.10
GPR65 H13R	149.4 ± 10.1	46.1 ± 3.7	7.04 ± 0.03	2.22 ± 0.09
GPR65 Y95F	169.5 ± 12.2	57.1 ± 5.1	7.18 ± 0.02	3.37 ± 0.13
GPR65 E142Q	337.0 ± 44.4	25.0 ± 2.8	7.50 ± 0.01	3.68 ± 0.13
GPR65 D172A	149.0 ± 33.7	10.0 ± 2.5	7.25 ± 0.02	2.08 ± 0.11
GPR65 D172N	666.7 ± 299.8	15.6 ± 3.2	6.99 ± 0.03	1.71 ± 0.07
GPR65 R249A	162.9 ± 22.5	39.8 ± 18.3	6.57 ± 0.02	4.88 ± 0.19
GPR65 R249D	86.4 ± 13.7	13.9 ± 3.9	6.97 ± 0.10	1.23 ± 0.07
GPR65 R273D	210.4 ± 24.7	29.6 ± 4.9	6.46 ± 0.03	3.08 ± 0.13
GPR65 R273H	1031.2 ± 142.7	5.6 ± 0.5	7.40 ± 0.04	1.72 ± 0.05
GPR65 R273N	438.9 ± 84.8	16.7 ± 3.6	6.80 ± 0.06	1.69 ± 0.09

Table 2.3 GPR68 mutation parameters

Receptor	Basal RLU	Emax (fold of basal)	pEC50	Hill
WT	95.2 ± 1.2	46.0 ± 1.3	6.66 ± 0.00	4.16 ± 0.07
H20D	83.3 ± 1.9	46.6 ± 1.5	6.52 ± 0.00	3.43 ± 0.04
H20N	77.2 ± 2.8	43.5 ± 2.3	6.45 ± 0.02	2.96 ± 0.03
H20R	80.8 ± 0.8	40.0 ± 1.0	6.55 ± 0.01	3.29 ± 0.09
E149Q	186.6 ± 74.7	27.4 ± 2.6	7.41 ± 0.05	2.60 ± 0.32
E174A	102.0 ± 1.3	23.6 ± 1.3	6.73 ± 0.01	2.76 ± 0.10
E174Q	93.3 ± 4.2	18.0 ± 1.2	6.16 ± 0.01	2.69± 0.05
R251A	113.6 ± 5.2	23.0 ± 1.6	5.84 ± 0.01	2.88 ± 0.12
R251D	116.4 ± 7.0	17.1 ± 1.6	5.77 ± 0.00	3.06 ± 0.01
H269D	104.3 ± 3.1	44.3 ± 1.5	6.34 ± 0.01	2.94 ± 0.06
H269K	115.9 ± 58.8	28.2 ± 7.4	6.87 ± 0.03	2.59 ± 0.11
H269N	99.4 ± 3.0	53.5 ± 3.0	6.46 ± 0.01	2.71 ± 0.09

2.7 Materials and Methods

GloSensor cAMP assays

Proton-sensing GPCR Gs activation and cAMP production were determined using the GloSensor cAMP assay. The following method was adopted from a previously published procedure [129] with minor modifications. In detail, HEK293T cells were maintained and cotransfected with receptor DNA and GloSensor cAMP reporter plasmids in DMEM containing 10% FBS. Overnight transfected cells were plated in PLL coated 384-well white clear-bottom plates in DMEM supplemented with 1% dialyzed fetal bovine serum (dFBS), about 15,000 cells in 40 μL per well, for a minimum of 6 h up to 24 h. To prepare for the assays, cells were removed of culture medium (gently shaking off) and incubated for 1 h at 37 °C with 20 μL per well of loading buffer (1x Calcium- and Magnesium-free HBSS, 20 mM TAPS, pH 8.40) containing 2 mM luciferin and 10 μM PDE inhibitor Ro 20-1724. Buffers were prepared in 1x Calcium- and Magnesium-free HBSS supplemented with different organic buffer agents for different pH ranges, 20 mM MES for pH 5.00–6.60, 20 mM HEPES for pH 6.70–8.20, and 20 mM TAPS for pH 8.30–8.60. Fresh buffers were prepared each day to use within 24 h, and pH was adjusted with KOH at room temperature. PDE inhibitor Ro 20-1724 was added to working solutions to maintain the final concentration of 10 μM. To stimulate cells with desired pH solutions, cells were first removed of luciferin loading solutions (gently shaking off) and added with 25 μL per well of buffer solutions. The cell plate was incubated at room temperature for 20 min before luminescence was counted.

GPR68 deep mutational scanning library generation and cloning

The DIMPLE platform [124] was used to generate a list of oligonucleotide primers to target the wildtype sequence of GPR68 and we followed their protocol for generating the library. We

designed this library to include all missense mutations, single amino acid deletions at each position, and synonymous mutations at each position along the gene. These oligos were ordered as a SurePrint Oligonucleotide library (Agilent Technologies). This DNA was resuspended and the sublibrary fragments were amplified using PrimeStar GXL DNA polymerase and fragragment-specific primers. These reactions were subjected to PCR cleanup using Clean and Concentrate kits (Zymo Research). The sequence of GPR68 WT was synthesized by Twist Bioscience in their High Copy Number Kanamycin backbone, BsmBI and BsaI cutsites were removed. For each library fragment, this plasmid was amplified to add BsaI sites, and the corresponding oligo sublibrary were assembled using BsaI-mediated Golden Gate assembly. These reactions were cleaned and transformed into MegaX DH10B cells and added to 30mL LB+Kan and grown while shaking until they reached OD 0.6. DNA was isolated by midiprep (Zymo Research). Each sublibrary was quantified using Qubit HS and pooled in equimolar ratios. This pooled library was then assembled into our landing pad compatible cAMP reporter vector containing a GSGSGS-P2A-PuroR cassettes for positive selection.

Cell line generation and cell culture

Cell lines for GPR68 WT and the GPR68 mutational library were generated as previously described[124]. The HEK 293T LLP-iCasp9 cells used in this study were a gift from Doug Fowler's lab. To make cell lines, 1ug of DNA was cotransfected with 1ug BxB1 recombinase (pCAG-NLS-BxB1) using 3.75uL lipofectamine 3000 and 5uL P3000 reagent in 6 wells of a 6 well plate. For GPR68 WT, 2 wells were transfected and pooled following selection. For the GPR68 library, 12 wells were transfected in parallel. Cells were cultured in "D10" media (DMEM, 10% dialyzed FBS, 1% sodium pyruvate, and 1% penicillin/streptomycin). The landing pad in the cell line

contains a Tet-on promoter upstream of the BxB1 recombination site and a split rapamycin analog inducible dimerizable Casp-9. Two days after transfection, we induce with doxycycline and treat with AP1903, 10nM. Recombined cells have shifted the iCasp-9 cassette out of frame, recombined cells will express the cassette and upon treatment with AP1903 die from iCasp-9 induced apoptosis. Cells were selected for 2 days in AP1903 after which they were transitioned back to D10 supplemented with doxycycline. After two days of recovery, cells were transitioned to D10 supplemented with both doxycycline and puromycin to select for cells that have proper in-frame and full-length assemblies. Following selection by puromycin for two days, cells were transitioned to D10 and expanded before freezing down or using in subsequent assays.

Flow cytometry, fluorescence activated cell sorting

For flow-based assays and cell sorting, frozen stocks of cells were thawed and allowed to recover for several days in D10 media. 48h prior to starting the experiment, cells were split into an appropriate sized dish such that they reach 75% confluency by the start of the assay. 36h prior to starting the assay, cells were induced with doxycycline. Doxycycline was subsequently washed out after 24h and cells were maintained in D10 for the remaining 12h prior to starting. For the pH and pH + 30uM ogerin conditions, the pH of D10 media was adjusted using HCl on the same day as the assay. The cAMP assay was run as follows: cells were swapped to D10 (at indicated pH) with trimethoprim for 8h. After this incubation, cells were detached using TrpyLE Express, washed, and resuspended in BD FACS buffer. The surface expression assay was run similarly, cells were simply detached using TrypLE after induction, stained with M2 FLAG APC-Surelight antibody (Abcam), washed, and then kept covered on ice until running. Flow cytometry was performed using a Cytoflex, and cell sorting was performed using a Cytoflex SRT. Briefly, cells

were gated on FSC-A and SSC-A to separate HEK293T while cells, then FSC-A and FSC-H to find single cells. For the cAMP assay, we assessed activity using eGFP on the FITC-A channel, and for surface expression assays, the APC-A. For the cAMP sorting experiments, the population was split into four roughly equal populations (% cells) based on the most active condition, pH 5.5 + 30uM Ogerin. These gates were maintained for all subsequent samples. For surface expression assays, the population was largely bimodal, and we gated using the peaks of each distribution and the intervening trough. For sorting experiments we aimed to collect cells equal or greater than 100x the expected number of variants in our library.

Mutational scanning data analysis

Following sorting, genomic DNA was extracted from each sample using a Micro kit from Zymo. We quantified the extracted DNA and used the cell_line_for_5 and P2A_cell_line_rev primers to amplify our gene as previously described[138]. PCR reactions were cleaned and concentrated and run on an agarose gel. Bands corresponding to our amplicon were extracted and purified. Amplicons were quantified using Qubit HS and 1ng of each was prepared for sequencing using the Nextera XT DNA kit from Illumina. Samples were indexes using the IDT for Illumina UD indices set C. SPRI select beads at a 0.9x ratio were used for cleanup and size selection of the final libraries, and the tagmented library was analyzed using a D5000 HS kit on an Agilent Tapestation. After confirming the libraries were properly tagmented, they were quantified using Qubit HS and then pooled for sequencing on a NovaSeq 6000 SP flow cell in pained end mode.

Expression and purification of proton sensor active-state complexes

The human *GPR4, GPR65,* and *GPR68* genes with an N-terminal influenza hemagglutinin signal sequence and Flag epitope tag were cloned into a pcDNA3.1/Zeo vector containing a tetracycline inducible cassette. The miniG proteins ($miniG_{s399}$ for GPR4 and GPR65 and $miniG_{s/q70}$ for GPR68[33]) were fused to the C terminus of each proton sensor preceded by a glycine/serine linker and rhinovirus 3C protease recognition site. The resulting fusion constructs were transfected into inducible Expi293F-TetR cells (Thermo Fisher) using the ExpiFectamine transfection reagent per manufacturer instructions. After 18 h, protein expression was induced with 1 µg/mL doxycycline hyclate for 24 h before collection by centrifugation. Pelleted cells were washed with 50 mL phosphate buffered saline, pH 7.5 before storage at −80 °C. For receptor purification, frozen cells were hypotonically lysed in 20 mM MES, pH 6, 1 mM EDTA, 160 µg/mL benzamidine, 2 µg/mL leupeptin for 10 min at 25 °C. The membrane fraction was collected by centrifugation, and the fusion protein was extracted with 20 mM MES, pH 6, 300 mM NaCl, 1% (w/v) lauryl maltose neopentyl glycol (L-MNG, Anatrace), 0.1% (w/v) cholesteryl hemisuccinate (CHS, Steraloids), 2 mM $MgCl_2$, 2 mM $CaCl_2$, 160 µg/mL benzamidine, 2 µg/mL leupeptin with dounce homogenization and incubation with stirring for one hour at 4 °C. The soluble fraction was separated from the insoluble fraction by centrifugation and was incubated in batch for 1 h at 4 °C with homemade M1–Flag antibody-conjugated Sepharose beads. Sepharose resin was then washed extensively with 20 mM MES, pH 6, 150 mM NaCl, 0.1% (w/v) L-MNG, 0.01% (w/v) CHS, 2 mM $MgCl_2$, 2 mM $CaCl_2$ and then with 20 mM MES, pH 6, 150 mM NaCl, 0.0075% (w/v) L-MNG, 0.00075% (w/v) CHS, 2 mM $MgCl_2$, 2 mM $CaCl_2$ prior to elution with 20 mM MES, pH 6, 150 mM NaCl, 0.0075% (w/v) L-MNG, 0.00075% (w/v) CHS, 5 mM EDTA, 0.2 mg/mL Flag peptide. Eluted protein was concentrated in a 100 kDa MWCO Amicon spin concentrator, and

injected onto a Superdex200 Increase 10/300GL (Cytiva) gel filtration column equilibrated in 20 mM MES, pH 6, 150 mM NaCl, 0.0075% (w/v) L-MNG, 0.0025% glyco-diosgenin (GDN, Anatrace), and 0.0005% CHS. Monodisperse fractions were complexed with $G_{\beta 1\gamma 2}$ heterodimer and Nb35 at 2 molar excess overnight at 4˚C. The next day, the heterotrimeric complex was concentrated with a 100 kDa MWCO spin concentrator and excess $G_{\beta 1\gamma 2}$ and Nb35 was removed via size-exclusion chromatography, using a Superdex200 Increase 10/300 GL column (GE Healthcare) equilibrated in 20 mM MES pH 6, 150 mM NaCl, 0.00075% (w/v) L-MNG, 0.00025% (w/v) GDN, and 0.0001% CHS. Resulting heterotrimeric complex was concentrated with a 100 kDa MWCO spin concentrator for preparation of cryo-EM grids. For GPR68 structures with Co^{2+}, 10 μM Co^{2+} was added to all buffers. For GPR68 structure at pH 7.5, 20 mM HEPES pH 7.5 was substituted for 20 mM MES pH 6.

Expression and purification of $G_{\beta 1\gamma 2}$

Human $G_{\beta 1\gamma 2}$ heterodimer was expressed in *Trichoplusia ni* Hi5 insect cells (Expression Systems) using a single baculovirus generated in *Spodoptera frugiperda* Sf9 insect cells (Expression Systems). A bicistronic pVLDual construct contained the $G_{\beta 1}$ subunit with a N-terminal 6 × His tag, and an untagged human $G_{\gamma 2}$ subunit. For expression, Hi5 insect cells were transduced with baculovirus at a density of ~3.0 × 10^6 cells per mL, grown with 27 °C shaking at 130 rpm. 48 h post-transduction, cells were collected and washed in a hypotonic buffer containing 20 mM HEPES, pH 8.0, 5 mM β-mercaptoethanol (β-ME), and protease inhibitors (20 μg/mL leupeptin, 160 μg/mL benzamidine). The membrane fraction was then separated by centrifugation and solubilized with 20 mM HEPES pH 8.0, 100 mM sodium chloride, 1.0% sodium cholate, 0.05% dodecylmaltoside (Anatrace), and 5 mM β-mercaptoethanol (β-ME). Solubilized $G_{\beta 1\gamma 2}$

heterodimer was then incubated with HisPur Ni-NTA resin (Thermo Scientific) in batch. Bound $G_{\beta 1\gamma 2}$ heterodimer was washed extensively and detergent was slowly exchanged to 0.1% (w/v) lauryl maltose neopentyl glycol (L-MNG, Anatrace) and 0.01% CHS before elution with 20 mM HEPES pH 7.5, 100 mM NaCl, 0.1% L-MNG, 0.01% CHS, 270 mM imidazole, 1 mM dithiothreitol (DTT), and protease inhibitors. Eluted $G_{\beta 1\gamma 2}$ heterodimer was pooled and rhinovirus 3C protease was added to cleave the N-terminal 6 × His tag during overnight dialysis in 20 mM HEPES pH 7.5, 100 mM NaCl, 0.02% L-MNG, 0.002% CHS, 1 mM DTT, and 10 mM imidazole. To remove uncleaved $G_{\beta 1\gamma 2}$, dialysed material was incubated with HisPur Ni-NTA resin in batch. The unbound fraction was then incubated for 1 h at 4 °C with lambda phosphatase (New England Biolabs), calf intestinal phosphatase (New England Biolabs), and Antarctic phosphatase (New England Biolabs) for dephosphorylation. Final anion exchange chromatography was performed using a MonoQ 4.6/100 PE (Cytiva) column to purify only geranylgeranylated heterodimer. The resulting protein was pooled and dialysed overnight in 20 mM HEPES pH 7.5, 100 mM NaCl, 0.02% L-MNG, and 100 µM TCEP, and concentrated with a 3 kDa centrifugal concentrator to a final concentration of 162 µM. Glycerol was added to a final concentration of 20%, and the protein was flash frozen in liquid nitrogen and stored at −80 °C until further use.

Expression and purification of Nb35

A pET-26b vector containing the Nb35 sequence with a carboxy-terminal Protein C affinity tag was transformed into BL21 Rosetta *Escherichia coli* cells (UC Berkeley QB3 MacroLab) and inoculated into 8 L of Terrific Broth supplemented with 0.1% glucose, 2 mM $MgCl_2$, and 50 µg/mL kanamycin. Cells were induced with 400 µM IPTG at *A*600 of 0.6 and allowed to express at 20 °C for 21 h. Collected cells were incubated SET Buffer (200 mM Tris pH 8.0, 500 mM sucrose, 0.5 mM EDTA) in the presence of protease inhibitors (20 µg/mL leupeptin,

160 µg/mL benzamidine) and benzonase. To initiate hypotonic lysis, two volumes of deionized water were added to the cell mixture after 30 min of SET buffer mixing. Following lysis, NaCl was added to 150 mM, $CaCl_2$ was added to 2 mM, and $MgCl_2$ was added to 2 mM and lysate was centrifuged to remove the insoluble fraction. Supernatant was incubated with homemade anti-Protein C antibody-coupled Sepharose. Nb35 was eluted with 20 mM HEPES pH 7.5, 100 mM NaCl, and 2 mM $CaCl_2$, 0.2 mg/mL protein C-peptide, and 5 mM EDTA pH 8.0, concentrated in a 10 kDa MWCO Amicon filter and injected over a Superdex S75 Increase 10/300 GL column (Cytiva) size-exclusion chromatography column equilibrated in 20 mM HEPES pH 7.5, 100 mM NaCl. Monodisperse Nb35 fractions were pooled, concentrated, and flash frozen in liquid nitrogen for storage at −80 °C until further use.

Cryo-EM vitrification, data collection and processing

GPR4-G_s pH 6 complex

The GPR4-G_s pH 6 complex was concentrated to 14 mg/mL supplemented with 0.05% CHAPS (Thermo Fisher) and 3 µL was applied onto a glow-discharged 300 mesh 1.2/1.3 gold grid covered in a holey gold film (UltrAufoil). Excess sample was removed with a blotting time of 4 s and a blotting force of 1 at 4 °C prior to plunge freezing into liquid ethane using a Vitrobot Mark IV (Thermo Fisher). A total of 9,018 movies were recorded with a K3 detector (Gatan) on a Titan Krios (Thermo Fisher) microscope operated at 300 keV with a BioQuantum post-column energy filter set to a zero-loss energy selection slit width set of 20 eV. Movies were recorded using dose-fractionated illumination at a nominal magnification of 86,000x (physical pixel size of 0.86 Å/pixel) and a defocus range of -1 to -2.1 µm for a total dose of 50.7 $e^-/Å^2$. Exposure areas were acquired with image shift collection using EPU (Thermo Fisher).

Movies of the GPR4-G_s pH 6 complex were motion-corrected and dose-fractionated using UCSF MotionCor2 [75]. Corrected micrographs were imported into cryoSPARC v3.1 [76] for CTF estimation via the Patch Estimation job. Micrographs with estimated CTF fit resolution > 5 Å were removed before further processing. Templates for particle picking were generated from the same complex reconstructed from a previous 200 keV imaging session. Particle picking templates were low-pass filtered to 20 Å and used to pick 8,608,607 particles. After picking, particles were extracted in a 288 pixel box and Fourier cropped to 48 pixels before 3D classification with alignment using a 20 Å low-pass filtered reconstruction and three random reconstructures generated from a prematurely truncated ab initio reconstruction job, called "garbage collectors," with the Heterogeneous Refinement job type. Two rounds of Heterogeneous Refinement yielded 2,501,915 particles that were re-extracted in the same box size cropped to 72 pixels and classified in a third Heterogeneous Refinement job. The resulting 1,453,906 particles were re-extracted in the same box cropped to 144 pixels. A fourth round of Heterogeneous Refinement and 2D classification, yielded 878,077 particles that were extracted without cropping. A final round of Heterogeneous Refinement yielded 439,296 particles that were refined using the Non-Uniform Refinement job type giving the final full-particle map. Finally, local refinement using an inclusion mask covering the 7TM domain was performed, using poses/shift Gaussian priors with standard deviation of rotational and shift magnitudes limited to 3° and 2 Å, respectively.

GPR65-G_s pH 6 complex

The GPR65-G_s pH 6 complex was concentrated to 11 mg/mL supplemented with 0.05% CHAPS (Thermo Fisher) and 3 μL was applied onto a glow-discharged 300 mesh 1.2/1.3 gold grid covered in a holey gold film (UltrAufoil). Excess sample was removed with a blotting time of 4 s and a

blotting force of 1 at 4 °C prior to plunge freezing into liquid ethane using a Vitrobot Mark IV (Thermo Fisher). A total of 8,294 movies were recorded with a K3 detector (Gatan) on a Titan Krios (Thermo Fisher) microscope operated at 300 keV with a BioQuantum post-column energy filter set to a zero-loss energy selection slit width set of 20 eV. Movies were recorded using dose-fractionated illumination at a nominal magnification of 105,000x (physical pixel size of 0.81 Å/pixel) and a defocus range of -1 to -2.1 µm for a total dose of 46 $e^-/Å^2$. Exposure areas were acquired with image shift collection using SerialEM 3.8[43].

Movies of the GPR65-G_s pH 6 complex were motion-corrected and dose-fractionated using UCSF MotionCor2 [75]. Corrected micrographs were imported into cryoSPARC v3.1 [76] for CTF estimation via the Patch Estimation job. Micrographs with estimated CTF fit resolution > 5 Å were removed before further processing. Templates for particle picking were generated from the same complex reconstructed from a previous 200 keV imaging session. Particle picking templates were low-pass filtered to 20 Å and used to pick 8,673,428 particles. After picking, particles were extracted in a 288 pixel box and Fourier cropped to 48 pixels before 3D classification with alignment using a 20 Å low-pass filtered reconstruction and "garbage collectors" with the Heterogeneous Refinement job type. Two rounds of Heterogeneous Refinement yielded 2,588,765 particles that were re-extracted in the same box size cropped to 74 pixels and classified in two Heterogeneous Refinement jobs. The resulting 1,637,819 particles were re-extracted in the same box cropped to 150 pixels and further classified with two rounds of Heterogeneous Refinement and 2D classification. The resulting 1,055,443 particles were refined using the Non-Uniform Refinement job type. Particles were exported using csparc2star.py from the pyem script package[77], and a mask covering the 7TM domain of GPR65 was generated using the Segger tool in UCSF ChimeraX[78] and the Volume Tools utility in cryoSPARC. The particles and mask were imported into Relion

v3.0[79] and classified in 3D without alignment through three separate iterations. Particles comprising the three highest resolution classes were reimported into cryoSPARC for Non-Uniform Refinement. Finally, particles were exported into cisTEM[80] for 7TM local refinements using the Manual Refinement job type and low-pass filtering outside of the mask.

GPR68-G_q pH 6 complex

The GPR68-G_q pH 6 complex was concentrated to 4 mg/mL and 3 µL was applied onto a glow-discharged 300 mesh 1.2/1.3 gold grid covered in a holey carbon film (Quantifoil). Excess sample was removed with a blotting time of 4 s and a blotting force of 1 at 4 °C prior to plunge freezing into liquid ethane using a Vitrobot Mark IV (Thermo Fisher). A total of 6,650 movies were recorded with a K3 detector (Gatan) on a Titan Krios (Thermo Fisher) microscope operated at 300 keV with a BioQuantum post-column energy filter set to a zero-loss energy selection slit width set of 20 eV. Movies were recorded using dose-fractionated illumination at a nominal magnification of 105,000x (physical pixel size of 0.855 Å/pixel) and a defocus range of -1 to -2.1 µm for a total dose of 50 $e^-/Å^2$. Exposure areas were acquired with image shift collection using EPU (Thermo Fisher).

Movies of the GPR68-G_q pH 6 complex were motion-corrected and dose-fractionated using UCSF MotionCor2 [75]. Corrected micrographs were imported into cryoSPARC v3.1 [76] for CTF estimation via the Patch Estimation job. Micrographs with estimated CTF fit resolution > 5 Å were removed before further processing. Templates for particle picking were generated from the same complex reconstructed from a previous 200 keV imaging session. Particle picking templates were low-pass filtered to 20 Å and used to pick 6,764,523 particles. After picking, particles were extracted in a 288 pixel box and Fourier cropped to 72 pixels before 3D classification with alignment using a 20

Å low-pass filtered reconstruction and "garbage collectors" with the Heterogeneous Refinement job type. Two rounds of Heterogeneous Refinement yielded 2,774,555 particles that were re-extracted in the same box size cropped to 192 pixels and classified in an additional Heterogeneous Refinement job. The resulting 1,144,750 particles were refined using the Non-Uniform Refinement job type. Particles were exported using csparc2star.py from the pyem script package[77], and a mask covering the 7TM domain of GPR68 was generated using the Segger tool in UCSF ChimeraX[78] and the mask.py pyem script. The particles and mask were imported into Relion v3.0[79] and classified in 3D without alignment. Particles comprising the highest resolution class were reimported into cryoSPARC for Non-Uniform Refinement. Finally, particles were exported into cisTEM[80] for 7TM local refinements using the Manual Refinement job type and low-pass filtering outside of the mask.

GPR68-G_s pH 6 complex

The GPR68-G_s pH 6 complex was concentrated to 4 mg/mL and 3 µL was applied onto a glow-discharged 300 mesh 1.2/1.3 gold grid covered in a holey carbon film (Quantifoil). Excess sample was removed with a blotting time of 4 s and a blotting force of 1 at 4 °C prior to plunge freezing into liquid ethane using a Vitrobot Mark IV (Thermo Fisher). A total of 6,812 movies were recorded with a K3 detector (Gatan) on a Titan Krios (Thermo Fisher) microscope operated at 300 keV with a BioQuantum post-column energy filter set to a zero-loss energy selection slit width set of 20 eV. Movies were recorded using dose-fractionated illumination at a nominal magnification of 105,000x (physical pixel size of 0.83 Å/pixel) and a defocus range of -1 to -2.1 µm for a total dose of 49 $e^-/Å^2$. Exposure areas were acquired with image shift collection using SerialEM 3.8[43].

Movies of the GPR68-G_s pH 6 complex were imported into cryoSPARC v3.1 [76] for motion-correction, dose-fractionation, and CTF estimation. Micrographs with estimated CTF fit resolution > 5 Å were removed before further processing. Templates for particle picking were generated from the same complex reconstructed from a previous 200 keV imaging session. Particle picking templates were low-pass filtered to 20 Å and used to pick 7,064,401 particles. After picking, particles were extracted in a 288 pixel box and Fourier cropped to 48 pixels before 3D classification with alignment using a 20 Å low-pass filtered reconstruction and "garbage collectors" with the Heterogeneous Refinement job type. Two rounds of Heterogeneous Refinement yielded 2,524,876 particles that were re-extracted in the same box size cropped to 144 pixels and classified in an Heterogeneous Refinement job. The resulting 804,228 particles were refined using the Non-Uniform Refinement job type. Particles were exported using csparc2star.py from the pyem script package[77], and a mask covering the 7TM domain of GPR68 was generated using the Segger tool in UCSF ChimeraX[78] and the mask.py pyem script. The particles and mask were imported into Relion v3.0[79] and classified in 3D without alignment. Particles comprising the highest resolution classes were reimported into cryoSPARC for Non-Uniform Refinement. Finally, particles were exported into cisTEM[80] for two local refinements using the Manual Refinement job type and low-pass filtering outside of masks. In the first local refinement, the previous 7TM mask was used, and the second local refinement used a full-particle mask.

GPR68-G_q pH 6 Co^{2+} complex

The GPR68-G_q pH 6 Co^{2+} complex was concentrated to 4 mg/mL and 3 µL was applied onto a glow-discharged 300 mesh 1.2/1.3 gold grid covered in a holey carbon film (Quantifoil). Excess sample was removed with a blotting time of 4 s and a blotting force of 1 at 4 °C prior to plunge freezing into liquid ethane using a Vitrobot Mark IV (Thermo Fisher). A total of 5,052 movies

were recorded with a K3 detector (Gatan) on a Titan Krios (Thermo Fisher) microscope operated at 300 keV with a BioQuantum post-column energy filter set to a zero-loss energy selection slit width set of 20 eV. Movies were recorded using dose-fractionated illumination at a nominal magnification of 105,000x (physical pixel size of 0.86 Å/pixel) and a defocus range of -1 to -2.1 µm for a total dose of 50 $e^-/Å^2$. Exposure areas were acquired with image shift collection using EPU (Thermo Fisher).

Movies of the GPR68-G_q pH 6 Co^{2+} complex were motion-corrected and dose-fractionated using UCSF MotionCor2 [75]. Corrected micrographs were imported into cryoSPARC v3.1 [76] for CTF estimation via the Patch Estimation job. Micrographs with estimated CTF fit resolution > 5 Å were removed before further processing. Templates for particle picking were generated from the same complex reconstructed from a previous 200 keV imaging session. Particle picking templates were low-pass filtered to 20 Å and used to pick 5,119,543 particles. After picking, particles were extracted in a 288 pixel box and Fourier cropped to 72 pixels before 3D classification with alignment using a 20 Å low-pass filtered reconstruction and "garbage collectors" with the Heterogeneous Refinement job type. Three rounds of Heterogeneous Refinement yielded 1,264,403 particles that were re-extracted in the same box size cropped to 144 pixels and classified in an additional Heterogeneous Refinement job. The resulting 1,060,866 particles were extracted without cropping and classified in three Heterogeneous Refinement jobs. The resulting 776,298 particles were refined using the Non-Uniform Refinement job type. Particles were exported using csparc2star.py from the pyem script package[77], and a mask covering the 7TM domain of GPR68 was generated using the Segger tool in UCSF ChimeraX[78] and the mask.py pyem script. The particles and mask were imported into Relion v3.0[79] and classified in 3D without alignment. Particles comprising the highest resolution classes were reimported into cryoSPARC for Non-

Uniform Refinement. Finally, particles were exported into cisTEM[80] for 7TM local refinements using the Manual Refinement job type and low-pass filtering outside of the mask.

GPR68-G_q pH 7.5 Co^{2+} complex

The GPR68-G_q pH 7.5 Co^{2+} complex was concentrated to 2 mg/mL and 3 µL was applied onto a glow-discharged 300 mesh 1.2/1.3 gold grid covered in a holey carbon film (Quantifoil). Excess sample was removed with a blotting time of 4 s and a blotting force of 1 at 4 °C prior to plunge freezing into liquid ethane using a Vitrobot Mark IV (Thermo Fisher). A total of 8,817 movies were recorded with a K3 detector (Gatan) on a Titan Krios (Thermo Fisher) microscope operated at 300 keV with a BioQuantum post-column energy filter set to a zero-loss energy selection slit width set of 20 eV. Movies were recorded using dose-fractionated illumination at a nominal magnification of 105,000x (physical pixel size of 0.855 Å/pixel) and a defocus range of -1 to -2.1 µm for a total dose of 60 $e^-/Å^2$. Exposure areas were acquired with image shift collection using EPU (Thermo Fisher).

Movies of the GPR68-G_q pH 7.5 Co^{2+} complex were motion-corrected and dose-fractionated using UCSF MotionCor2 [75]. Corrected micrographs were imported into cryoSPARC v3.1 [76] for CTF estimation via the Patch Estimation job. Micrographs with estimated CTF fit resolution > 5 Å were removed before further processing. Templates for particle picking were generated from the same complex reconstructed from a previous 200 keV imaging session. Particle picking templates were low-pass filtered to 20 Å and used to pick 5,658,056 particles. After picking, particles were extracted in a 288 pixel box and Fourier cropped to 72 pixels before 3D classification with alignment using a 20 Å low-pass filtered reconstruction and “garbage collectors” with the Heterogeneous Refinement job type. Three rounds of Heterogeneous Refinement yielded

1,303,214 particles that were re-extracted in the same box size without cropping to 144 and classified in two additional Heterogeneous Refinement jobs. The resulting 787,438 particles were refined using the Non-Uniform Refinement job type. Particles were exported using csparc2star.py from the pyem script package[77], and a mask covering the 7TM domain of GPR68 was generated using the Segger tool in UCSF ChimeraX[78] and the mask.py pyem script. The particles and mask were imported into Relion v3.0[79] and classified in 3D without alignment. Particles comprising the highest resolution classes were reimported into cryoSPARC for Non-Uniform Refinement. Finally, particles were exported into cisTEM[80] for 7TM local refinements using the Manual Refinement job type and low-pass filtering outside of the mask.

Model building and refinement

Model building and refinement began with the Alphafold2[81] predicted structures as the starting models, which were fitted into the experimental cryoEM maps using UCSF ChimeraX. The model was iteratively refined with real space refinement in Phenix[82] and manually in Coot[83] and Isolde[51]. The cholesteryl hemisuccinate model and rotamer library were generated with the PRODRG server[84], docked using Coot, and refined in Phenix and Isolde. Final map-model validations were carried out using Molprobity and EMRinger in Phenix.

2.8 Acknowledgements

This work was supported by the National Institutes of Health (NIH) Ruth L. Kirschstein Predoctoral Fellowship F31HL164045 (N.H.). Cryo-EM equipment at UCSF is partially supported by NIH grants S10OD020054 and S10OD021741. Some of this work was performed national electron microscopy facilities including: Stanford-SLAC Cryo-EM Center (S2C2), which is

supported by the National Institutes of Health Common Fund Transformative High-Resolution Cryo-Electron Microscopy program (U24 GM129541), National Cancer Institute Cryo-Electron Microscopy Facility, and Pacific Northwest Center for Cryo-EM. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. A.M. acknowledges support from the Edward Mallinckrodt, Jr. Foundation and the Vallee Foundation. A.M. is a Chan Zuckerberg Biohub Investigator.

2.9 Author Contributions

N.H. cloned, expressed, and biochemically optimized the purification of proton sensor constructs for structural studies. N.H. performed cryo-EM data collection, with help from cryo-EM facilites, and data processing. N.H. built and refined models of the proton sensors. N.H., M.H, and X.H. generated receptor constructs and determined expression levels by flow cytometry and performed signaling studies, and analyzed the data. M.H. cloned, performed and analyzed deep-mutational scans under the supervision of W.C.M. All authors contributed to figures. N.H., M.H., and A.M., wrote the manuscript, with edits and approval from all authors. A.M. supervised the overall project.

2.10 References

1. Dixon, R. A. *et al.* Cloning of the gene and cDNA for mammalian beta-adrenergic receptor and homology with rhodopsin. *Nature* **321**, 75–79 (1986).
2. Lodes, M. J., Dillon, D. C., Houghton, R. L. & Skeiky, Y. A. W. Expression cloning. *Methods Mol. Med.* **94**, 91–106 (2004).
3. Libert, F., Vassart, G. & Parmentier, M. Current developments in G-protein-coupled receptors. *Curr. Opin. Cell Biol.* **3**, 218–223 (1991).
4. Thorens, B. Expression cloning of the pancreatic beta cell receptor for the gluco-incretin hormone glucagon-like peptide 1. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 8641–8645 (1992).
5. Parmentier, M. *et al.* Molecular cloning of the thyrotropin receptor. *Science* **246**, 1620–1622 (1989).
6. Sakurai, T. *et al.* Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. *Cell* **92**, 573–585 (1998).
7. Laschet, C., Dupuis, N. & Hanson, J. The G protein-coupled receptors deorphanization landscape. *Biochem. Pharmacol.* **153**, 62–74 (2018).
8. Mukhopadhyay, S. *et al.* The Ciliary G-Protein-Coupled Receptor Gpr161 Negatively Regulates the Sonic Hedgehog Pathway via cAMP Signaling. *Cell* **152**, 210–223 (2013).
9. Kim, S.-E. *et al.* Dominant negative GPR161 rare variants are risk factors of human spina bifida. *Hum. Mol. Genet.* **28**, 200–208 (2019).
10. Begemann, M. *et al.* Germline GPR161 Mutations Predispose to Pediatric Medulloblastoma. *J. Clin. Oncol.* **38**, 43–50 (2020).
11. Feigin, M. E., Xue, B., Hammell, M. C. & Muthuswamy, S. K. G-protein–coupled receptor GPR161 is overexpressed in breast cancer and is a promoter of cell proliferation and

invasion. *Proceedings of the National Academy of Sciences* **111**, 4191–4196 (2014).

12. Civelli, O., Saito, Y., Wang, Z., Nothacker, H.-P. & Reinscheid, R. K. Orphan GPCRs and their ligands. *Pharmacol. Ther.* **110**, 525–532 (2006).
13. Hwang, S.-H., Somatilaka, B. N., White, K. & Mukhopadhyay, S. Ciliary and extraciliary Gpr161 pools repress hedgehog signaling in a tissue-specific manner. *Elife* **10**, (2021).
14. Shimada, I. S. *et al.* Basal Suppression of the Sonic Hedgehog Pathway by the G-Protein-Coupled Receptor Gpr161 Restricts Medulloblastoma Pathogenesis. *Cell Rep.* **22**, 1169–1184 (2018).
15. Hwang, S.-H. *et al.* The G protein-coupled receptor Gpr161 regulates forelimb formation, limb patterning and skeletal morphogenesis in a primary cilium-dependent manner. *Development* **145**, (2018).
16. Kim, S.-E. *et al.* Wnt1 Lineage Specific Deletion of Gpr161 Results in Embryonic Midbrain Malformation and Failure of Craniofacial Skeletal Development. *Front. Genet.* **12**, 761418 (2021).
17. Shimada, I. S. *et al.* Derepression of sonic hedgehog signaling upon Gpr161 deletion unravels forebrain and ventricular abnormalities. *Dev. Biol.* **450**, 47–62 (2019).
18. Li, B. I. *et al.* The orphan GPCR, Gpr161, regulates the retinoic acid and canonical Wnt pathways during neurulation. *Dev. Biol.* **402**, 17–31 (2015).
19. Matteson, P. G. *et al.* The orphan G protein-coupled receptor, Gpr161, encodes the vacuolated lens locus and controls neurulation and lens development. *Proceedings of the National Academy of Sciences* **105**, 2088–2093 (2008).
20. Karaca, E. *et al.* Whole-exome sequencing identifies homozygous GPR161 mutation in a family with pituitary stalk interruption syndrome. *J. Clin. Endocrinol. Metab.* **100**, (2015).

21. Website. https://bpspubs.onlinelibrary.wiley.com/doi/10.1111/bph.16053.
22. McMahon, A. P., Ingham, P. W. & Tabin, C. J. 1 Developmental roles and clinical significance of Hedgehog signaling. in *Current Topics in Developmental Biology* vol. 53 1–114 (Academic Press, 2003).
23. Kopinke, D., Norris, A. M. & Mukhopadhyay, S. Developmental and regenerative paradigms of cilia regulated hedgehog signaling. *Semin. Cell Dev. Biol.* **110**, 89–103 (2021).
24. Truong, M. E. *et al.* Vertebrate cells differentially interpret ciliary and extraciliary cAMP. *Cell* **184**, 2911-2926.e18 (2021).
25. Hilgendorf, K. I., Johnson, C. T. & Jackson, P. K. The primary cilium as a cellular receiver: organizing ciliary GPCR signaling. *Curr. Opin. Cell Biol.* **39**, 84–92 (2016).
26. Mukhopadhyay, S. & Rohatgi, R. G-protein-coupled receptors, Hedgehog signaling and primary cilia. *Semin. Cell Dev. Biol.* **33**, (2014).
27. Tschaikner, P. M. *et al.* Feedback control of the Gpr161-Gαs-PKA axis contributes to basal Hedgehog repression in zebrafish. *Development* **148**, dev192443 (2021).
28. Pal, K. *et al.* Smoothened determines β-arrestin–mediated removal of the G protein–coupled receptor Gpr161 from the primary cilium. *J. Cell Biol.* **212**, 861–875 (2016).
29. Kroeze, W. K. *et al.* PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nat. Struct. Mol. Biol.* **22**, 362–369 (2015).
30. Pusapati, G. V. *et al.* G protein–coupled receptors control the sensitivity of cells to the morphogen Sonic Hedgehog. *Sci. Signal.* **11**, eaao5749 (2018).
31. Foster, S. R. *et al.* Discovery of Human Signaling Systems: Pairing Peptides to G Protein-Coupled Receptors. *Cell* **179**, (2019).
32. Bachmann, V. A. *et al.* Gpr161 anchoring of PKA consolidates GPCR and cAMP signaling.

Proc. Natl. Acad. Sci. U. S. A. **113**, 7786–7791 (2016).

33. Nehmé, R. *et al.* Mini-G proteins: Novel tools for studying GPCRs in their active conformation. *PLoS One* **12**, e0175642 (2017).
34. 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. Sel.* **29**, 583–594 (2016).
35. Rasmussen, S. G. F. *et al.* Crystal structure of the β2 adrenergic receptor–Gs protein complex. *Nature* **477**, 549–555 (2011).
36. Zhou, Q. *et al.* Common activation mechanism of class A GPCRs. *Elife* **8**, (2019).
37. Lin, X. *et al.* Structural basis of ligand recognition and self-activation of orphan GPR52. *Nature* **579**, 152–157 (2020).
38. Ye, F. *et al.* Cryo-EM structure of G-protein-coupled receptor GPR17 in complex with inhibitory G protein. *MedComm (2020)* **3**, e159 (2022).
39. Wong, T.-S. *et al.* Cryo-EM structure of orphan G protein-coupled receptor GPR21. *MedComm (2020)* **4**, e205 (2023).
40. 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).
41. Dixon, A. S. *et al.* NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. *ACS Chem. Biol.* **11**, 400–408 (2016).
42. Copp, A. J. *et al.* Spina bifida. *Nat Rev Dis Primers* **1**, 1–18 (2015).
43. Eaton, S. Multiple roles for lipids in the Hedgehog signalling pathway. *Nat. Rev. Mol. Cell Biol.* **9**, 437–445 (2008).
44. Luchetti, G. *et al.* Cholesterol activates the G-protein coupled receptor Smoothened to

promote Hedgehog signaling. *Elife* **5**, (2016).

45. Cooper, M. K. *et al.* A defective response to Hedgehog signaling in disorders of cholesterol biosynthesis. *Nat. Genet.* **33**, 508–513 (2003).
46. Budelier, M. M. *et al.* Photoaffinity labeling with cholesterol analogues precisely maps a cholesterol-binding site in voltage-dependent anion channel-1. *J. Biol. Chem.* **292**, 9294–9304 (2017).
47. Krishnan, K. *et al.* Validation of Trifluoromethylphenyl Diazirine Cholesterol Analogues As Cholesterol Mimetics and Photolabeling Reagents. *ACS Chem. Biol.* **16**, 1493–1507 (2021).
48. Castellano, B. M. *et al.* Lysosomal cholesterol activates mTORC1 via an SLC38A9-Niemann-Pick C1 signaling complex. *Science* **355**, 1306–1311 (2017).
49. Shin, H. R. *et al.* Lysosomal GPCR-like protein LYCHOS signals cholesterol sufficiency to mTORC1. *Science* **377**, 1290–1298 (2022).
50. Chen, M.-H. *et al.* Cilium-independent regulation of Gli protein function by Sufu in Hedgehog signaling is evolutionarily conserved. *Genes Dev.* **23**, 1910–1928 (2009).
51. Haycraft, C. J. *et al.* Gli2 and Gli3 localize to cilia and require the intraflagellar transport protein polaris for processing and function. *PLoS Genet.* **1**, e53 (2005).
52. Kim, J., Kato, M. & Beachy, P. A. Gli2 trafficking links Hedgehog-dependent activation of Smoothened in the primary cilium to transcriptional activation in the nucleus. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 21666–21671 (2009).
53. Dessauer, C. W. Adenylyl cyclase--A-kinase anchoring protein complexes: the next dimension in cAMP signaling. *Mol. Pharmacol.* **76**, (2009).
54. Moore, B. S. *et al.* Cilia have high cAMP levels that are inhibited by Sonic Hedgehog-regulated calcium dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 13069–13074 (2016).

55. Somatilaka, B. N. *et al.* Ankmy2 Prevents Smoothened-Independent Hyperactivation of the Hedgehog Pathway via Cilia-Regulated Adenylyl Cyclase Signaling. *Dev. Cell* **54**, 710-726.e8 (2020).

56. Jiang, J. Y., Falcone, J. L., Curci, S. & Hofer, A. M. Direct visualization of cAMP signaling in primary cilia reveals up-regulation of ciliary GPCR activity following Hedgehog activation. *Proceedings of the National Academy of Sciences* **116**, 12066–12071 (2019).

57. Smith, F. D. *et al.* Local protein kinase A action proceeds through intact holoenzymes. *Science* **356**, 1288–1293 (2017).

58. May, E. A. *et al.* Time-resolved proteomics profiling of the ciliary Hedgehog response. *J. Cell Biol.* **220**, e202007207 (2021).

59. Calebiro, D. *et al.* Persistent cAMP-signals triggered by internalized G-protein-coupled receptors. *PLoS Biol.* **7**, (2009).

60. Crilly, S. E. & Puthenveedu, M. A. Compartmentalized GPCR Signaling from Intracellular Membranes. *J. Membr. Biol.* **254**, (2021).

61. Irannejad, R. *et al.* Conformational biosensors reveal GPCR signalling from endosomes. *Nature* **495**, (2013).

62. Vilardaga, J.-P., Jean-Alphonse, F. G. & Gardella, T. J. Endosomal generation of cAMP in GPCR signaling. *Nat. Chem. Biol.* **10**, 700–706 (2014).

63. Ogden, S. K. *et al.* G protein Gαi functions immediately downstream of Smoothened in Hedgehog signalling. *Nature* **456**, 967–970 (2008).

64. Ayers, K. L. & Thérond, P. P. Evaluating Smoothened as a G-protein-coupled receptor for Hedgehog signalling. *Trends Cell Biol.* **20**, 287–298 (2010).

65. Happ, J. T. *et al.* A PKA inhibitor motif within SMOOTHENED controls Hedgehog signal

transduction. *Nat. Struct. Mol. Biol.* **29**, 990–999 (2022).

66. Stubbs, T., Bingman, J. I., Besse, J. & Mykytyn, K. Ciliary signaling proteins are mislocalized in the brains of Bardet-Biedl syndrome 1-null mice. *Frontiers in cell and developmental biology* **10**, (2023).
67. Badgandi, H. B., Hwang, S. H., Shimada, I. S., Loriot, E. & Mukhopadhyay, S. Tubby family proteins are adapters for ciliary trafficking of integral membrane proteins. *J. Cell Biol.* **216**, (2017).
68. Sheu, S. H. *et al.* A serotonergic axon-cilium synapse drives nuclear signaling to alter chromatin accessibility. *Cell* **185**, (2022).
69. Chou, C.-H. *et al.* Bisdemethoxycurcumin Promotes Apoptosis and Inhibits the Epithelial-Mesenchymal Transition through the Inhibition of the G-Protein-Coupled Receptor 161/Mammalian Target of Rapamycin Signaling Pathway in Triple Negative Breast Cancer Cells. *J. Agric. Food Chem.* **69**, 14557–14567 (2021).
70. Bock, A. *et al.* Optical Mapping of cAMP Signaling at the Nanometer Scale. *Cell* **182**, (2020).
71. Zhang, J. Z. *et al.* Phase Separation of a PKA Regulatory Subunit Controls cAMP Compartmentation and Oncogenic Signaling. *Cell* **182**, (2020).
72. Ring, A. M. *et al.* Adrenaline-activated structure of β2-adrenoceptor stabilized by an engineered nanobody. *Nature* **502**, 575–579 (2013).
73. Staus, D. P. *et al.* Sortase ligation enables homogeneous GPCR phosphorylation to reveal diversity in β-arrestin coupling. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3834–3839 (2018).
74. Faust, B. *et al.* Autoantibody mimicry of hormone action at the thyrotropin receptor. *Nature* **609**, 846–853 (2022).

75. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

76. 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).

77. Asarnow, D., Palovcak, E. & Cheng, Y. *asarnow/pyem: UCSF pyem v0.5*. (2019). doi:10.5281/zenodo.3576630.

78. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

79. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, (2018).

80. Grant, T., Rohou, A. & Grigorieff, N. cisTEM, user-friendly software for single-particle image processing. *Elife* **7**, (2018).

81. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

82. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

83. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

84. Schüttelkopf, A. W. & van Aalten, D. M. F. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 1355–1363 (2004).

85. Darbandi-Tonkabon, R. *et al.* Photoaffinity labeling with a neuroactive steroid analogue. 6-azi-pregnanolone labels voltage-dependent anion channel-1 in rat brain. *J. Biol. Chem.* **278**,

13196–13206 (2003).

86. Zhang, L. & Hermans, J. Hydrophilicity of cavities in proteins. *Proteins* **24**, 433–438 (1996).
87. Qu, C. *et al.* Ligand recognition, unconventional activation, and G protein coupling of the prostaglandin E_2 receptor EP2 subtype. *Sci Adv* **7**, (2021).
88. Lomize, M. A., Lomize, A. L., Pogozheva, I. D. & Mosberg, H. I. OPM: orientations of proteins in membranes database. *Bioinformatics* **22**, 623–625 (2006).
89. Betz, R. *Dabble*. (2017). doi:10.5281/zenodo.836914.
90. Huang, J. *et al.* CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).
91. Klauda, J. B. *et al.* Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).
92. D.A. Case, H.M. Aktulga, K. Belfon, I.Y. Ben-Shalom, J.T. Berryman, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, G.A. Cisneros, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, K. Kasavajhala, M.C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, A. Shajan, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, J. Wang, H. Wei, R.M. Wolf, X. Wu, Y. Xiong, Y. Xue, D.M. York, S. Zhao, and P.A. Kollman. Amber21. *University of California, San Francisco* (2022).
93. Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S. & Walker, R. C. Routine

Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **9**, 3878–3888 (2013).

94. Hopkins, C. W., Le Grand, S., Walker, R. C. & Roitberg, A. E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **11**, 1864–1874 (2015).
95. Roe, D. R. & Cheatham, T. E., 3rd. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).
96. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–8, 27–8 (1996).
97. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
98. Norman, R. X. *et al.* Tubby-like protein 3 (TULP3) regulates patterning in the mouse embryo through inhibition of Hedgehog signaling. *Hum. Mol. Genet.* **18**, 1740–1754 (2009).
99. Reeh, P. W. & Steen, K. H. Chapter 8. Tissue acidosis in nociception and pain. in *Progress in Brain Research* (eds. Kumazawa, T., Kruger, L. & Mizumura, K.) vol. 113 143–151 (Elsevier, 1996).
100. Jordt, S. E., Tominaga, M. & Julius, D. Acid potentiation of the capsaicin receptor determined by a key extracellular site. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 8134–8139 (2000).
101. Jasti, J., Furukawa, H., Gonzales, E. B. & Gouaux, E. Structure of acid-sensing ion channel 1 at 1.9 A resolution and low pH. *Nature* **449**, 316–323 (2007).
102. Yang, J. *et al.* PAC, an evolutionarily conserved membrane protein, is a proton-activated chloride channel. *Science* **364**, 395–399 (2019).
103. Rahman, M. F., Askwith, C. & Govindarajan, R. Molecular determinants of acidic pH-

dependent transport of human equilibrative nucleoside transporter 3. *J. Biol. Chem.* **292**, 14775–14785 (2017).

104. Li, B., Rietmeijer, R. A. & Brohawn, S. G. Structural basis for pH gating of the two-pore domain K+ channel TASK2. *Nature* **586**, 457–462 (2020).

105. Wang, L., Hall, C., Li, J., Choi, E. & Bai, X.-C. Structural basis of the alkaline pH-dependent activation of insulin receptor-related receptor. *Nat. Struct. Mol. Biol.* (2023) doi:10.1038/s41594-023-00974-0.

106. Choi, J. W., Lee, S. Y. & Choi, Y. Identification of a putative G protein-coupled receptor induced during activation-induced apoptosis of T cells. *Cell. Immunol.* **168**, 78–84 (1996).

107. Ludwig, M.-G. *et al.* Proton-sensing G-protein-coupled receptors. *Nature* **425**, 93–98 (2003).

108. Kumar, N. N. *et al.* PHYSIOLOGY. Regulation of breathing by CO_2 requires the proton-activated receptor GPR4 in retrotrapezoid nucleus neurons. *Science* **348**, 1255–1260 (2015).

109. Lassen, K. G. *et al.* Genetic Coding Variant in GPR65 Alters Lysosomal pH and Links Lysosomal Dysfunction with Colitis Risk. *Immunity* **44**, 1392–1405 (2016).

110. Vullo, S. *et al.* Conformational dynamics and role of the acidic pocket in ASIC pH-dependent gating. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 3768–3773 (2017).

111. Ruan, Z., Osei-Owusu, J., Du, J., Qiu, Z. & Lü, W. Structures and pH-sensing mechanism of the proton-activated chloride channel. *Nature* **588**, 350–354 (2020).

112. Zhang, K., Julius, D. & Cheng, Y. Structural snapshots of TRPV1 reveal mechanism of polymodal functionality. *Cell* **184**, 5138-5150.e12 (2021).

113. Wang, S., Yan, R., Zhang, X., Chu, Q. & Shi, Y. Molecular mechanism of pH-dependent substrate transport by an arginine-agmatine antiporter. *Proc. Natl. Acad. Sci. U. S. A.* **111**,

12734–12739 (2014).

114. Thomaston, J. L. *et al.* High-resolution structures of the M2 channel from influenza A virus reveal dynamic pathways for proton stabilization and transduction. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 14260–14265 (2015).

115. Liang, R. *et al.* Acid activation mechanism of the influenza A M2 proton channel. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E6955–E6964 (2016).

116. Robert, R. & Mackay, C. R. Gαs-coupled GPCRs GPR65 and GPR174. Downers for immune responses. *Immunology and cell biology* vol. 96 341–343 (2018).

117. Wirasinha, R. C. *et al.* GPR65 inhibits experimental autoimmune encephalomyelitis through CD4+ T cell independent mechanisms that include effects on iNKT cells. *Immunol. Cell Biol.* **96**, 128–136 (2018).

118. Xu, J. *et al.* GPR68 Senses Flow and Is Essential for Vascular Physiology. *Cell* **173**, 762-775.e16 (2018).

119. Wang, J.-Q. *et al.* TDAG8 Is a Proton-sensing and Psychosine-sensitive G-protein-coupled Receptor*. *J. Biol. Chem.* **279**, 45626–45633 (2004).

120. Huang, X.-P., Kenakin, T. P., Gu, S., Shoichet, B. K. & Roth, B. L. Differential Roles of Extracellular Histidine Residues of GPR68 for Proton-Sensing and Allosteric Modulation by Divalent Metal Ions. *Biochemistry* **59**, 3594–3614 (2020).

121. Rowe, J. B., Kapolka, N. J., Taghon, G. J., Morgan, W. M. & Isom, D. G. The evolution and mechanism of GPCR proton sensing. *J. Biol. Chem.* **296**, 100167 (2020).

122. Fowler, D. M. & Fields, S. Deep mutational scanning: a new style of protein science. *Nat. Methods* **11**, 801–807 (2014).

123. Jones, E. M. *et al.* Structural and functional characterization of G protein-coupled receptors

with deep mutational scanning. *Elife* **9**, (2020).

124. Macdonald, C. B. *et al.* DIMPLE: deep insertion, deletion, and missense mutation libraries for exploring protein variation in evolution, disease, and biology. *Genome Biol.* **24**, 36 (2023).

125. Coyote-Maestas, W. *et al.* Probing ion channel functional architecture and domain recombination compatibility by massively parallel domain insertion profiling. *Nat. Commun.* **12**, 7114 (2021).

126. Coyote-Maestas, W., Nedrud, D., He, Y. & Schmidt, D. Determinants of trafficking, conduction, and disease within a K+ channel revealed through multiparametric deep mutational scanning. *Elife* **11**, (2022).

127. Huang, X.-P. *et al.* Allosteric ligands for the pharmacologically dark receptors GPR68 and GPR65. *Nature* **527**, 477–483 (2015).

128. Weis, W. I. & Kobilka, B. K. The Molecular Basis of G Protein-Coupled Receptor Activation. *Annu. Rev. Biochem.* **87**, 897–919 (2018).

129. Yu, X. *et al.* Design, Synthesis, and Characterization of Ogerin-Based Positive Allosteric Modulators for G Protein-Coupled Receptor 68 (GPR68). *J. Med. Chem.* **62**, 7557–7574 (2019).

130. Harris, J. A. *et al.* Selective G protein signaling driven by substance P-neurokinin receptor dynamics. *Nat. Chem. Biol.* **18**, 109–115 (2022).

131. Koehl, A. *et al.* Structure of the μ-opioid receptor-Gi protein complex. *Nature* **558**, 547–552 (2018).

132. Hong, C. *et al.* Structures of active-state orexin receptor 2 rationalize peptide and small-molecule agonist recognition and receptor activation. *Nat. Commun.* **12**, 815 (2021).

133. Hoppe, N. *et al.* GPR161 structure uncovers the redundant role of sterol-regulated ciliary cAMP signaling in the Hedgehog pathway. *bioRxiv* 2023.05.23.540554 (2023) doi:10.1101/2023.05.23.540554.

134. Thal, D. M., Glukhova, A., Sexton, P. M. & Christopoulos, A. Structural insights into G-protein-coupled receptor allostery. *Nature* **559**, 45–53 (2018).

135. Honig, B. H. & Hubbell, W. L. Stability of "salt bridges" in membrane proteins. *Proc. Natl. Acad. Sci. U. S. A.* **81**, 5412–5416 (1984).

136. Gao, J., Bosco, D. A., Powers, E. T. & Kelly, J. W. Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins. *Nat. Struct. Mol. Biol.* **16**, 684–690 (2009).

137. Zhuang, Y. *et al.* Molecular recognition of morphine and fentanyl by the human μ-opioid receptor. *Cell* **185**, 4361-4375.e19 (2022).

138. Robertson, M. J., Meyerowitz, J. G., Panova, O., Borrelli, K. & Skiniotis, G. Plasticity in ligand recognition at somatostatin receptors. *Nat. Struct. Mol. Biol.* **29**, 210–217 (2022).

139. Mastronarde, D. N. SerialEM: A Program for Automated Tilt Series Acquisition on Tecnai Microscopes Using Prediction of Specimen Position. *Microsc. Microanal.* **9**, 1182–1183 (2003).

140. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* **74**, 519–530 (2018).

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
Nicholas Hoppe
6C5FABBCB0624E6...
Author Signature

8/10/2023
Date