

---

**GPR68**  
*A Proton Sensing  
G Protein-Coupled Receptor*

---

**Dissertation**

In fulfilment of the requirements for the doctoral degree at the  
Faculty of Mathematics and Natural Sciences at  
Kiel University

Submitted by

**Bhav Kapur**

Biberach an der Riß, 2024

First examiner: Prof. Dr. Eric Beitz

Second examiner: Dr. Herbert Nar

Third examiner: Prof. Dr. Axel Scheidig

Date of oral examination: 9<sup>th</sup> July, 2024

*“I put my heart and my soul into my work and have lost my mind in the process.”*

– Vincent Van Gogh

## Abstract

G protein-coupled receptor 68 (GPR68) is an orphan receptor that falls under the Rhodopsin-like family of G protein-coupled receptors (GPCRs). It is a proton sensing GPCR which is activated by extracellular acidity and plays a crucial role in various physiological processes. It is highly expressed in various types of cancers and has implications in chronic inflammatory diseases, fibrosis, and bone homeostasis. Pharmacological explorations of GPR68 have discovered allosteric modulators in the form of non-specific small molecules, peptides, or divalent metal ions. The precise role of the proton sensing receptor in a variety of disorders remains undetermined. It is an emerging and challenging therapeutic target due to its association with a wide array of pathophysiological conditions. The current understanding of how GPR68 senses extracellular protons is mostly hindered by the absence of a publicly available 3D structure. Existing theories propose that extracellular histidines or a buried carboxylic triad play a part in proton sensing function of the receptor.

In my doctoral thesis, I utilized protein engineering to express and purify adequate amounts of the human GPR68 in presence of a proprietary antagonist, with detergent. However, the protein quality was unsuitable for determining the structure of the receptor in its inactive conformation via X-ray crystallography and single particle cryogenic electron microscopy (cryo-EM). Following a mitigation strategy, I started to purify the homologous zebrafish GPR68 in presence of the proprietary antagonist. Subsequently, I pursued structure determination via cryo-EM with a specific protein construct which had a fusion protein (BRIL) in the intracellular loop 3, that was exploited to form a complex with Anti-BRIL FAB and Anti-BRIL-FAB nanobody. In view of substantial noise in the dataset and limited time for the thesis, I did not process the data further after achieving a satisfactory 2D classification that showed a clearly visible fiducial marker. Future cryo-EM work with improved protein material and cryo-EM samples holds great promise to enable the structure determination of this GPR68 ortholog.

Additionally, I established a protocol for validating structural models of GPR68 based on various strategies and used the best possible model for classical molecular dynamics to

identify crucial proton sensing regions of the receptor. I performed atomistic simulations with specific protonation states for previously described functionally relevant residues in GPR68 and applied graph theory and hydrogen bond calculations to describe key proton sensing structural elements by assessing the impact of the protonation states. I found that GPR68 has a robust and highly dynamic water-mediated hydrogen-bond network that connects the extracellular histidine and carboxylate clusters to the buried carboxylates via a previously described putative allosteric pocket of GPR68. The hydrogen bond network was found to be sensitive to changes in the protonation states of specific buried carboxylates and may induce sufficient structural changes to allow further connection of the buried acidic residues to the intracellular G protein binding site.

These observations and hypotheses were experimentally validated by establishing a homogeneous time-resolved fluorescence (HTRF) based functional assay for GPR68. I used this assay to measure GPR68 signaling via  $G\alpha_{q/11}$  pathway over a range of pHs. After performing mutation studies using the HTRF assay, I have shown that specific surface carboxylates, and a buried acidic residue that had no prior indication to be functional, play a substantial role in the proton sensing ability of GPR68.

## Zusammenfassung

Der G-Protein-gekoppelte Rezeptor 68 (GPR68) ist ein Orphan-Rezeptor, der zur Rhodopsin-Familie der G-Protein-gekoppelten Rezeptoren (GPCRs) gehört. Es handelt sich bei GPR68 um einen pH-abhängigen Rezeptor, der durch extrazelluläres azides Milieu aktiviert wird und eine entscheidende Rolle bei verschiedenen physiologischen Prozessen spielt. Er ist bei verschiedenen Krebsarten überexprimiert und wird als Faktor in chronisch entzündliche Erkrankungen, Fibrose und Knochenhomöostase diskutiert. Bei pharmakologischen Untersuchungen von GPR68 wurden verschiedene organische Moleküle, Peptide oder zweiwertige Metallionen als allosterische Modulatoren entdeckt. Die genaue Rolle des pH-abhängigen Rezeptors bei einer Vielzahl von Erkrankungen bleibt ungeklärt. Aufgrund seiner vermuteten Rolle in einer Vielzahl pathophysiologischer Erkrankungen ist er ein interessantes Zielobjekt für therapeutische Ansätze. Das derzeitige Verständnis darüber, wie GPR68 extrazelluläre Protonen erkennt, wird hauptsächlich durch das Fehlen einer öffentlich zugänglichen 3D-Struktur begrenzt. Bestehende Theorien gehen davon aus, dass extrazelluläre Histidine oder eine Triade saurer Aminosäuren im Rezeptorinneren eine Rolle bei der Protonenerkennungsfunktion des Rezeptors spielen.

In meiner Doktorarbeit habe ich Protein-Engineering genutzt, um das menschliche GPR68 in Detergens in ausreichenden Mengen in Gegenwart eines proprietären Antagonisten zu exprimieren und zu reinigen. Allerdings war die Proteinqualität nicht geeignet, um die Struktur des Rezeptors in seiner inaktiven Konformation mittels Röntgenkristallographie und Einzelpartikel-Kryo-Elektronenmikroskopie (cryo-EM) zu bestimmen. Als alternativen Ansatz habe ich das homologe Zebrafisch-GPR68 in Gegenwart des proprietären Antagonisten gereinigt. Mit einem spezifischen Proteinkonstrukt, ein Fusionsprotein mit einer BRIL-domäne, die in dem intrazellulären Loop 3 eingebettet ist, das wiederum zur Bildung eines Komplexes mit Anti-BRIL-FAB und Anti-BRIL-FAB-Nanokörper geeignet ist, begann ich mit der Strukturbestimmung mittels cryo-EM. Angesichts des erheblichen Rauschens im Datensatz und der begrenzten Zeit für die Arbeit habe ich die Daten bisher nicht weitergehend bearbeiten können. Aktueller Stand ist, dass eine vielversprechende 2D-Klassifizierung der Einzelpartikel vorliegt, die einen deutlich sichtbaren Referenzmarker zeigt. Zukünftige Arbeiten mit cryo-EM mit verbessertem

Proteinmaterial und cryo-EM Proben sollten mit Wahrscheinlichkeit die Strukturbestimmung dieses GPR68 Orthologen ermöglichen.

Darüber hinaus habe ich ein Protokoll zur Validierung von unterschiedlich hergeleiteten Strukturmodellen von GPR68 erstellt und das bestmögliche Modell für klassische Molekulardynamik verwendet, um wichtige Protonen-Sensorregionen des Rezeptors zu identifizieren. Ich führte atomistische Simulationen mit spezifischen Protonierungszuständen für die zuvor beschriebenen funktionsrelevanten Reste in GPR68 durch und nutzte Graphentheorie und Wasserstoffbrückenbindungsberechnungen, um Strukturelemente der Protonenerkennung zu beschreiben. Ich fand heraus, dass GPR68 über ein robustes, hochdynamisches wasservermitteltes Wasserstoffbrückennetzwerk verfügt, das den extrazellulären Histidin- und Carboxylatcluster über eine in der Literatur beschriebene mutmaßliche allosterische Tasche von GPR68 verbindet. Das Wasserstoffbindungsnetzwerk reagiert empfindlich auf Änderungen der Protonierungszustände der Carboxylat-Triade im Proteininneren und kann ausreichende strukturelle Veränderungen induzieren, um eine weitere Verbindung der vergrabenen sauren Reste mit der intrazellulären G-Protein-Bindungsstelle zu ermöglichen.

Schließlich habe ich einen Funktionstest zur Messung der pH-abhängigen GPR68-Signalübertragung über den  $G\alpha_q/11$ -Signalweg etabliert, der auf homogener, zeitaufgelöster Fluoreszenz (HTRF) beruht. Mit den Ergebnissen dieses Assays wurden die strukturabgeleiteten Hypothesen experimentell validiert. Nach der Durchführung von Mutationsstudien mit dem HTRF-Assay habe ich gezeigt, dass spezifische Oberflächencarboxylate und ein verborgener saurer Rest, für den es zuvor keine Hinweise auf eine funktionelle Rolle gab, eine wesentliche Rolle bei der Protonenerkennungsfähigkeit von GPR68 spielen.

The translation of the abstract in German was kindly provided by Dr. Herbert Nar.

## Publications

Parts of this doctoral thesis were originally written for a publication, as stated below, as well as for the deliverable reports of the Innovative Training Network on Proton and proton coupled transport project (Grant agreement ID: 860592).

**Kapur, B.,** Baldessari, F., Lazaratos, M., Nar, H., Schnapp, G., Giorgetti, A., & Bondar, A. N. (2023). Protons taken hostage: Dynamic H-bond networks of the pH-sensing GPR68. *Computational and structural biotechnology journal*, 21, 4370–4384. <https://doi.org/10.1016/j.csbj.2023.08.034>

The copyright description from Elsevier can be found below:

“Authors can include their articles in full or in part in a thesis or dissertation for non-commercial purposes”. “As an Elsevier journal author, you have the right to Include the article in a thesis or dissertation (provided that this is not to be published commercially) whether in full or in part, subject to proper acknowledgment. No written permission from Elsevier is necessary.”

Source: <https://www.elsevier.com/about/policies-and-standards/copyright/permissions>.

## Acknowledgements

The journey towards the completion of this thesis was not a solitary one, and I am deeply indebted to all those who provided their support and inspiration along the way.

Firstly, I would like to thank my academic supervisor, Prof. Dr. Eric Beitz, at the Christian-Albrechts University of Kiel for accepting to supervise my external thesis work performed at Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach. I am extremely thankful that Prof. Beitz agreed to supervise my work, which facilitated my swift enrollment at the university, and rescued me from a significant dilemma by resolving a multitude of immigration issues. Additionally, I appreciate your highly valuable feedback during routine project updates.

This research work was funded by the European Union's Horizon 2020 Research and Innovation Program under MSCA-ITN Marie Skłodowska-Curie Innovative Training Networks (ITN) grant agreement No. 860592 for the Innovative Training Network 'Proton and proton-coupled transport'.

I would like to express my special thanks, with the utmost sincere and heartfelt gratitude, to my industry supervisor, Dr. Gisela Schnapp, for being my beacon of light and helping me traverse the unknown lands of the dark scientific world. I wholeheartedly thank her for devoting a tremendous amount of time and energy to my project and for her continuous support. Furthermore, thanks very much for giving me a wonderful opportunity to work on this project while allowing me to pursue all my "crazy" ideas under the veil of scientific curiosity with no bounds.

I also want to express my immense gratitude towards my industry co-supervisor, Dr. Herbert Nar for believing in me and providing me with the rare opportunity to work on such a great project in an industry setting. I would like to thank him for the enormous amount of time expended in guiding me, providing me with extremely valuable feedback, and for all the wonderful discussions.

Additionally, I would like to thank Dr. Rebecca Ebenhoch and Dr. Dietmar Weichert for their time, encouragement, invaluable support, and feedback throughout my thesis. Thanks very much for all the lively and engaging discussions we had regarding the hurdles in the project. I am forever indebted to them for their guidance and assistance in strengthening my comprehension of various structure determination techniques.

I would also like to thank Dr. Matthias Hoffmann for the very first discussion about my project and for providing me with proprietary ligands, which were a key part of my thesis, as well as Dr. Christofer Tautermann for discussions regarding potential thermostabilizing mutants. Also, I want to express my gratitude to Karen Köhler for supporting my work with all the mass spectrometry measurements.

I would like to express my deepest gratitude to Dr. Esther Schmidt for her time and valuable insights that enabled me to finish functional studies. Furthermore, I would also like to thank Ursula Schmid and Claudia Weber, who helped me set up the cellular assays.

I feel honored to have collaborated with Prof. Dr. Ana-Nicoleta Bondar, and to have worked under supervision to embark on my expedition in the computational realm of science. I want to express immense gratitude to her for her guidance, support, energy, and time she has devoted to my project and to solve all my doubts. I learnt great many things from her which I can apply not only in the world of science, but also to the other areas of my life. Additionally, I would like to thank Dr. Michalis Lazaratos, who patiently helped me understand concepts related to the computational studies I had pursued, and I am grateful to have met and made friends with such a kind soul.

I would also like to thank my other collaborators, Prof. Dr. Alejandro Giorgetti and Filippo Baldessari, whose crucial structural modeling work helped me finish my computational investigation of the target receptor. Additionally, I would like to thank Boehringer Ingelheim's IT support and HPC teams for providing me with all the necessary infrastructure to help me wrap up my project and for promptly troubleshooting all the issues.

Prominently, I would like to thank all my lovely colleagues in the structural research department at Boehringer Ingelheim for creating such a warm and wonderful work

environment. Without the support from the entire team, my thesis would have not been possible.

I want to thank Sarah-Ana Mitrovic for being such a terrific friend and my continual venting place for my frustrations over disastrous experiments, and for tolerating me in the shared office space and more so my random thoughts and ideas over the years. I also want to thank my three musketeers and dear friends Dr. Tobias Claff, Dr. Manuel-Antonio Liaci and Dr. Carolina Sanchez, whom I unfortunately only met in the latter half of my thesis. I only wish you three were here in Biberach earlier to make my journey even more memorable and enriching. Special gratitude to Tobias Claff, with whom I had the closest relationship and the most treasured experiences. I think I finally found my long-lost brother from another mother.

I would also like to specially thank Adelheid Löhle for showing me around the ways of the lab and for her constant support. In addition to Adelheid, I would like to thank Jessica Bretzel and Heike Rapp for always making me feel at home and essentially babysitting me during my time in Biberach. Anita Bloching has my sincere gratitude for sharing with me her dark art mastery of freezing cryo-EM grids. Waldemar Eberhard has my gratitude as well for all the entertaining and fantastic conversations about our shared interests.

Moreover, I want to thank Prof. Dr. Peter Pohl and Dr. Denis Knyazev for coordinating and managing the entire PROTON-ITN project and for organizing all the meetings at beautiful locations. I'd like to thank all my colleagues who were part of PROTON-ITN, as well as the new friends I made because of the innovative training network.

I also want to thank my previous professors and mentors, Prof. Dr. Alexander Kastaniotis, Prof. Dr. Lloyd Ruddock, and Dr. Rosario Recacha, for giving me my very first opportunity to step into the world of science, and for their invaluable time and teachings that helped me propel forward in life.

My journey so far would have not been possible without the constant support and guidance of my mother, Manisha Kapur. I also want to thank all my dear friends who, despite their distance, provided unwavering support throughout my life. Lastly, I want to thank myself for not giving up. It has been a wonderful journey and a humbling experience to pursue a doctoral degree. There was much that I learned and unlearned about science and myself.

# Table of Contents

Abstract	I
Zusammenfassung	III
Publications	V
Acknowledgements	VI
Table of Contents	IX
List of Abbreviations	XII
1. Introduction	1
1.1 <i>G PROTEIN-COUPLED RECEPTORS</i>	1
1.1.1 <i>Classification, nomenclature &amp; signaling pathways</i>	5
1.1.2 <i>Constitutive activity &amp; ligands</i>	8
1.1.3 <i>Class A GPCRs</i>	9
1.2 <i>IMPACT OF PROTONATION ON GPCRS</i>	13
1.2.1 <i>Proton sensing GPCRs</i>	13
1.2.2 <i>Influence of pH on GPCRs</i>	16
1.2.3 <i>Role of protonation in the activation of class A GPCRs</i>	17
1.3 <i>GPR68</i>	19
1.3.1 <i>An orphan proton sensing GPCR</i>	19
1.3.2 <i>Expression and pathology</i>	19
1.3.3 <i>Proton sensing</i>	20
1.3.4 <i>Mechanosensing</i>	21
1.3.5 <i>Modulation via divalent metal ions</i>	22
1.3.6 <i>Small molecule ligands</i>	22
1.4 <i>WORKING WITH G PROTEIN-COUPLED RECEPTORS</i>	23
1.4.1 <i>Expression and purification</i>	24
1.4.2 <i>Structure determination techniques</i>	28
1.4.3 <i>Molecular dynamics simulations</i>	30
2. Aims & Objective	32

3. Materials & Methods	34
3.1 MATERIALS	34
3.1.1 GPR68 gene synthesis	34
3.1.2 Cell lines	34
3.1.3 Cell culture	34
3.1.4 Equipments	35
3.1.5 Disposables	37
3.1.6 Chemicals and reagents	39
3.1.7 Commercial kits	41
3.1.8 Antibodies and enzymes	41
3.1.9 Softwares	42
3.1.10 Small molecule ligands	43
3.2 METHODS	44
3.2.1 Molecular biology	44
3.2.2 Protein expression and production	46
3.2.3 Protein purification	49
3.2.4 Protein analysis	53
3.2.5 X-ray crystallography	56
3.2.6 Single particle cryogenic electron microscopy	58
3.2.7 Molecular dynamics simulations	60
3.2.8 Functional assay	65
4. Results	70
4.1 EXPRESSION, PURIFICATION AND STRUCTURE DETERMINATION TRIALS OF GPR68	70
4.1.1 Expression of constructs	72
4.1.2 Screening of different protein fusions to GPR68	75
4.1.3 Effect of termini truncations	76
4.1.4 Combination of fusions, truncations, and variations of fusion insertion site	77
4.1.5 Effect of cobalt ions on GPR68 stability	79
4.1.6 Addition of small molecules enhanced protein stability	82
4.1.7 Effect of potential thermostabilizing mutations	83
4.1.8 Anti-BRIL FAB complexation trials	85
4.1.9 X-ray crystallography trials with GPR68	89
4.1.10 Single particle cryogenic electron microscopy trials with human GPR68	91
4.1.11 Single particle cryogenic electron microscopy trials with zebrafish GPR68	94

4.2	<i>MOLECULAR DYNAMICS SIMULATIONS</i>	96
4.2.1	<i>Sequence and structural features of GPR68</i>	96
4.2.2	<i>Structural model validation</i>	99
4.2.3	<i>Insights from standard protonation simulation</i>	100
4.2.4	<i>Insights from select residue protonation simulations</i>	103
4.2.5	<i>Insights into crucial residues orientations</i>	106
4.3	<i>FUNCTIONAL ASSAY</i>	107
4.3.1	<i>Homogenous time-resolved fluorescence assay</i>	107
4.3.2	<i>Insights into extracellular and transmembrane histidines</i>	108
4.3.3	<i>Insights into buried charged residues</i>	110
4.3.4	<i>Insights into surface charged residues</i>	111
4.3.5	<i>Insights into other polar residues</i>	112
5.	<b>Discussion</b>	114
5.1	<i>EXPRESSION, PURIFICATION AND STRUCTURE DETERMINATION TRIALS OF GPR68</i>	115
5.2	<i>MOLECULAR DYNAMICS SIMULATIONS</i>	121
5.3	<i>FUNCTIONAL ASSAY</i>	124
	<b>References</b>	XV
	<b>Appendices</b>	LI
	<i>APPENDIX 1: GPR68 CONSTRUCTS</i>	LI
	<i>APPENDIX 2: SUPPLEMENTARY FIGURES</i>	LXXIV
	<i>APPENDIX 3: SUPPLEMENTARY TABLES</i>	LXXXV
	<b>List of Figures</b>	LXXXVII
	<b>List of Supplementary Figures</b>	XCVI
	<b>List of Tables</b>	XCVIII
	<b>List of Supplementary Tables</b>	XCIX
	<b>Eidesstattliche Erklärung / Declaration</b>	C

## List of Abbreviations

2D – 2 dimensional

3D – 3 dimensional

AI – artificial intelligence

ATP – adenosine triphosphate

Bluo-Gal – 5-bromo-3-indolyl- $\beta$ -D-galactopyranoside

BRIL – thermostabilized apocytochrome b562 (with mutations M7W, H102I, R106L)  
PDB ID: 1M6T <sup>[1]</sup>

BW – Ballesteros-Weinstein numbering scheme for class A GPCRs

cAMP – cyclic adenosine monophosphate

CHO – Chinese hamster ovary

CHS – cholesteryl hemisuccinate

CV – column volume

cryo-EM – cryogenic electron microscopy

DAG – diacylglycerol

DCyFIR – dynamic cyan induction by functional integrated receptors

DDM – n-dodecyl- $\alpha$ -D-maltopyranoside

DMEM – Dulbecco's modified eagle medium

DMSO – dimethyl sulfoxide

DNA – deoxyribonucleic acid

DSF – differential scanning fluorimetry

DTT – dithiothreitol

EC<sub>50</sub> – half maximal effective concentration required for activation

ECL – extracellular loop

EDTA – ethylenediamine tetraacetic acid

ER – endoplasmic reticulum

FAB – Anti-BRIL FAB

FDA – Food and Drug Administration

FRET – Fluorescence resonance energy transfer  
GDN – glyco-diosgenin  
GDP – guanine diphosphate  
GOMoDo – GPCR Online Modeling and Docking  
GTP – guanine triphosphate  
GTPases – nucleotide guanine triphosphate hydrolases  
GPCRs – G protein-coupled receptors  
GPCRdb – G protein-coupled receptor database - <https://gpcrdb.org/>  
GRKs – G protein-coupled receptor kinases  
H-bond – hydrogen bond  
HBSS – Hank’s balanced salt solution  
HEK – human embryonic kidney  
HEPES – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid  
Hi5 – *Trichoplusia ni* High Five™ cells  
HPLC – high-performance liquid chromatography  
HTRF – homogenous time-resolved fluorescence  
IBMX – 3-isobutyl-1-methylxanthine  
ICL – intracellular loop  
IgG – immunoglobulin G  
IMAC – immobilized metal affinity chromatography  
IP<sub>1</sub> – inositol monophosphate  
IP<sub>3</sub> – inositol 1,4,5-triphosphate  
IPTG – Isopropyl-β-D-1-thiogalactopyranoside  
IUPAC – International Union of Pure and Applied Chemistry  
LB – Luria Bertani  
LCP – lipidic cubic phase  
LMNG – Lauryl maltose neopentyl glycol  
MD – molecular dynamics  
MES – (2-(N-morpholino)ethanesulfonic acid)  
MSA – multiple sequence alignment

MW – molecular weight

NB – Anti-BRIL-FAB nanobody

NBT/BCIP – nitro blue tetrazolium chloride and 5-bromo-4-chloro-3-indolyl-phosphate, toluidine-salt

NMR – nuclear magnetic resonance

PAGE – polyacrylamide gel electrophoresis

PCR – polymerase chain reaction

PEI – polyethylenimine

RMSD or rmsd – root mean square deviation

SAXS – small angle X-ray scattering

SDS – sodium dodecyl sulfate

SEC – size exclusion chromatography

Sf9 – *Spodoptera frugiperda* 9

TCEP – tris(2-carboxyethyl) phosphine

TEV – tobacco etch virus

TM – transmembrane or transmembrane helix (if followed by a number e.g., TM6)

UV – ultraviolet

WHO – World Health Organization

WT – wild-type

# 1. Introduction

## 1.1 G Protein-Coupled Receptors

G protein-coupled receptors (GPCRs) form the largest family of integral membrane proteins, consisting of about 800 receptors <sup>[2]</sup>, representing roughly 5% of the expressed genes in humans <sup>[3]</sup>. GPCRs are divided into two major classes of olfactory and non-olfactory receptors. Approximately 400 out of the total 800 genes are responsible for encoding olfactory receptors <sup>[2]</sup>. The first structure of a GPCR, specifically bovine rhodopsin bound to 11-cis retinal, was resolved in 2000 <sup>[4]</sup>. Subsequently, the first structure of a human GPCR, specifically the  $\beta_2$ -adrenergic receptor, was resolved in 2007 <sup>[5]</sup>. A total of 187 unique non-olfactory receptor structures have been determined as of August 2023 based on GPCRdb <sup>[6]</sup>.

Structurally, GPCRs have seven transmembrane (TM) alpha helices. These helices are accompanied by alternating intracellular and extracellular loops <sup>[7]</sup> (Figure 1). Certain GPCRs even have an additional eighth helix, which is amphipathic in nature and runs parallel to the membrane <sup>[8]</sup>. GPCRs are usually tilted in the lipid bilayer, mostly because of helix-helix interactions and variations in tilts, kinks, lengths and hydrophobicity of the helices <sup>[7]</sup>. GPCRs undergo multiple post-translation modifications. They have an extracellular *N*-terminus that is often *N*-glycosylated <sup>[9, 10]</sup>. Additionally, the *C*-terminus may undergo several modifications like phosphorylation, palmitoylation, and ubiquitination <sup>[10]</sup>. These receptors have a disulfide bond between TM3 and extracellular loop 2 (ECL2) which is highly conserved (92%) amongst all GPCRs <sup>[11]</sup>. Certain GPCRs may possess an additional disulfide bond that may connect different regions of the protein. For example: *N*-terminus to ECL3 <sup>[12]</sup>, *N*-terminus to the junction of ECL3/TM7 <sup>[13]</sup>, *N*-terminus to ECL2 <sup>[14]</sup>, between ECLs <sup>[15]</sup> or as an intra-loop disulfide bond <sup>[16]</sup>.

Newly synthesized GPCRs undergo exocytic trafficking from the endoplasmic reticulum (ER) to the plasma membrane via the Golgi apparatus. After the synthesis of GPCRs in the endoplasmic reticulum, these receptors undergo processing and folding by quality control proteins called chaperones such as calnexin and calreticulin, which retain misfolded

GPCRs in the ER. Properly folded receptors are packaged into coat protein complex II (COPII)-coated vesicles derived from the ER. These transport vesicles visit various intracellular compartments and carry the GPCRs through ER-Golgi intermediate complex, the Golgi apparatus, and the trans-Golgi network. Throughout this process of migration, the receptors may undergo additional post-translational modifications such as glycosylation. Subsequently, these receptors are packed in exocytic vesicles in which they are transported to the plasma membrane as mature proteins [17, 18]. Certain key motifs for efficient export of the GPCRs have been identified. The motifs are in the *N*-terminus / *C*-terminus of the receptor and typically the part that is proximal to the membrane. The following motifs can be involved at different stages from the ER export to the plasma membrane. *C*-terminal motifs are E/D(X)<sub>3</sub>LL], [F(X)<sub>3</sub>F(X)<sub>3</sub>F], [FN(X)<sub>2</sub>LL(X)<sub>3</sub>L] and [F(X)<sub>6</sub> L/I L/I]. *N*-terminal motifs are [YS] and N-linked glycosylation (NXS/T). Interestingly, specific motifs exist that prevent the transport of properly folded GPCRs to the plasma membrane, which are called ER retention motifs. Examples of such motifs are RXR or a double RXR motif, which can be found either in the proximal part of the *C*-terminus or in the intracellular loop 3 (ICL3) [19].

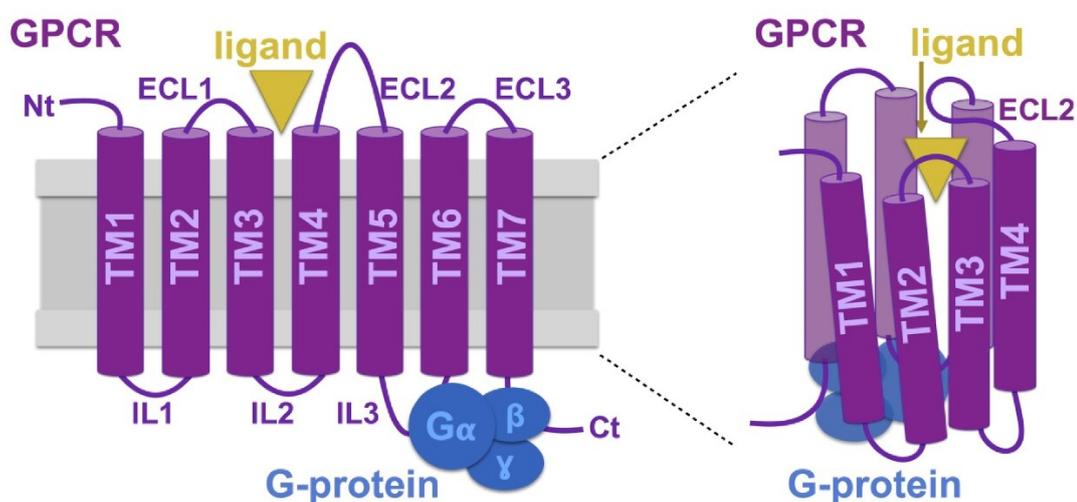


Figure 1: Overall architecture of a G protein-coupled receptor (GPCR) highlighting the seven transmembrane (TM) helices with the alternating extracellular loops (ECL), intracellular loops (IL), ligand binding site and G protein coupling site. This figure has been adapted and reused [20].

Mature GPCRs are located at the plasma membrane to facilitate their crucial function as stimuli mediators, allowing them to recognize a wide range of physical and chemical stimuli from the extracellular environment, such as protons, photons, peptides, nucleotides, lipids, mechanical forces, and more [21]. GPCRs undergo a conformational change upon

binding of these molecules, which enables binding of the G protein at the cytoplasmic side (Figure 1). This, in turn, triggers a suitable signaling pathway leading to a series of intracellular events ultimately resulting in an effect in response to the extracellular stimuli. These receptors were termed G Protein-Coupled Receptors because they can cooperate with G proteins for their function [22, 23]. The compounds that facilitate the binding of receptors to G proteins, resulting in receptor activation are referred to as agonists. Conversely, molecules that bind to same receptors but prevent G proteins from binding, resulting in receptor inactivation, are known as antagonists [24].

G proteins are heterotrimeric proteins comprising of the three subunits:  $\alpha$ ,  $\beta$ , and  $\gamma$ . When a GPCR is inactive, guanosine diphosphate (GDP) is bound to the  $\alpha$  subunit of the G protein due to the internal GTPase activity of the  $\alpha$  subunit. In the inactive state, heterotrimeric proteins exist as a complex. Typically, the  $\alpha$  and  $\gamma$  subunits are anchored to the inner leaflet of the lipid bilayer. The activation of a GPCR causes the exchange of GDP with guanosine triphosphate (GTP), which enables G protein to undergo favorable conformational changes enabling binding to the GPCR. Consequently, heterotrimeric G protein components dissociate into  $\alpha$ -GTP and a  $\beta$ - $\gamma$  complex. These components induce downstream signaling via different effectors and secondary messengers, while the receptor remains free and bound to the agonist and may activate additional G proteins initiating a process called catalytic activation of GPCRs [25–27]. The signaling of GPCRs is terminated when the GTP is hydrolyzed back to GDP in the G protein alpha sub-unit and the heterotrimer is reassembled, or dissociation of the ligand or specific modifications of the GPCRs leads to its desensitization, subsequently leading to recycling or degradation of the GPCRs [28, 29].

The internalization or endocytosis of the receptor is initiated when GPCR kinases (GRK) mediated GPCR phosphorylation, sequestration or ubiquitination occurs [28, 29]. Additionally, GPCRs can also get regulated independent of G proteins via a  $\beta$ -arrestin mediated pathway that promotes receptor internalization with additionally recruiting other adaptor or accessory proteins like clathrin [30, 31]. While ubiquitination has been associated with receptor internalization and degradation (Figure 2), S-palmitoylation participates in both GPCR export to the plasma membrane and its internalization [32].

GPCRs continuously undergo a cycle of endocytosis (recycled or degraded) and exocytosis (Figure 2). Thus, internalization of the receptor can occur independent of agonist

stimulation. Although, the rates of internalization differ. For instance, the  $\beta_2$ -adrenergic receptor is internalized from the cell surface with an estimated half-life of approximately one hour in absence of an agonist and about 10 min in presence of an agonist [18, 33]. The endocytosis of the GPCRs can occur via clathrin coated vesicles, uncoated vesicles or caveolae [34]. Caveolae can be described as special cases of lipid rafts that form small invaginations or cup-shaped pits in the plasma membrane which are approximately 50-100 nm in size [35]. Certain sequence-based motifs have also been identified for the internalization of the GPCRs. Tyrosine-based sorting motifs like YXXHy, where Hy is a hydrophobic residue and NXXY are found typically on exposed cytoplasmic domains of the receptors [19, 36]. Clusters of threonines and serines act as sites of receptor phosphorylation, and the acidic di-leucine motif [D/E XX L L/I or D/E XXX L L/I] is not only a key motif for receptor export as described above, but also play an important role in receptor internalization [32]. Additionally, the NPxxY motif (present in the TM7), which is more popularly known as one of the structural rearrangement motifs in GPCRs has been previously described to be important for receptor internalization in specific GPCRs [37, 38].

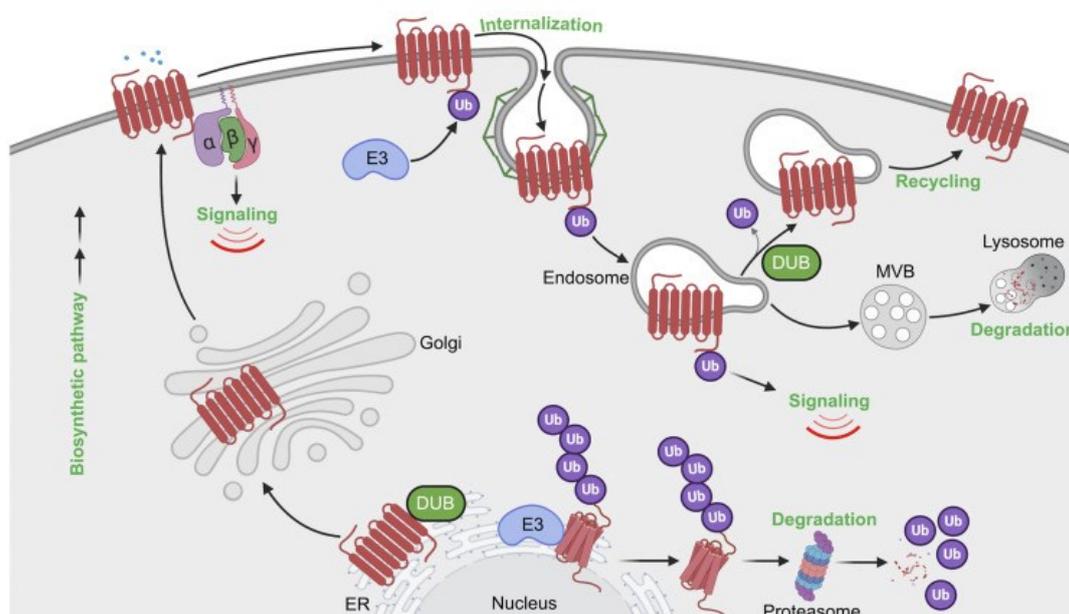


Figure 2: Outline of the life cycle of a G protein-coupled receptor, emphasizing the trafficking, recycling, and degradation via ubiquitination of GPCRs. This figure has been reused [39].

Upon entering the endosome, GPCRs have two possible fates: they can either be recycled to the plasma membrane or guided to other locations along the endocytic pathway, such as the lysosome. In most cases GPCRs are recycled to the plasma membrane, but the fate of the GPCRs is determined by specific signals within early endosomes. An example of such

a specific signal in  $\beta_2$ -adrenergic receptor would be the disruption of crucial domain interactions of an adaptor protein with the GPCR that promote degradation over recycling [40].

### 1.1.1 Classification, nomenclature & signaling pathways

Classification of GPCRs is crucial to gain a more thorough comprehension of the fascinating molecular machinery in a systematic way. The initial classification of GPCRs that was put forth is known as the GRAFS system, which stands for Glutamate (G), Rhodopsin (R), Adhesion (A), Frizzled/Taste2 (F) and Secretin (S) classes (Figure 3). Later, the Class F family of receptors was reclassified, where Taste2 (T) family of receptors were allocated a separate class [41]. The Rhodopsin class is the largest with around 700 receptors and consists of four major sub-classes  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  with 13 sub-divisions [42].

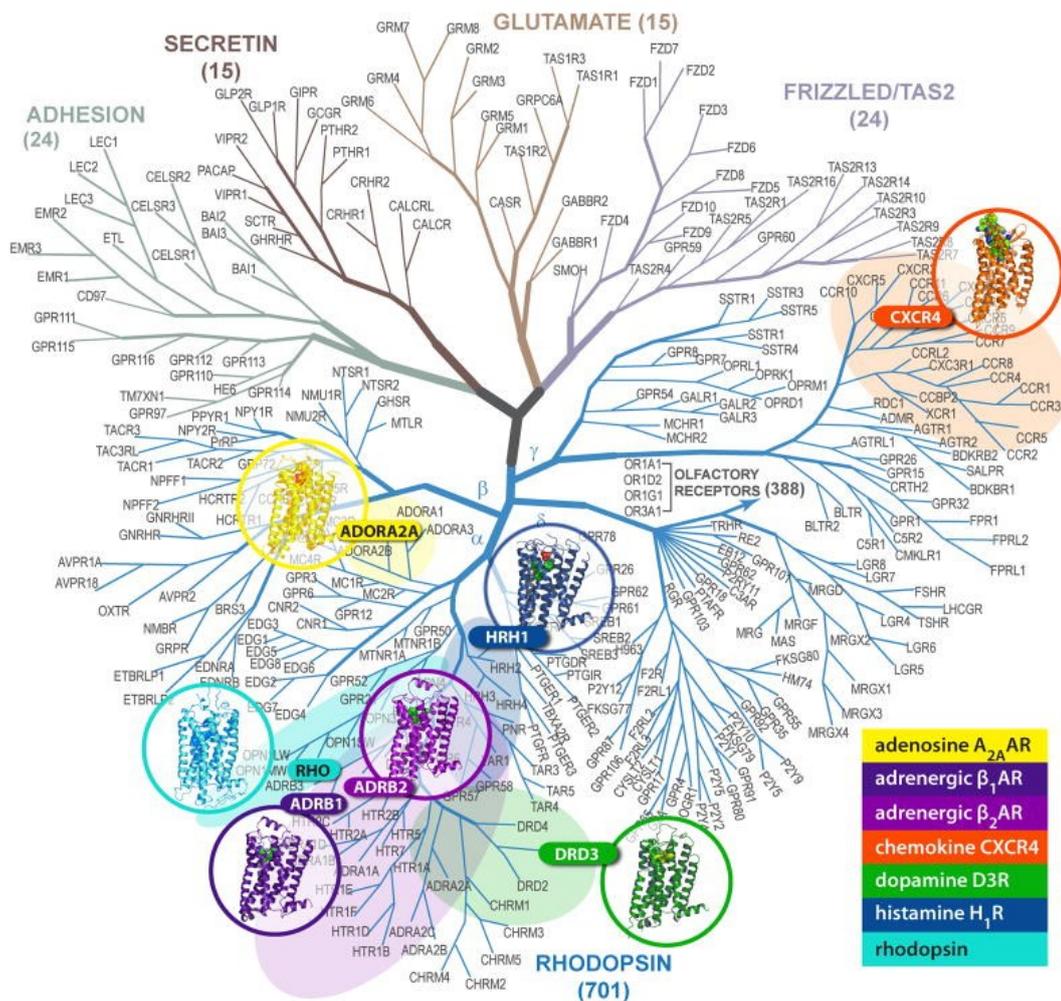


Figure 3: Classification of G protein-coupled receptor family based on the GRAFS classification system. This figure from [43] has been reused with permission.

In another classification system GPCRs were divided into seven groups A-F and O, but this was later modified to only six classes i.e., A-F [44]. The Class A is Rhodopsin-like receptors, Class B is Secretin-like, Class C is for Metabotropic glutamate-like receptors, Class D is Pheromone receptors, Class E cAMP receptors and Class F for Frizzled/Smoothed receptors. The receptors that do not have a defined function or endogenous ligand but have structural similarities to the GPCRs are termed as “orphan receptors” [45].

Comparison of a particular residue position within GPCRs has been made easier due to existing well-established sequence-based numbering methods for Class A [46], Class B [47], Class C [48] and Class F [49]. The nomenclature used in these numbering systems consists of two parts. The first number represents the helix, ranging from TM1 to TM7. The second number indicates the position of the residue relative to the most conserved position, which is designated as 50, for example – N1.50 (97%) D2.50 (94%), R3.50 (96%), W4.50 (96%), P5.50 (77%), P6.50 (100%), and P7.50 (96%) [50]. For instance, the value 3.49 represents a residue located in TM3, which is one position before the most conserved residue (3.50). The conserved positions of a helix under question vary among the different GPCR classes. These systems were further improved by GPCR database (GPCRdb) nomenclature that additionally gives insights into the helix constrictions or bulges. This system is based on the GPCRs whose structures have been determined and accounts for structure-based sequence alignment. A unique separator ‘x’ is added to the previous described systems followed by a number that is based on the structural data. For instance, 6.40x41. The conserved residues in the loops can also be written with the GPCRdb nomenclature, especially residues of the ECL2. For instance, C45.50 – here the number 45 tells us that the loop is between TM4 and TM5. Although, as loops are less structured there are limited conserved residues that can be labelled with GPCRdb nomenclature. Superscript can be used to integrate this numbering scheme with the general residue numbers and receptor-specific numbers e.g., R119<sup>3.50x50</sup> or R3.50x50<sup>119</sup> [51].

Furthermore, it is possible to further categorize the three subunits that constitute G proteins. Although, the major classification is based on the  $\alpha$  subunit. G proteins can be mainly classified into four types based on the  $\alpha$  sub-unit:  $G\alpha_q$ ,  $G\alpha_s$ ,  $G\alpha_i$  and  $G\alpha_{12/13}$ . Each of these family of the  $\alpha$  sub-unit can further be sub-divided. Likewise, the  $\beta$  and  $\gamma$  subunits can also be classified and have multiple members. Such functional and structural diversity makes it a complicated task in understanding the downstream signaling pathways of the GPCRs [52].

In  $G\alpha_s$  the letter 's' denotes stimulatory [53], in  $G\alpha_i$  the letter 'i' denotes inhibitory [53] and in  $G\alpha_q$  the letter 'q' denotation remains unclear.

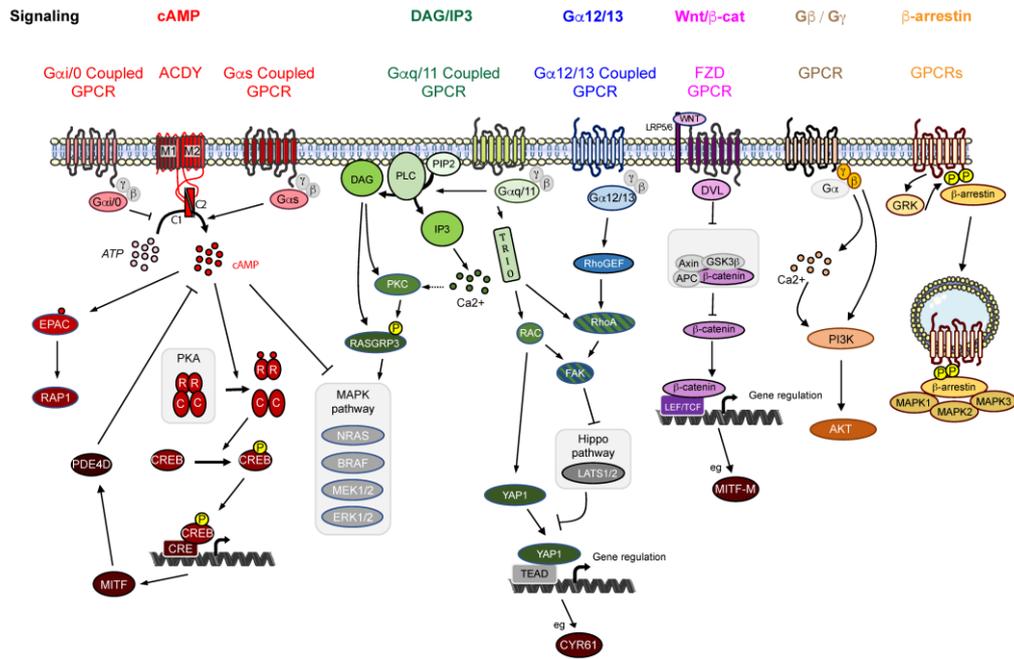


Figure 4: Signaling pathways for G protein-coupled receptors. This figure has been reused [54] under the Creative Commons Attribution license.

Signaling via both  $G\alpha_s$  and  $G\alpha_i$  modulates a common effector, adenylyl cyclase, which catalyzes the conversion of cytosolic adenosine triphosphate (ATP) to cyclic adenosine monophosphate (cAMP), a secondary messenger (Figure 4). It is directly stimulated by the  $G\alpha_s$  class of proteins. Conversely, binding of  $G\alpha_i$  inhibits conversion of ATP to cAMP by inhibiting adenylyl cyclase. The levels of cytosolic cAMP can activate members of the protein kinase A family or ion channels representing secondary effector molecules [55]. Signaling via  $G\alpha_q$  activates an effector molecule known as phospholipase C. This enzyme catalyzes the breakdown of membrane-bound phosphatidylinositol 4,5-bisphosphate (PIP<sub>2</sub>) into the secondary messenger inositol 1,4,5-triphosphate (IP<sub>3</sub>) and diacyl glycerol (DAG). The messenger IP<sub>3</sub> binds to calcium ion channel/IP<sub>3</sub> receptors located in the ER membrane, causing the release of calcium ions from the ER. Meanwhile, DAG can activate any membrane-bound variants of a secondary effector molecule known as protein kinase C. The calcium ions released from the ER act as tertiary messengers that trigger activation of other proteins like calmodulin [56]. Lastly, signaling via the  $G\alpha_{12/13}$  stimulates guanine nucleotide

exchange factors for the Rho small GTPases which in turn regulate the actin cytoskeleton [57, 58].

### 1.1.2 Constitutive activity & ligands

Constitutive or basal activity refers to GPCR activation in the absence of a ligand subsequently triggering signal transduction pathway(s). Constitutive activity in GPCRs was first discovered in the 1980s, when  $\beta_2$ -adrenergic and  $\delta$ -opioid receptor showed basal activity [59, 60].

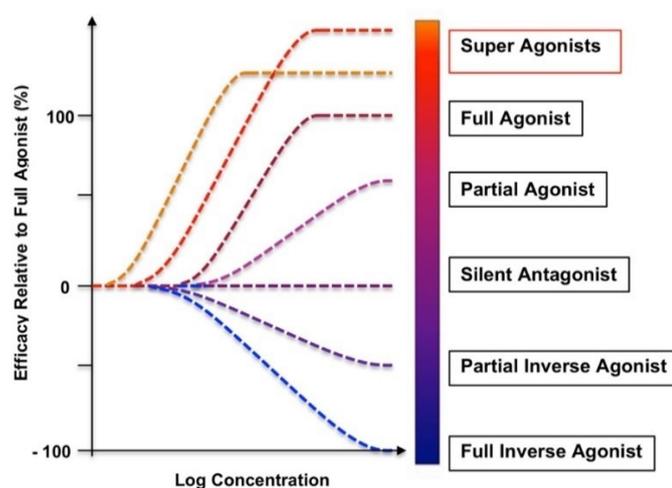


Figure 5: Various types of dose-response curves and a spectrum of relative efficacy for GPCR ligands. This figure has been reused [61] under the Creative Commons Attribution license.

A diverse array of endogenous or synthetic ligands can activate GPCRs, and these ligands can be categorized based on their capacity to facilitate the coupling of different signal transducers to the receptors. Ligands that maximally activate the GPCRs are considered as full agonists (usually endogenous ligands), and a ligand is known as a super-agonist when the maximal effect of the ligand exceeds that of the endogenous ligand [62]. Agonists which do not induce 100% activation are defined as partial agonists. Ligands that block agonist activity are called antagonists and can be of two types: inverse agonists, which reduce constitutive activity, and silent or neutral antagonists, which limit the actions of agonists without affecting constitutive activity [63] (Figure 5).

These group of ligands typically bind at the same site where the endogenous agonists bind their cognate GPCRs, and this site is known as the orthosteric binding site or pocket. Sometimes this orthosteric pocket maybe occupied by the *N*-terminus (after proteolytic cleavage) [64] or the ECL2 of the same GPCR [65] and lead to receptor activation e.g.,

calcitonin receptor-like receptor<sup>[66]</sup> or block ligand access to the receptor e.g., sphingosine-1-phosphate receptor 3<sup>[67]</sup>. In addition, GPCRs can also be modulated allosterically by compounds that bind to a site different than the orthosteric pocket. These allosteric ligand binding sites may present on the extracellular side or in the transmembrane region<sup>[68]</sup> and in some cases even at the intracellular side<sup>[69]</sup>. The allosteric ligands can modulate the activity of the receptors negatively or positively and are categorized as negative allosteric modulators (NAMs) or positive allosteric modulators (PAMs) respectively. Allosteric modulators can also occur naturally, such as sodium ions and cholesterol. Bitopic ligands interact with both orthosteric and allosteric binding sites and can either be either agonists or antagonists<sup>[63, 70]</sup>. Some ligands can selectively activate a specific signal transduction pathway, and this phenomenon is known as functional selectivity or biased signaling. The ligands that cause this effect are called as biased ligands<sup>[71–73]</sup>.

### 1.1.3 Class A GPCRs

Class A GPCRs or Rhodopsin-like family of the receptors is the most widely investigated class and it accounts for 85% of GPCRs. The Rhodopsin-like family includes 435 sensory GPCRs (olfaction, vision, taste and pheromone), 197 non-sensory receptors (excluding orphans) and 87 orphan receptors<sup>[74]</sup>. The non-sensory receptors can be categorized into eight different broad classes according to the type of ligand that stimulates them: protein, peptide, nucleotide, lipid, steroid, alicarboxylic acid, melatonin, and aminergic based on GPCRdb<sup>[6]</sup>. Such diversity of ligand recognition makes this group of receptors an interesting therapeutic target area. More than 500 drugs that are available on the market target class A GPCRs of which three-quarters are synthesized against aminergic receptors and 10% for peptide-activated receptors. These drugs cover disease areas like cardiovascular diseases, pulmonary diseases, hypertension, depression, glaucoma, Parkinson's disease, cancer associated symptoms and many more<sup>[75]</sup>.

Class A GPCRs have a standard GPCR structure as described previously in this thesis (Figure 1, section 1.1). GPCR activation requires intricate rearrangement within the core receptor which is facilitated by certain key motifs. Different classes of GPCRs possess different set of functional motifs that undergo major rearrangement to cause a global conformational change in the receptor. These functional motifs are highly conserved within every class. The unique set of motifs observed within Class A GPCRs are: CWxP motif

(C6.47, W6.48, P6.50) also known as toggle switch <sup>[76, 77]</sup>, PIF motif (P5.50, I3.40, F6.44) <sup>[78, 79]</sup>, allosteric sodium pocket <sup>[80]</sup>, NPxxY (N7.49, P7.50, Y7.53) motif <sup>[81–83]</sup>, and DRY (D3.49, R3.50, Y3.51) motif <sup>[84]</sup> (Figure 6).

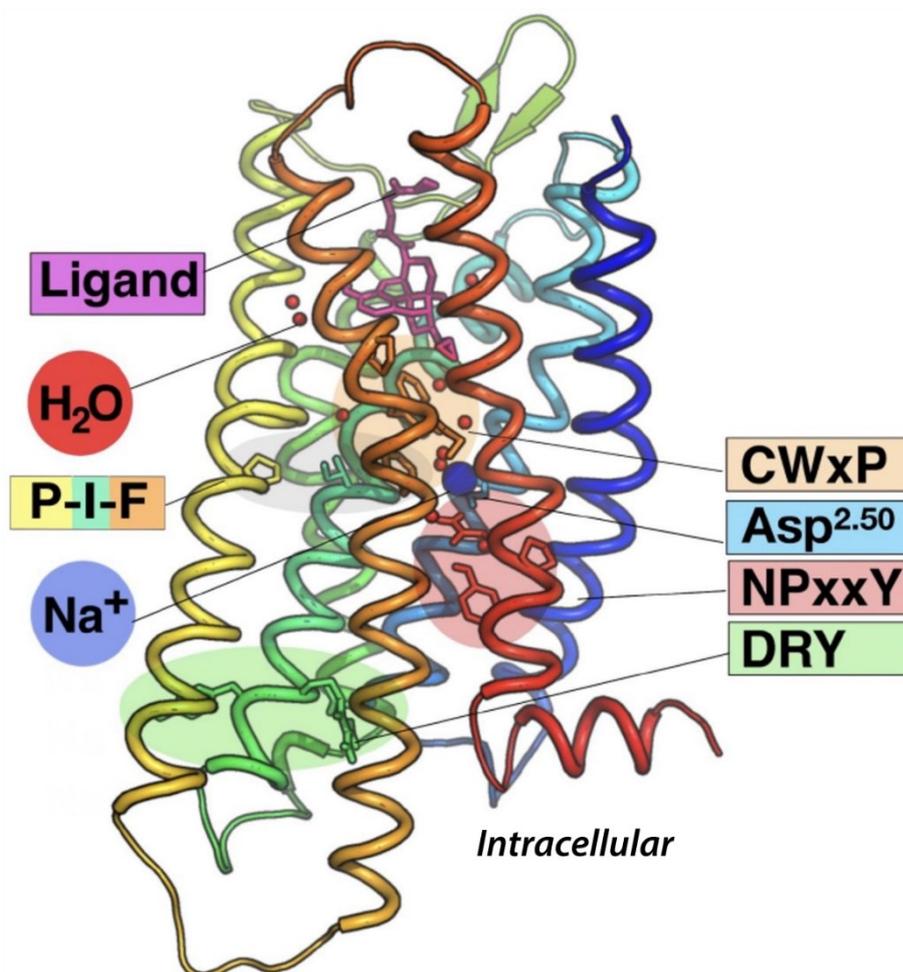


Figure 6: Conserved motifs of Class A GPCRs. The figure has been adapted and reused <sup>[85]</sup> under the Creative Commons Attribution license.

The CWxP motif residues (C6.47, W6.48, P6.50) have a sequence conservation of 71%, 78% and 98% amongst the rhodopsin family of receptor <sup>[77]</sup>. The proline causes a helix kink of about 35°, which is 15° more than usually found for prolines in TM helices <sup>[86]</sup>. C6.47 and W6.48 are the most important residues of the motif that undergo a sidechain rotamer shift depending on the receptor state <sup>[87, 88]</sup>, but this data was found to be inconsistent within the rhodopsin family <sup>[77]</sup>. Although, site directed mutagenesis of C6.47 or W6.48 is known to affect receptor signaling or ligand recognition <sup>[89, 90]</sup>. Interestingly, the CWxP motif is known to influence the residues involved in the NPxxY motif and the sodium pocket of class A GPCRs <sup>[77]</sup>.

The PIF and CWxP motifs are spatially proximal, and together form a larger motif known as the transmission switch <sup>[76,91]</sup>. This larger motif confers rearrangements in the orthosteric ligand binding pocket affecting ligand interaction and consequently conformational changes on the cytosolic side. The PIF motif (P5.50, I3.40, F6.44) position 3.40 may possess smaller hydrophobic amino acids like valine or isoleucine forming PVF or PIL motifs respectively in class A GPCRs <sup>[92]</sup>.

The activation of class A GPCRs involves the expulsion of the sodium ion from the receptor which leads to the collapse of the sodium pocket. Thus, the sodium ion acts as a negative allosteric modulator in class A GPCRs <sup>[80]</sup>. This sodium ion is attached via multiple polar sidechains and a few directly coordinated water molecules. The sodium pocket is typically comprised of the residues D2.50 (90% conserved), S3.39 (75% conserved), N7.45 (70% conserved) and N7.49 (75% conserved or an aspartate here is roughly 20% conserved), but in certain GPCRs (e.g., neurotensin receptor 1) other residues at positions N1.50 (97% conserved) and S7.46 (66% conserved) may also become a part of the allosteric sodium ion pocket <sup>[80]</sup>. Note that N1.50 is the most conserved position within the rhodopsin family <sup>[50]</sup>. Similar allosteric effects can also be observed in presence of other small organic cations like amiloride instead of sodium ion <sup>[93]</sup>.

The NPxxY motif is recognized for its important contribution to the receptor activation pathway and is located in the TM7 <sup>[94]</sup>. Inward movement of TM7 during activation allows for crucial interactions like a hydrogen bond formation between Y7.53 and Y5.58, and Y7.53 with residues in the TM3 (3.43, 3.46 and R3.50 of the DRY motif), which stabilizes the active state of the receptor <sup>[95, 96]</sup>. The position 7.49 maybe be an asparagine (75% conserved) or an aspartate (20% conserved) and 7.53 is a tyrosine (89% conserved). This N(D)7.49 has been reported to also be a part of sodium pocket <sup>[80]</sup>. The other roles of the NPxxY motif include stabilizing the receptor for receptor internalization <sup>[97]</sup>.

Lastly, the DRY motif (D3.49, R3.50, Y3.51), where the positions 3.49 is 87% conserved with a negatively charged receptor, 3.50 is 93% conserved with arginine and 3.51 is 64% conserved with tyrosine. The R3.50 forms an ionic lock with the D3.49 in TM3 and with E/D6.30 in TM6 <sup>[84, 98, 99]</sup>. In certain GPCRs like chemokine receptor CCR5, the position 6.30 may contain a lysine, where the R3.50 interacts with the K6.30 via water mediated

hydrogen bonds <sup>[100]</sup>. The breaking of these ionic interactions is essential for the activation of the receptor as the R3.50 can interact directly with the G $\alpha$  sub-unit of the G proteins <sup>[101]</sup>.

Apart from these motifs another hydrophobic pocket has been observed in Class A GPCRs. This hydrophobic lock consists of residue positions 6.40, 6.41, 3.43. The breaking of the hydrophobic lock between TM3-TM6 allows the outward movement of TM6 which is key for receptor activation. The rearrangement of these residues happens in conjunction with the other conserved motifs and crucial for receptor activation <sup>[102, 103]</sup>.

A conserved rearrangement of the motifs has been previously described after analyzing active and inactive state structures of several GPCRs. This rearrangement of the conserved motifs residues sidechains leads to larger changes in the TM helices leading to their inward or outward movement, rotation, and tilt. An inter-helical or inter-residue contacts map of over 400 class A GPCR structures revealed a common mechanistic pathway from the extracellular ligand binding site to the intracellular G protein coupling site via the conserved motifs. Sequence conservation analysis of each GPCR class shows that most residue-residue contacts can be formed in at least 30% of receptors <sup>[104]</sup>.

A study that involved the active and inactive state structures of 13 class A GPCRs revealed that the outward movement of the TM6 supported by TM5 and TM7 is a common hallmark activating feature within the rhodopsin family. The movement at the extracellular side of TM6 is by about 2 Å, whereas the movement on the intracellular side can range from 7-13 Å. Moreover, TM6 also undergoes a rotation of about 38° upon activation. Similarly, the cytosolic end movement of TM5 is about 2 Å on an average and TM7 also may experience either extracellular or cytosolic side movements ranging from 2-4 Å roughly <sup>[104]</sup>. These movements within the three helices are supported by the presence of a helix kink which exists due to highly conserved proline located at the position X.50, where X can be helices 5, 6 or 7 <sup>[104, 105]</sup>. Additionally, TM3 acts as an important helix in stabilization of a specific receptor state due to its high residue contact frequency to the other regions of the receptor <sup>[96, 106]</sup>.

Hence, the activation of class A GPCRs may be described as concerted complex events that includes the binding of a ligand, alterations in the residue sidechain rotamers of highly

conserved class A GPCR motifs, and concomitant modifications in the overall movement (inward or outward) and rotation of TM helices, ultimately followed by G protein coupling.

## 1.2 Impact of protonation on GPCRs

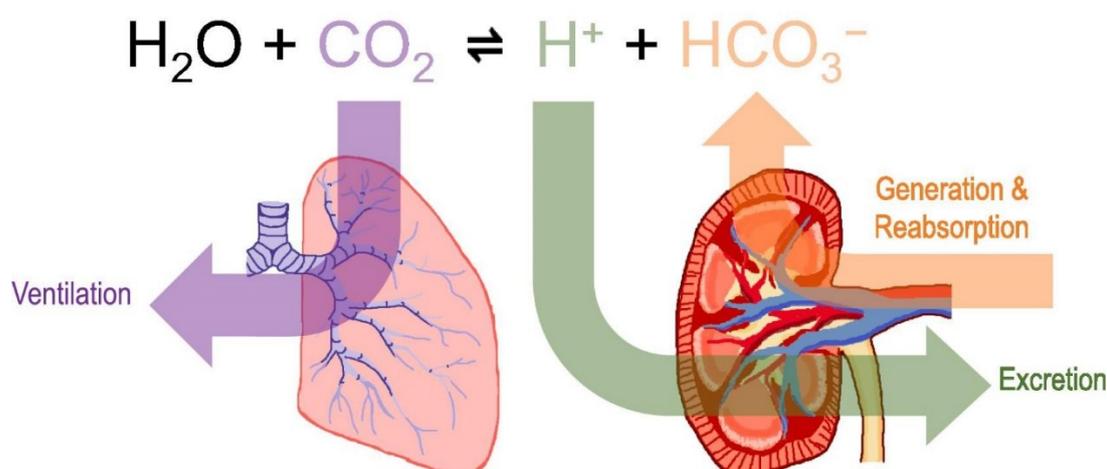


Figure 7: Mechanism of the acid base homeostasis in the human body. This figure has been adapted and reused <sup>[107]</sup>.

The blood pH in humans ranges from 7.35-7.45 in absence of pathological states. The bicarbonate buffer system (Figure 7) alone is not sufficient to maintain pH homeostasis during consistent and large acid-base imbalance. The biological membranes allow for passive diffusion of certain selected ions, although a controlled selective ion transfer is more efficient in regulating pH within different cell compartments <sup>[108]</sup>. Several membrane proteins play a role in pH regulation by transporting ions selectively along (passive transport – often regulated by ion channels) <sup>[109]</sup> or against a gradient (active transport – often regulated by ion pumps, which can be either antiporters or symporters) <sup>[110]</sup>, converting byproducts that create pH gradients <sup>[111]</sup>, and detecting protons or pH changes to trigger different signaling pathways <sup>[112]</sup>. A few examples are  $\text{Na}^+$ - $\text{H}^+$  exchangers (NHE) <sup>[113]</sup>, monocarboxylate- $\text{H}^+$  efflux cotransporters (MCT) <sup>[114]</sup>,  $\text{Na}^+$ / $\text{HCO}_3^-$  cotransporters (NBC) <sup>[115]</sup>, carbonic anhydrases <sup>[111]</sup> and pH sensing GPCRs <sup>[112]</sup>. The least understood membrane proteins that respond to the acid-base imbalance are pH sensing GPCRs.

### 1.2.1 Proton sensing GPCRs

Proton sensing GPCRs sense extracellular pH to undergo conformational change and subsequent activation of various signaling pathways <sup>[112]</sup>. Proton sensing GPCRs play an important role in different pathological states like cardiovascular diseases, tumor biology,



experimental protein structures of the proton sensing GPCRs are publicly unavailable making it difficult to establish structure function relationships.

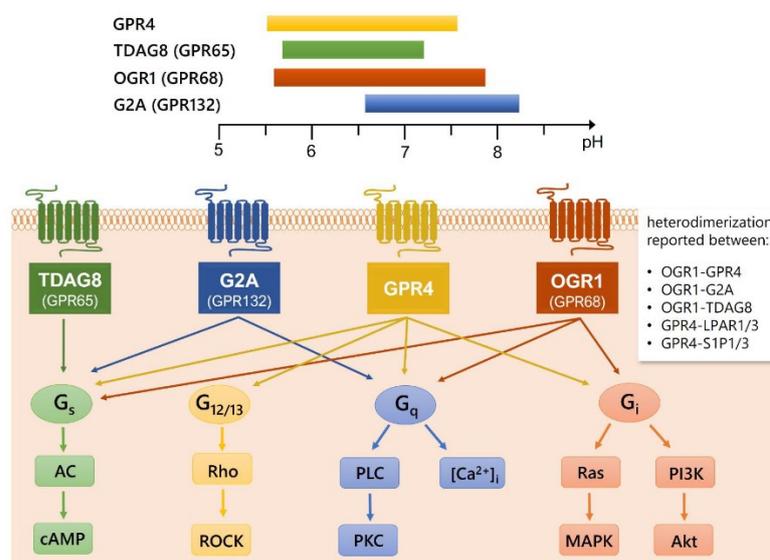


Figure 9: In vitro pH activation range, preferences for signaling pathways, and signaling of the proton sensing GPCRs. The following figure displays the pH activation range for each of the four receptors, namely TDAG8 (GPR65, green), G2A (GPR132, blue), GPR4 (yellow), and OGR1 (GPR68, red). Abbreviations: AC: adenylyl cyclase, cAMP: cyclic adenosine monophosphate, Rho: Ras homologue, ROCK: Rho-associated coiled-coil-containing protein kinase, PLC: phospholipase C, PKC: protein kinase C,  $[Ca^{2+}]_i$ : intracellular calcium concentration, Ras: rat sarcoma protein, MAPK: mitogen-activated protein kinase(s), PI3K: phosphoinositide 3-kinase, Akt: protein kinase B, LPAR: lysophosphatidic acid receptor, S1PR: sphingosine-1-phosphate receptor. This figure has been reused <sup>[116]</sup> under Creative Commons Attribution license.

Certain lipid molecules have been described to bind to proton sensing GPCRs: 1- $\beta$ -D-galactosylsphingosine (psychosine) binds to GPR65 <sup>[135]</sup>, and 12-(S)-hydroxyeicosatetraenoic acid binds to GPR31 <sup>[136]</sup>. GPR68, GPR4, and GPR132 were proposed to bind sphingosylphosphorylcholine and lysophosphatidylcholine <sup>[137–139]</sup>, but these three papers were later retracted. A biologically active neuropeptide, galanin was shown to slightly activate GPR151 <sup>[140]</sup>, but later studies revealed that galanin is unlikely to activate GPR151 <sup>[141]</sup>.

The proton sensing GPR68 and GPR4 undergo oligomerization (both homo and hetero) with other proton sensing sub-family members and other lipid sensing GPCRs like lysophosphatidic acid receptors or sphingosine-1-phosphate receptors <sup>[142]</sup> (Figure 9). Though, later studies have suggested the heterodimerization of sphingosine-1-phosphate receptors with distantly related receptors to be unlikely <sup>[143]</sup>.

## 1.2.2 Influence of pH on GPCRs

An increase in the extracellular proton concentration affects proton sensing GPCRs. Yet, it is extensively questioned if protons are the true endogenous ligands for the proton sensing GPCRs. There is a divide over the question whether the proton sensing GPCRs are directly activated by protons or rather allosterically modulated [116, 144].

There is evidence that additional GPCRs can be activated or modulated by pH but are not classified as pH sensing GPCRs. A study investigating proton modulation of 28 unique class A GPCRs using 10 independent C-terminal  $G\alpha$  chimera strains showed that the receptors were positively, negatively, or unmodulated by protons in the presence of their respective agonists, and modulation by protons may depend on the concentration of the agonist. Furthermore, the study reported that the adenosine  $A_{2A}$  receptor was activated by acidic pH ( $pH_{50}$  of  $6.25 \pm 0.75$ ) in the absence of an endogenous or synthetic ligand. This study was performed in a yeast-based platform using Dynamic Cyan Induction by Functional Integrated Receptors (DCyFIR), which exploits the single pheromone GPCR activation pathway in the yeast and the fact that the intracellular pH of the system is insensitive to the changes in the extracellular pH [145, 146]. Another example of protonation dependent activation has previously been reported for the  $\beta_2$ -adrenergic receptor, where the protonation of the D2.50 (sodium pocket) destabilizes the inactive state of the receptor [147]. The observations described above raises the question – whether this protonation-dependent activation is a coincidental characteristic shared by all class A GPCRs? Modulation by pH has also been witnessed in the calcium sensing receptor, a class C GPCR, where the binding of the endogenous agonist, i.e., calcium ions or other polycationic agonists, is modulated by pH [148].

G proteins interact with GPCRs at the intracellular side of the receptor after agonist binding. In various pathological states like ischemia and neuroinflammation [149–151] or other metabolic stress [152], the intracellular pH can drop to about 6.6–6.8. The pH sensing structural motifs in the Ras-like domain and helical domains have been described for the mammalian  $G\alpha$  subunit, which govern its stability and conformation. For instance, one study indicated that the  $G\alpha_i$  subunit may possess ionizable residues in the buried regions of the protein whose  $pK_a$  is upshifted. Typically, the residues that act as proton sensors have

pK<sub>a</sub> values close to the physiological range <sup>[152]</sup>. Thus, it is crucial to not only focus on extracellular pH but also on the intracellular pH in understanding the behavior of a GPCR.

Sub-cellular compartments like the endosomes which play an important role in GPCR transport to the plasma membrane and recycling <sup>[18]</sup> expose GPCRs to an intraluminal acidic pH (pH 4.5-6.5) <sup>[153, 154]</sup>. GPCR signaling from the endosomes is possible via G proteins or  $\beta$ -arrestin. Most of the endosome based GPCR signaling has been studied mainly with class B GPCRs like vasopressin receptor type 2 <sup>[155]</sup> and parathyroid hormone receptor type 1 <sup>[156]</sup> but has been also witnessed in class A GPCRs like  $\beta$ 2-adnergic receptor <sup>[157]</sup>. Larger complexes termed as ‘megaplexes’, which consist of the GPCR,  $\beta$ -arrestin and the G protein has also been observed in certain cases where  $\beta$ -arrestin binds to the C-terminus of the receptor <sup>[155, 158]</sup>. Such data has opened new avenues for novel drug discovery opportunities where ligands could sustain or inhibit endosomal GPCR signaling <sup>[159, 160]</sup>.

In special cases, pH can influence ligand behavior and GPCR activation, where the ligand binds to an ionizable residue in the binding pocket. Such an example is the  $\mu$ -opioid receptor that has a histidine residue in the ligand pocket, and under acidic conditions, H6.52 gets protonated and prevents ligand binding <sup>[161]</sup>.

### 1.2.3 Role of protonation in the activation of class A GPCRs

Hydrogen bonds play a role in proton-transfer processes and can be potentially seen as intermediates to these processes <sup>[162]</sup>. Some studies highlight the role of complex water mediated hydrogen bond networks in the activation of GPCRs like the  $\mu$ -opioid receptor and  $\delta$ -opioid receptor <sup>[163]</sup>, and the importance of conserved water networks in class A GPCRs <sup>[164]</sup>.

The impact of protonation on GPCRs can be better understood by investigating the structural details at an atomistic level. The energy required for the conformational change within class A GPCRs is governed by the proton movement from the extracellular side to the intracellular side of the membrane driven mainly by the differences in the membrane potential <sup>[165]</sup>.

Proton transfer would require ionizable residues, polar residues, or water molecules <sup>[166]</sup>. Highly conserved class A GPCR motifs seem to contain highly conserved residue positions

like D2.50 (sodium pocket) in the core of the receptor, Y7.53 (NPxxY motif) and D/E 3.49 (DRY motif) at the cytosolic end <sup>[77]</sup>. The D2.50 and D3.49 are titratable and can activate the receptor when protonated. These motifs and conserved titratable residues are supported by other polar residues and ordered water molecules within GPCRs to form a proton transfer network allowing proton transfer from the extracellular side to the intracellular side, where the source of protons is likely to be the extracellular side. Movement of protons would alter the dynamic properties within different regions of protein causing helix rotations or tilts and formation of new interactions favoring a particular conformation (active or inactive) <sup>[167]</sup>.

Neutralization of D2.50, which is situated in the buried core of the class A GPCRs, abolishes receptor activation as studied by mutating to an asparagine <sup>[168]</sup>. The sodium ion that acts as a negative allosteric modulator of class A GPCRs binds mainly to D2.50, only to show the importance of negatively charged D2.50 <sup>[80]</sup>. The energy barrier for sodium ion transfer is higher than that for proton transfer. Although, protonation of D2.50 leads to the expulsion of the sodium ion to the cytoplasmic side by reducing the energy barrier allowing movement of the sodium ion <sup>[169]</sup>. This would cause the collapse of the sodium pocket causing proximity of S3.39 to D2.50 which leads to deprotonation of D2.50 and by involvement of ordered water molecules and other polar residues mainly from the conserved motifs e.g., Y7.53 would allow eventual protonation of D3.49. Upon its protonation, the salt bridge between R3.50 and D3.49 would break leading to a series of altered interactions and TM helix rearrangement leading to G protein recruitment and subsequent activation <sup>[167]</sup>. Apart from the above-mentioned motifs, C6.47 (CWxP motif) is regarded to play a crucial role in the supporting the hydrogen bond network formed by D2.50 and N7.45 of the sodium pocket, and N.7.49 of the NPxxY motif <sup>[77]</sup>. Additionally, computational studies on  $\delta$ -opioid receptor and  $\kappa$ -opioid receptor have shown a comparable activation mechanism to the one described before for rhodopsin. In the inactive state, the DRY motif is not involved in the larger hydrogen bond network, but it becomes part of it upon receptor activation. These studies focused on protein-water H-bond network and conformational dynamics using graph theory <sup>[163]</sup>.

## 1.3 GPR68

### 1.3.1 An orphan proton sensing GPCR

GPR68 is an orphan class A proton sensing GPCR [112]. GPR68 is also known as ovarian cancer GPCR 1 (OGR1), as it was first isolated from ovarian cancer cells. GPR68 has been mapped to chromosome 14 at band 14q31. GPR68 has 365 amino acids and has a molecular weight of about 41 kDa [118]. Structurally, GPR68 is predicted to have the typical class A GPCR structure (Figure 1A). It has two N-glycosylation sites in the *N*-terminus along with a putative site proposed in the extracellular loop 1 [118]. Overexpressed GPR68 in the human embryonic kidney (HEK293) cells has reported to be glycosylated at two sites with a molecular weight of 58 kDa [170]. Apart from the highly conserved disulfide bond within class A GPCRs i.e., C3.25-C45.50, GPR68 may have an additional disulfide between C13<sup>1.24</sup>-C258<sup>ECL3</sup> [171]. GPR68 can be activated by protons alone and is fully active at about pH 6.6-6.8 and fully inactive at about pH 7.6-7.8 [112, 172]. GPR68 can signal via  $G\alpha_q$  [112, 173, 174],  $G\alpha_s$  [175, 176],  $G\alpha_i$  [177] or  $G\alpha_{12/13}$  pathways [178], but there is no evidence for  $\beta$ -arrestin coupling [170]. Weak homodimerization and heteromerization of GPR68 with other proton sensing GPCRs like GPR132 [179] and GPR4 or other lipid sensing GPCRs like lysophosphatidic acid or sphingosine-1-receptor has been previously described [142] (Figure 9).

### 1.3.2 Expression and pathology

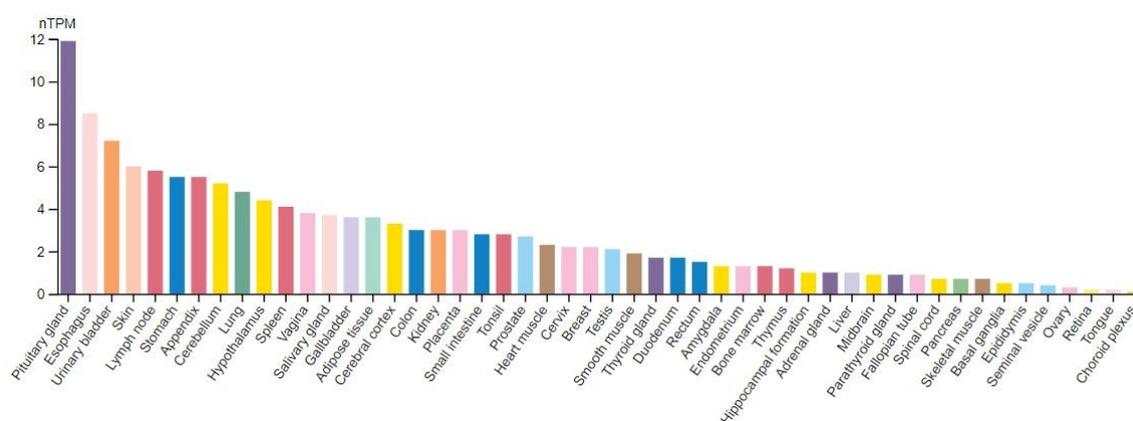


Figure 10: The normalized expression (nTPM) levels of the GPR8 mitochondrial RNA (mRNA) in 55 different tissue types. This figure has been reused from the Human Protein Atlas database [124]. The website (<https://www.proteinatlas.org/ENSG00000119714-GPR68/tissue>) was accessed on 05.03.2024.

Overall, GPR68 is expressed quite ubiquitously at a low to moderate level <sup>[170]</sup>. It has been observed to be highly expressed in the pituitary gland, small intestine, placenta, lungs, brain, and kidney <sup>[118, 180]</sup> (Figure 10). GPR68 is highly expressed in certain cell types like the immune cells like neutrophils <sup>[181]</sup> and T-cells <sup>[182]</sup>, smooth muscle cells of the aorta <sup>[183]</sup> and esophageal mucosa <sup>[184]</sup>, skeletal myocytes and epithelial cells of lungs and intestines <sup>[112]</sup>. GPR68 overexpression has been observed in various pathological states like solid tumors <sup>[170, 185, 186]</sup>, inflammatory bowels disease (IBD) <sup>[187]</sup>, and idiopathic pulmonary fibrosis <sup>[188]</sup>. Interestingly, a single residue mutation L74P in GPR68 causes *amelogenesis imperfecta* <sup>[189]</sup>.

### 1.3.3 Proton sensing

The proton sensing function of GPR68 was initially linked to the extracellular histidines <sup>[112]</sup> (Figure 11). A study based on cytosolic inositol monophosphate 1 (IP<sub>1</sub>) accumulation involving GPR68 single mutations of the histidine to phenylalanine showed that H17<sup>1.28</sup>, H20<sup>1.31</sup>, H84<sup>2.67</sup>, H169<sup>ECL2</sup> and H269<sup>7.36</sup> were described crucial for the proton sensing function of the receptor, whereas mutation of H89<sup>23.52</sup>, H159<sup>4.63</sup> or H175<sup>ECL2</sup> had no significant effect on the proton sensing function <sup>[112]</sup>. The proton sensing function was inhibited by the single mutants H269<sup>7.36</sup> and H245<sup>6.52</sup>, where H245<sup>6.52</sup> was proposed to be important for the structural integrity of the receptor based on homology modeling <sup>[112]</sup>. A four-histidine mutant of H17<sup>1.28</sup>, H20<sup>1.31</sup>, H84<sup>2.67</sup> and H269<sup>7.36</sup> where all residues are mutated to phenylalanine, does not completely abolish the receptor function. Interestingly, some receptor activity was regained after exposure to more acidity (~pH 5.6) for the above-described mutants <sup>[112]</sup>. Additionally, early homology modelling has predicted hydrogen bonds between H17<sup>1.28</sup>-H84<sup>2.67</sup> and H20<sup>1.31</sup>-H269<sup>7.36</sup> <sup>[112]</sup>. A separate study examining the function of extracellular histidines based on cAMP accumulation revealed that the majority of single extracellular histidines mutants involving alanine and phenylalanine exhibit distinct behavior and suggested that the histidines may have a lesser impact on protonation compared to their role maintaining the structural integrity of the receptor. H20 and H169 were identified as the most important histidines individually crucial for proton sensing function of the receptor. The joint replacement of seven histidines by alanine (H17<sup>1.28</sup>, H20<sup>1.31</sup>, H84<sup>2.67</sup>, H89<sup>23.52</sup>, H159<sup>4.63</sup>, H169<sup>ECL2</sup>, H175<sup>ECL2</sup>) decreased the efficacy to less than 10% compared to the wild-type (WT) GPR68 and a 300-fold decrease in potency <sup>[190]</sup>. Later studies revealed that the buried acidic triad (D67<sup>2.50</sup>, E149<sup>4.53</sup>, D282<sup>7.49</sup>) (Figure 11) was

responsible for the proton sensing function of GPR68, where E149<sup>4.53</sup> was predicted to have an upshifted pK<sub>a</sub> [134]. Although, GPR68 has been suggested to be involved in co-incidence detection of protons [134, 145, 191] suggesting an alternate endogenous ligand.

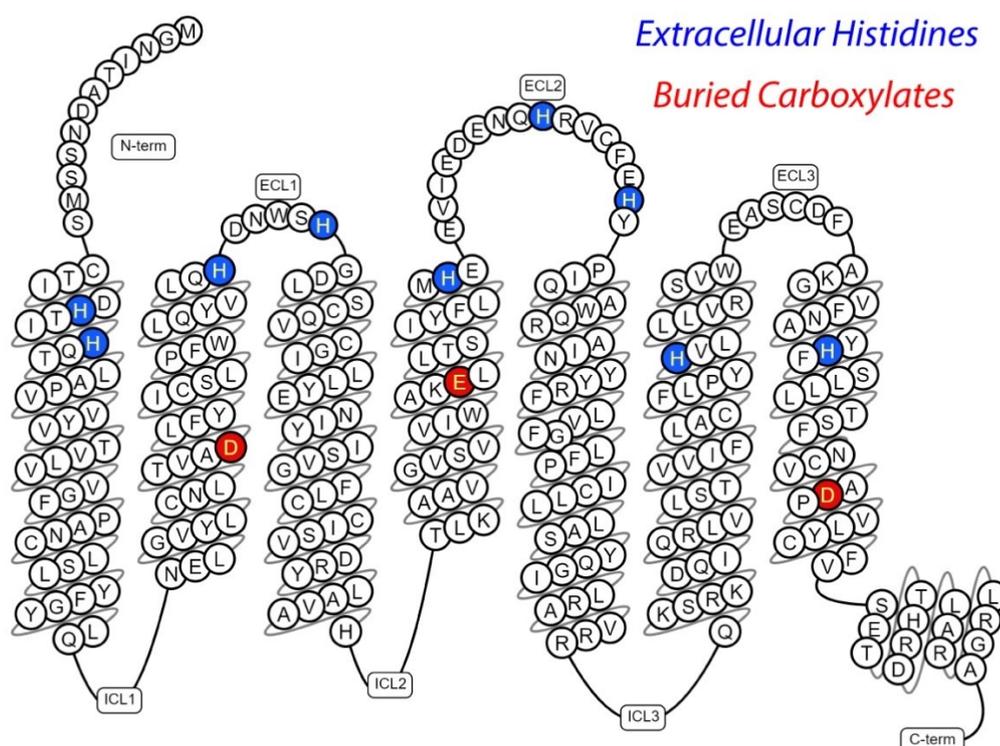


Figure 11: The snake diagram of GPR68 highlighting the proposed structural elements of GPR68 for pH sensing. This figure was prepared using the GPCRdb website [6].

### 1.3.4 Mechanosensing

Mechanosensing is suggested to be an additional putative function of GPR68 [125, 192]. Mechano-sensation may occur via membrane stretch or shear stress [193, 194]. GPR68 may potentially recognize shear stress of the blood flow via *N*-terminal glycosylation like  $\beta$ -adrenergic receptor [195] or via influence of membrane stretch by the movement of the amphipathic helix 8 [196]. Additionally, a study where all the crucial histidines, H17<sup>1.28</sup>, H20<sup>1.31</sup>, H84<sup>2.67</sup>, H169<sup>ECL2</sup>, and H269<sup>7.36</sup>, were changed to phenylalanine, showed a loss in both the proton sensing and mechanosensing function of GPR68 [125].

### 1.3.5 Modulation via divalent metal ions

Varying results have been reported for GPR68 modulation via divalent metal ions. Earlier studies suggested  $Zn^{2+}$  and  $Cu^{2+}$  ions to inhibit GPR68 based on  $G\alpha_q$  signaling when it is fully activated at approximately pH 6.9 [112]. Another  $G\alpha_q$  signaling based study indicated  $Fe^{2+}$ ,  $Zn^{2+}$ ,  $Co^{2+}$ ,  $Ni^{2+}$  and  $Mn^{2+}$  as agonists as they activate GPR68 at neutral pH [197]. Similar results were reported by a different group that investigated cAMP based signaling [171]. Although, the agonistic behavior of divalent metal ions was not observed for GPR68 when checked for signaling via the  $G\alpha_{12}/G\alpha_{13}$  pathway [198]. More recent studies showed that GPR68 can be allosterically modulated by divalent metal ions ( $Cd^{2+}$ ,  $Co^{2+}$ ,  $Cu^{2+}$ ,  $Fe^{2+}$ ,  $Mn^{2+}$ ,  $Ni^{2+}$ , and  $Zn^{2+}$ ) and the modulation is influenced by pH and the concentration of metal ions [171]. Studies that performed homology modeling of GPR68 with two disulfide bonds suggested a tetrahedron model for metal ion co-ordination, where the four extracellular histidines (H17<sup>1.28</sup>, H20<sup>1.31</sup>, H84<sup>2.67</sup>, and H169<sup>ECL2</sup>) could co-ordinate a zinc ion [171]. Additionally, GPR68 with multiple extracellular histidine replacements (H17<sup>1.28</sup>F, H20<sup>1.31</sup>F, H84<sup>2.67</sup>F, H169<sup>ECL2</sup>F, H269<sup>7.36</sup>F) did not respond to metal ions suggesting the intertwined role of protonation and metal ion modulation [171].

### 1.3.6 Small molecule ligands

In addition to the above-mentioned metal ions, there have also been reports of various small molecule ligands that can act as positive allosteric modulators of GPR68. A yeast-based screen revealed that benzodiazepines like lorazepam, diazepam and, ogerin are non-selective modulators for GPR68 [175] (Figure 12), and lorazepam promotes biased signaling of GPR68 via the cAMP pathway [199]. Docking studies of lorazepam using a homology model of GPR68 followed by mutagenesis studies highlights that lorazepam may contact polar residues like E160<sup>4.64</sup>, R189<sup>5.42</sup>, Y244<sup>6.51</sup> and Y268<sup>7.35</sup>, and form non-polar contacts with W77<sup>2.60</sup>, L101<sup>3.32</sup>, F173<sup>4.51</sup>, H269<sup>7.36</sup> and L272<sup>7.39</sup> [175]. An improved form of ogerin, called “Compound 71” was later synthesized that is 33-fold more potent than ogerin [200] (Figure 12). Three peptide ligands were discovered that are positive allosteric modulators of GPR68: osteocrin, cocaine and amphetamine-regulated protein (CART), and rat corticotropin [21]. The osteocrin peptide is reported to be 2-fold more potent than ogerin [21].

A dibenzazepine derivative that acts as a GPR4 antagonist had an antagonistic impact on GPR68 at concentrations exceeding 1  $\mu\text{M}$  [201]. A vinca alkaloid, conophylline may inhibit GPR68 in cancer associated fibroblasts [202]. Recently, the inhibitory effects of *Cephalotaxus harringtonia* var. *nana* extract along with homoharringtonine on the generation of inflammatory cytokines have been reported with the mechanism being dependent on the expression of GPR68 [203]. A small molecule screen against zebrafish GPR68 led to a discovery of a selective inhibitor called as ogremorphin [204] (Figure 12).

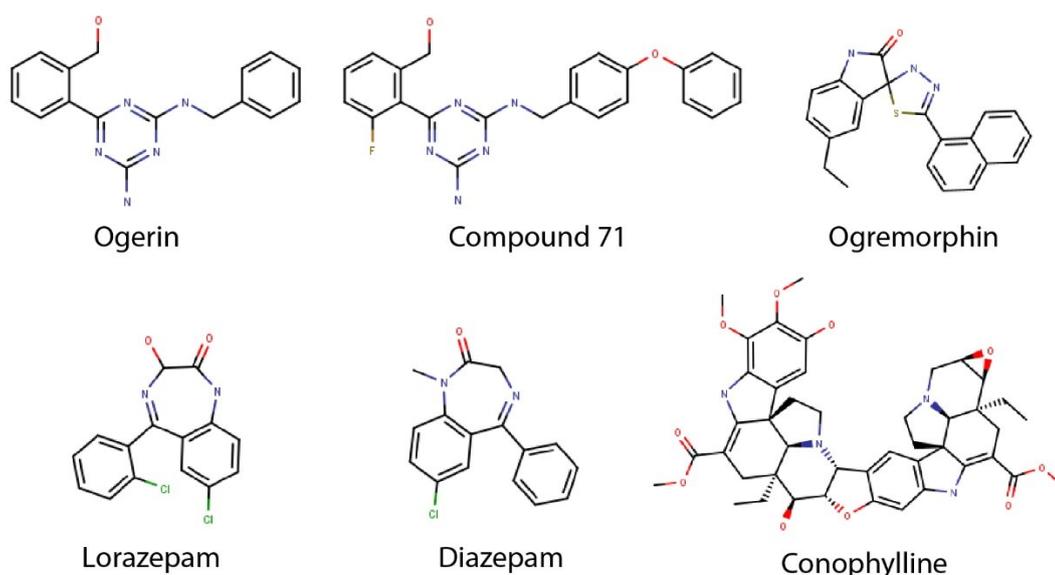


Figure 12: 2D structures of GPR68 antagonists (Ogremorphin) and allosteric modulators (Lorazepam, Diazepam, Ogerin, Compound 71 and Conophylline). The structures are represented without explicit hydrogens.

## 1.4 Working with G Protein-Coupled Receptors

It has been estimated that 60–85% of human GPCRs with potential therapeutic benefits have not yet been targeted with drugs [205]. The paths to reach a potential novel drug for a GPCR are mostly via a combined effort in functional *high-throughput* screening of compound libraries and structure-based *in silico* drug discovery [206–208]. Both the routes are lined up with complex challenges. This sub-section highlights the challenges faced during the structure-based drug discovery route.

Several techniques are the core of drug discovery research such as protein structure determination via X-ray crystallography, small angle X-ray scattering (SAXS), nuclear magnetic resonance (NMR) and cryogenic electron microscopy (cryo-EM), molecular dynamics simulations and molecular docking for structure-based drug discovery [209–211]. In

scenarios where the 3D structure is not determined for GPCRs, homology or structural modeling has been extremely helpful for gaining insights into such GPCRs [3, 212]. Recently developed AI-based structural modeling tools like AlphaFold [213] are expected to play a huge role in drug discovery in the future [214]. Other widely used techniques to study and characterize GPCRs and perform ligand screening are radioligand binding assays [215], DNA encoded libraries [216], surface plasmon resonance [217], resonance energy transfer [218] and other assays for detecting receptor activation [219].

#### 1.4.1 Expression and purification

Selection of an appropriate heterologous expression of human GPCRs is a crucial as different expression systems come with different drawbacks, and the expression system can have a substantial influence on the protein purity and yields. The initial challenge of working with the membrane proteins like GPCRs is that they are difficult to express and produce in sufficient quantities for structural studies due to various factors like their complex structure and poor stability which stems from several reasons such as the membrane embedded hydrophobic regions of the receptor along with intrinsically disordered regions like loops and termini [220, 221], and the inherent propensity of GPCRs to be highly dynamic proteins, as they rely on conformational change for their function. Additionally, a low expression level is a frequent problem when working with GPCRs. GPCRs are well known to contain post-translational modifications which further complicates the expression and production of GPCRs. The importance of cholesterol binding to GPCRs and its role in allosteric modulation [222] makes it challenging to work with expression systems like *Escherichia coli*. Although, successful GPCR expression in *Escherichia coli* have been reported for some GPCRs like rat neurotensin receptor [223] in inclusion bodies, human adenosine receptor [224] with an antagonist and fusion protein modification, and human opioid receptors [225]. Furthermore, another major set of challenges arrive during downstream protein purification process where the GPCRs when extracted from their native environment (membrane) tend to be highly unstable which leads to unfolding or degradation of the GPCRs [226].

The baculovirus expression system in insect cells is the most frequently used expression system for GPCRs and have yielded the highest number of GPCR structures [227]. Most GPCR studies use the following insect cell lines: *Spodoptera frugiperda* (Sf9), *Spodoptera*

*frugiperda* (Sf21) or *Trichoplusia ni* (High Five or Hi5) for receptor overexpression and production [227]. Insect cells are more expensive and complex to work with in comparison to the *E. coli* or yeast-based systems, but they are cheaper and easier to work with than mammalian cell lines [228]. Nevertheless, insect cells are easier to scale-up for high yields and the receptors overexpressed in insect cells are likely to have the correct post-translation modifications necessary for structural and functional integrity. Another inherent drawback of insect cell expression systems is the variations in the plasma membrane lipid composition in comparison to mammalian cells. The membranes of insect cells have a low cholesterol level, a high quantity of phosphatidylinositol, and an absence of phosphatidylserine [226, 229]. Furthermore, the lytic pathway of viral infection might lead to protein degradation [226, 230]. Despite the limitations of insect cells, they are the most popular choice of expression system for pursuing structural studies of GPCRs.

Mammalian cells lines like human embryonic kidney (HEK) and Chinese hamster ovary (CHO) are the next widely used choice in GPCR studies. These cell lines can be used for GPCR functional assays usually as adherent cells and for protein overexpression as suspension culture [226, 231]. Mammalian cell lines provide a native-like environment to work with human GPCRs, as they have the most favorable lipid composition of the plasma membrane and enzymes for post-translational modifications. Although, mammalian cells provide lower yields in comparison to other expression systems, and they are more expensive and difficult to work with [226, 232].

After the selection of an appropriate expression system, the next challenge is to utilize techniques to improve overexpression of the receptor and the successful purification of a stable and correctly folded protein. Below is a compilation of commonly employed strategies for increased cell-surface expression of GPCRs and overall protein stabilization throughout the purification process.

GPCRs that were expressed in high amounts were discovered to possess optimal codons for efficient translation, whereas the GPCRs expressed in low amounts were found to have sub-optimal codons which decreases translation [232]. Thus, codon optimization by replacement of sub-optimal codons with more frequently used codons in a particular expression system is a frequently used strategy for GPCRs. Targeted expression of GPCRs to the plasma membrane is often achieved by using the influenza derived hemagglutinin

signal peptide (MKTIIALSYIFCLVFAA) [233]. Overall improved expression for GPCRs has been witnessed by addition of the Kozak sequence (GCCACCATGG) at the 5' end for eukaryotic expression systems [226] and by usage of a strong promoter e.g., polyhedrin [234] for insect cell expression, and human CMV (cytomegalovirus) promoter for mammalian cell expression [235].

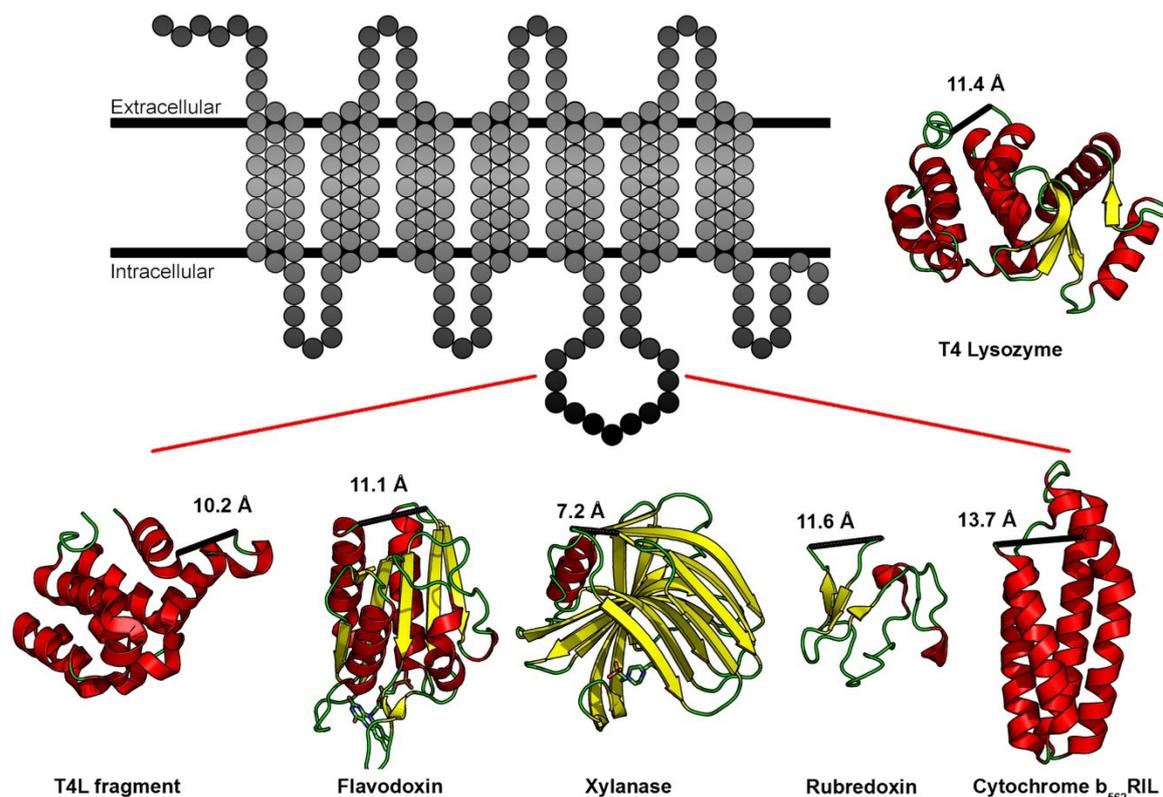


Figure 13: This figure illustrates commonly used fusion proteins which are inserted in the ICL3 of a GPCR, shown as a snake diagram. This figure has been reused [236].

Modifying the receptor is a major strategy employed for achieving increased receptor stabilization or for enabling smooth receptor purification. Addition of fusion proteins in the intracellular loops (ICLs), mainly ICL3, acts as a two-point fusion attachment strategy and assists in further stabilization of the GPCRs (Figure 13) [236]. A few examples of commonly used fusion proteins are rubredoxin (PDB ID: 1FHM), C-terminal fragment of T4 lysozyme (PDB ID: 2O7A), xylanase (PDB ID: 2B45), flavodoxin (PDB ID: 1I1O), thermostabilized apocytochrome b<sub>562</sub>RIL (BRIL) (PDB ID: 1M6T) and thermostable glycogen synthase domain from *Pyrococcus abyssi* (PGS) (PDB ID: 2FBW) [236, 237]. Recently, a three-point fusion attachment strategy using rat calcineurin (PDB ID: 4IL1) was established mainly aimed at cryo-EM based structure determination [238]. It is important to note that this strategy likely favors the inactive state of the receptors because of the modification made

in ICL3 which is an important interaction site for G proteins coupling, and it also has variable lengths amongst GPCRs as well as high flexibility <sup>[236]</sup>. Furthermore, addition of fusion in ICL3 is known to improve crystal contact formation or fusion can act as a fiducial marker during 2-dimensional (2D) classification in cryo-EM data processing e.g., addition of calcineurin fusion <sup>[238]</sup> or BRIL fusion in ICL3 allows for complexation with Anti-BRIL FAB and Anti-BRIL-FAB nanobody <sup>[237]</sup>. Another strategy often implemented for increased receptor stability and structure determination via cryo-EM is the co-expression of G proteins together with the target receptor to generate an active state complex that may be more stable than just purifying the receptor alone <sup>[239]</sup>. Specifically, addition of BRIL fusion to the *N*-terminus of recombinantly expressed GPCRs has proven to improve overall overexpression and stability of receptors <sup>[240]</sup>. Moreover, truncating the termini or other flexible region strategically may also improve protein expression, reduce chances of protein degradation, and promote overall protein stability <sup>[241, 242]</sup>.

Apart from these modifications, strategies like usage of receptor specific ligands or other modulators like metal ions that bind to the target receptor can be added during the protein expression and purification process to impart additional stability to the receptor <sup>[243]</sup>. Lastly, mutations of residues that form the conserved motifs can be mutated to lock and stabilize the receptor in a particular conformation <sup>[244]</sup>. Other strategies can be employed to identify point mutations which likely increase receptor thermostability are alanine/leucine scanning <sup>[245]</sup>, and directed evolution mutation approach <sup>[245]</sup>. Additionally, mutation of residues can be performed based on mutation prediction tools like CompoMug <sup>[246]</sup>, which uses sequence and structure-based criteria along with machine learning to help predict thermostabilizing mutations.

Some modifications to the receptor sequence are also made to specifically ease the purification process like addition of tags for enabling purification of recombinantly expressed protein (e.g., poly-histidine tag, FLAG tag, Strep-tactin tag etc.) <sup>[243, 247, 248]</sup>. Protease cleavage sites for endopeptidases may be added after tags or fusion proteins mainly when introduced at either terminus to assist in remove or cleavage of the protein if required <sup>[249]</sup>. Moreover, thiol modification <sup>[250]</sup> may improve receptor stability if the receptor contains free cysteines. In GPCR purification protocols, DTT is often used for reversible modification and iodoacetamide for irreversible modification of cysteines.

Extracting GPCRs from their native environment in the cell is the first essential step of the purification and is termed as solubilization. The most used reagent for solubilization of GPCRs are detergents. The charge of the detergents and the length of the hydrophobic tail (the shorter the tail, the harsher the detergent) play an important role in the ability to extract the proteins from the cells. Ionic detergents are considered harsh and can either be anionic (FC-12: n-dodecylphosphocholine 12) or cationic (LDAO: N,N-Dimethyl-n-dodecylamine N-oxide). The zwitterionic detergents like CHAPS: 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate are less harsh than the ionic detergents but more harsh than the non-ionic detergents. Non-ionic detergents like DDM: n-dodecyl- $\beta$ -d-maltopyranoside or LMNG: lauryl maltose neopentyl glycol are considered as mild detergents. The mild detergents may not affect the protein-protein interactions but may impact lipid-lipid or protein-lipid interactions, whereas harsh detergents even though more effective in protein extraction may cause the denaturation of the membrane protein <sup>[251, 252]</sup>. DDM and LMNG are the most widely used detergents for GPCR studies in combination with CHS: cholesteryl hemisuccinate (a detergent often used to replace cholesterol during the purification process) <sup>[247]</sup>. A major drawback of using detergents is that they replace the lipids which are bound to the target receptor (annular lipids), which may contribute to additional instability of the receptor <sup>[253]</sup>. Other membrane mimetic tools that are popularly used to stabilize GPCRs are bicelles <sup>[254, 255]</sup>, amphipols <sup>[256]</sup>, synthetic co-polymers like SMALPS <sup>[257]</sup>, saposin-lipoprotein nanoparticles like Salipro <sup>[258]</sup>, nanodiscs <sup>[259]</sup>, liposomes <sup>[260]</sup>, and peptidiscs <sup>[261]</sup>.

#### **1.4.2 Structure determination techniques**

The two most widely used techniques for protein structure determination are cryo-EM <sup>[262]</sup> and X-ray crystallography <sup>[263]</sup>. Structure determination of GPCRs is crucial for structure-based drug discovery to yield and optimize novel compounds. Within the last years there has been a significant rise in the number of cryo-EM determined structures of GPCRs in comparison to X-ray crystallography <sup>[264]</sup>.

The significant bottleneck for both the structure determination technique is a stable, pure, and monodispersed protein <sup>[265, 266]</sup>. X-ray crystallography of GPCRs is carried out using lipidic cubic phase (LCP) crystallography <sup>[267, 268]</sup>. Obtaining crystals can be time consuming and may even take years, as it is an iterative process. This iterative process may

involve optimizing the protein construct with mutations for thermostabilizing the receptor [269], different fusion proteins for establishing optimal crystal contacts and overall stability [236], and the receptor that remains stable throughout the crystal formation process. The flexibility of the construct, difficulties in establishing crystal contacts, and the time required for obtain the first crystals can be very difficult to address [264, 270].

With single particle cryo-EM a lot of the problems of X-ray crystallography are solved as it does not require crystals. In 2017, the first structure of a GPCR was determined using cryo-EM [271]. Typically, cryo-EM as technique is optimal for structure determination if the size of the target protein is ~100 kDa, but GPCRs usually have a molecular weight in the range of 35-45 kDa [272]. To circumvent this problem, GPCRs are often complexed with G proteins [273] or  $\beta$ -arrestin [274] which results in a complex size of ~110-135 kDa or the GPCR can be modified by adding a BRIL fusion protein in the ICL3 which can further complex with a BRIL-specific antibody FAB fragment along with an anti-BRIL-FAB nanobody [275, 276] leading to a complex size of ~120 kDa. Another strategy involved inserting a chimeric BRIL with an extra linker sequence derived from the  $A_{2A}$  adenosine receptor. These approaches have been previously described and proven effective [236].

The overall stability issue for the structure determination techniques can be counteracted with the usage of a ligand, which preferentially locks the receptor in a stable conformation [264, 277]. Cryo-EM being a faster technique lessens the burden of the overall stability for the protein in comparison to structure determination via X-ray crystallography. Other advantage of cryo-EM is that it requires less amounts of protein and is more tolerant towards sample heterogeneity and presence of small amounts of contaminants [264].

Before the last decade, the resolution obtained with X-ray crystallography was higher than cryo-EM, but the development in the past decades have significantly improved the resolutions of the protein map that can be determined via cryo-EM [278-280]. Thus, opening new avenues for structure determination of complex proteins.

### 1.4.3 Molecular dynamics simulations

Atomic resolution structures of GPCRs are crucial for structure-based drug discovery and for establishing a structure-function relationship. At the same time, one must consider that GPCRs are highly dynamic molecules whose function depends on their propensity to undergo conformational changes (section 1.1). Additionally, some information about the GPCR functionality may be lost due to modifications by fusion proteins, tags, truncations or point mutations for improved thermostability, and placement into a non-native environment e.g., detergents (section 1.4.1).

Molecular dynamics (MD) simulations can be used to complement experimental structural data or functional assays. The first MD simulation performed for GPCRs was for the dopamine D2 receptor [281]. Technology improvements in CPUs, GPUs, and parallel processing within GPUs in the last decades have enabled simulations of larger systems [282, 283]. Additionally, continuous developments of molecular dynamics simulation and other analysis tools have significantly helped investigating GPCRs. A few examples of such tools are: the widely used input generation tool, CHARMM-GUI [284], internal water molecules prediction tool for GPCRs, HomolWat [285], GPCR specific structural modeling tool, GOMoDo [212, 286], tool for analyzing MD simulations, MDAnalysis package [287], and the GPCR molecular dynamics database, GPCRmd [288]. MD simulations have improved the understanding of GPCRs from a pharmacological and GPCR conformational dynamics perspective to unravel different binding modes for allosteric modulators, the role of the loop regions, and G protein coupling interfaces [289]. Additionally, post-translational modifications like palmitoylation [290], the influence of cholesterol and the lipid bilayer composition on GPCRs [291, 292], and GPCR oligomerization [293] were simulated using molecular dynamics.

Moreover, MD has proven to be helpful in this investigation of conserved water molecules in GPCRs [294] that revealed the formation of internal continuous water wires with significant residence times [295], and how these water molecules support conformational changes within GPCRs for receptor activation in co-ordination with conserved GPCR motifs [163]. The role of the sodium ion as a negative allosteric modulator with investigations into residues influencing the kinetics of sodium binding, alternate entry points for sodium

into GPCRs, and importance of sodium exchange for GPCR activation was established by work performed using MD <sup>[169]</sup>.

Techniques like steered MD may be useful in understanding of the transitions between receptor conformations, but these are likely to be less unbiased and stressed to fit into the non-physiological time scales that are typically required for these events. Often, information about the atomic details is lost when using coarse-grained MD simulations to investigate broader questions about the conformational landscape over longer timeframes. Even though, MD requires experimental validation, but still forms a vital tool when investigating questions related to GPCRs that cannot be addressed due to experimental limitations <sup>[296]</sup>.

## 2. Aims & Objective

The primary objective of this thesis was to determine the structure of the inactive state of GPR68 and to perform functional studies to understand the mechanisms involved in proton sensing.

GPR68 is an orphan proton sensing GPCR that senses protons, but the proposed proton sensing mechanisms of GPR68 lack consensus [112, 134, 144]. GPR68 is a therapeutically relevant target as it is overexpressed in various pathological states [297–299], and acts as a marker in different types of cancer [170]. Currently, only positive allosteric modulators have been described for the receptor [21, 175].

The objective of my thesis was to obtain the structure of GPR68 in its inactive state, with an antagonist bound. An atomic resolution 3D structure of GPR68 would enable me to establish structure function relationships and would be valuable for structure-based drug discovery. To determine the structure of GPR68, lipidic cubic phase crystallization for X-ray crystallography and single particle cryo-EM should be used. For crystallography, GPR68 constructs with N-terminal or ICL3 fusions should be generated to promote crystal contacts. For cryo-EM, primarily GPR68 constructs with ICL3 BRIL fusion should be used, which can be complexed with Anti-BRIL FAB and an Anti-BRIL FAB nanobody to increase the receptor's size to approximately 110 kDa, as this modification would facilitate cryo-EM analysis.

Whilst pursuing structure determination, a structural model of GPR68 should be investigated using classical MD. This involves comparing a standard protonation simulation to a set of protonation simulation where specific residues were protonated. Additionally, the Bridge2 software [300, 301] should be used to generate and investigate the dynamic water-mediated hydrogen bond networks from each simulation to further enhance our understanding of the proton sensing mechanism. Further, a homogeneous time-resolved fluorescence assay [302] should be established to assess the functionality of GPR68 using IP<sub>1</sub> accumulation via the G $\alpha_{q/11}$  signaling pathway as a measure of receptor activation. This assay should complement the computational work by testing the function of identified, putatively relevant residues.

During the thesis work, an artificial intelligence based structural modeling tool, AlphaFold was first described <sup>[213]</sup>. Its description compelled me to incorporate AlphaFold into my research to keep up with the latest advancements in the field and to test the potential of this software to predict rigid continuous helices when substituting an intracellular loop with a fusion protein and investigating the behavior of AlphaFold based structural models in classical molecular dynamics.

## 3. Materials & Methods

### 3.1 Materials

#### 3.1.1 GPR68 gene synthesis

All the constructs designed for the investigation of GPR68 were synthesized by GeneArt™ (ThermoFisher Scientific). The GPR68 constructs were codon-optimized for the selected expression system i.e., HighFive™ cells (insect) or HEK293 cells (human). The GPR68 constructs that are designed for insect cell expression were subcloned in pFastbac1 using BamHI and EcoRI restriction sites, whereas the constructs designed for HEK293 cell expression were subcloned in pcDNA5/FRT, using BamHI and NotI restriction sites. Additional construct details include the presence of a Kozak sequence followed by a start codon at the *N*-terminus and a stop codon at the *C*-terminus. The amino acid sequences for the GPR68 constructs can be found in the Appendices (Supplementary 1).

#### 3.1.2 Cell lines

<u>Name</u>	<u>Supplier / Distributor</u>
<i>Escherichia coli</i> (DH10Bac) competent cells	Gibco™ (ThermoFisher Scientific)
<i>Escherichia coli</i> (DH5α) competent cells	Invitrogen™ (ThermoFisher Scientific)
HEK293H Cells	Gibco™ (ThermoFisher Scientific)
HEK293F Cells	Sigma Aldrich (Merck)
High Five™ Cells	Invitrogen™ (ThermoFisher Scientific)
Sf9 Cells	Gibco™ (ThermoFisher Scientific)

#### 3.1.3 Cell culture

<u>Name</u>	<u>Supplier / Manufacturer</u>
10X Heat-Inactivated FBS (HI FBS) Performance plus	Gibco™ (ThermoFisher Scientific)
Carbenicillin Disodium Salt	Gibco™ (ThermoFisher Scientific)

Cellfectin™ II Reagent	Gibco™ (ThermoFisher Scientific)
Dulbecco's Modified Eagle Medium (DMEM), high glucose, GlutaMAX™, pyruvate	Gibco™ (ThermoFisher Scientific)
Dulbecco's Phosphate Buffered Saline (No Ca <sup>2+</sup> or Mg <sup>2+</sup> )	Sigma Aldrich (Merck)
Expi293™ Expression Medium, GlutaMAX™	Gibco™ (ThermoFisher Scientific)
Gentamycin	Gibco™ (ThermoFisher Scientific)
Insect-XPRESS™ Protein-free Insect Cell Medium	Lonza
Kanamycin sulfate	Gibco™ (ThermoFisher Scientific)
Lipofectamine™ Transfection Reagent	Invitrogen™ (ThermoFisher Scientific)
Luria Bertani (LB) (Lennox) broth / agar	Sigma Aldrich (Merck)
Opti-MEM™ I Reduced Serum Medium	Gibco™ (ThermoFisher Scientific)
Penicillin-Streptomycin (10,000 U/mL)	Gibco™ (ThermoFisher Scientific)
Polyethylenimine (PEI) Max (MW 40,000)	PolySciences
Sf-900™ III SFM	Gibco™ (ThermoFisher Scientific)
Sodium butyrate	Sigma Aldrich (Merck)
Tetracycline	Sigma Aldrich (Merck)
TrypLE™ Express Enzyme (1X), no phenol red	Gibco™ (ThermoFisher Scientific)

### 3.1.4 Equipments

<u>Name</u>	<u>Supplier / Distributor</u>
Äkta FPLC/Explorer/Purifier	GE Healthcare
AQUALine AL 5 – water bath	Lauda
Avanti J-20 XP/ J-25/ J-26 XP centrifuge	Beckman Coulter
Axiovert 25	Zeiss
BBD6220 CO <sub>2</sub> -incubator	Heraeus
Benchtop pH meter pH 50 VioLab Basic	Carl Roth
Blue/Green LED Transilluminator FG-09WS	Nippon Genetics
Casy TTC Cell counter	Innovatis
Centrifuge 5415 R/ 5810 R/ 5702 R/ 5424 R	Eppendorf

Comfort / Premium Nofrost Fridge	Liebherr
Cooling cabinet unichromat 1500	UniEquip
Dionex Ultimate HPLC 3000	Thermoscientific
Dounce homogenizer	Wheaton
Dri-Block DB100/2	Techne
EasiGlow – glow discharge system	Pelco
EB170S incubator	Jouan
Econo-Column for chromatography	Bio-Rad
Ecotron HT Multitron standard/ Multitron	Infors
E-gel powerBase v4	Invitrogen™ (ThermoFisher Scientific)
ELx405 HT – washer	Bio-Tek
FINNPIPETTE F1	ThermoFisher Scientific
Freezer –80°C	Sanyo
Freezer –80°C HERA freezer	Heraeus
Freezer –80°C TSX Series	Thermo Scientific
Glass syringes for LCP	Hamilton
Heracell 240i CO2 incubator	ThermoFisher Scientific
Heracell™ 150i CO <sub>2</sub> -incubator	ThermoFisher Scientific
Highlight 3001 / SZX16	Olympus
HiLoad Superdex 200 16/600 preparative grade	Cytiva
iBlot 2	Invitrogen™ (ThermoFisher Scientific)
Innova 4000 incubator shaker	New Brunswick Scientific
Lab armor bead bath	VLM GmbH
Laminar Air Flow BDK	Luft und Reinraum Technik GmbH
Lipidic cubic phase (LCP) mosquito	SPT Labtech
Magnetic Stirrer RH basic 2	IKA
MicroLoops LD	Mitogen
Multidrop Combi	ThermoFisher Scientific
Multifuge X1R centrifuge	Thermo Scientific
Mx3005P	Stratagene
NanoDrop One	Thermo Scientific
Nitrogen decompression 4635 cell disruption vessel	Parr Instrument Company

Pheonix liquid handler	Art Robbins Intruments
PHERASTAR FSX	BMG Labtech
PIPETBOY 2	Integra
Pipettes (2.5 µL, 10 µL, 20 µL, 100 µL, 200 µL, 1 mL)	Eppendorf
PowerPac Basic	Bio-Rad
Powerpack 25 T	Biometra
Prometheus NT.48	NanoTemper Technologies GmbH
RI-1000 incubator and imager	Formulatrix
Rotator drive STR4	Stuart scientific
Rotor HE121, IL117	Eppendorf
Rotor JA 25.50/ 9.100/ 8.100	Beckman Coulter
See-Saw Rocker SSL4	Stuart Scientific
Sonorex RK	Bandelin
Sprout – Mini centrifuge	Biozym
Superdex 200 increase 10/300 GL	Cytiva
Superose 6 increase 10/300 GL	Cytiva
Superose 6 increase 5/150 GL	Cytiva
Swinging bucket rotor – JS 5.3	Beckman Coulter
T3 Thermocycler	Biometra
ThermoMixer C	Eppendorf
Unix-C SEC-300 4.6/300	Sepax Technologies
Unstirred water bath	Grant
Vitrobot Mark IV – plunge freezer	ThermoFisher Scientific
Vortex VV3	VWR
Weighing balance ED6202S/ TE214S/ ED4202-CW/ ED6202S-OCE	Sartorius

### 3.1.5 Disposables

<u>Name</u>	<u>Supplier / Distributor</u>
Amicon – Ultra centrifugal filters (Molecular weight cutoff: 30 kDa, 50 kDa, and 100 kDa)	Sigma Aldrich (Merck)
Anti-FLAG M2 affinity gel	Sigma Aldrich (Merck)

Cell culture petri dish (100 mm × 20 mm)	Corning
Corning, CoStar, Stripette serological Pipettes (5 mL, 10 mL, 25 mL, 50 mL)	Sigma Aldrich (Merck)
DeepWell Block (96-well)	ThermoFisher Scientific
Disposal bag	Sarstedt
E-gel 1.2%	Invitrogen™ (ThermoFisher Scientific)
Erlenmyer flask (250 mL)	Corning
Falcon – plasma vacuum gas treated tissue culture flasks (25 cm <sup>3</sup> , 75 cm <sup>3</sup> , 175 cm <sup>3</sup> )	Corning
Falcons – High clarity polypropylene tubes (15 mL, 50 mL)	Corning
Laminex Glass base (0.2 mm) and cover	Molecular Dimensions
Matrix™ Reagent Reservoirs (10 mL)	ThermoFisher Scientific
Monolith NT.115 series High sensitivity capillaries	NanoTemper
Multiwell plates (6-well)	Corning
NuPAGE™ 4-12% Bis-Tris Mini Gel	Invitrogen™ (ThermoFisher Scientific)
Parafilm	Bemis Flexible packaging
Pipette Tips (10 µL, 200 µL, 1 mL)	Eppendorf
Poly-D-lysine Solid white flat bottom plate (384-well)	Corning
Precision wipes	Kimtech / Kimberly-Clark professional
Purple Nitrile Xtra gloves	Kimtech
Quantifoil R0.6/1.0 300 Au	Quantifoil
Quantifoil R0.6/1.0 300 Cu	Quantifoil
Quantifoil R1.2/1.3 300 Au	Quantifoil
Quantifoil R1.2/1.3 300 Cu	Quantifoil
Roller bottles (850 cm <sup>3</sup> )	Corning
Safe-Lock Tubes (0.5 mL, 1.5 mL, 2 mL, 5 mL)	Eppendorf
Safeseal surphob pipet tips (10 µL, 100 µL, 300 µL 1 mL)	Biozym
Soft tissues	Henry Schein
Strep-Tactin XT 4Flow high capacity	IBA Lifesciences
TALON superflow cobalt-based resin for immobilized metal affinity chromatography (IMAC)	Clontech (Takara Biotech)
Tissue culture flask Nuclon delta surface (25cm <sup>3</sup> , 75 cm <sup>3</sup> , 175 cm <sup>3</sup> )	ThermoFisher Scientific

TouchNTuff nitrile gloves	Ansell
Zeba spin desalting columns	ThermoFisher Scientific

### 3.1.6 Chemicals and reagents

<u>Name</u>	<u>Supplier / Distributor</u>
(2-(N-morpholino)ethanesulfonic acid) (MES)	Carl Roth
1-oleoyl-rac-glycerol (Monoolein)	Sigma Aldrich
3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulfonate] (CHAPS)	Anatrace
3-isobutyl-1-methylxanthine (IBMX)	Sigma Aldrich
3x FLAG peptide	NeoBiotech
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	Sigma Aldrich
5-bromo-3-indolyl- $\beta$ -D-galactopyranoside (Bluo-Gal)	ThermoFisher Scientific
5x Colorless GoTaq Flexi Buffer	Promega
7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM)	ThermoFisher Scientific
Adenosine triphosphate (ATP)	Sigma Aldrich
Bicine	Sigma Aldrich
Biotin	Sigma Aldrich
Bovine serum albumin	Sigma Aldrich
Cholesterol	Sigma Aldrich
Cholesteryl hemisuccinate tris salt (CHS)	Anatrace
cOmplete™ EDTA free protease inhibitor cocktail	Roche (Sigma Aldrich)
Copper chloride	Carl Roth
D-glucose	Fluka Analytical
Dithiothreitol (DTT)	SERVA Electrophoresis GmbH
DNA molecular weight marker XIV (100 bp ladder) or XVI (250 bp ladder)	Roche (Sigma Aldrich)
DNase I	Roche (Sigma Aldrich)
Ethane 2.5 HiQ minican	Linda
Ethanol	Honeywell
Ethylenediamine tetraacetic acid (EDTA)	Carl Roth
Forskolin	Sigma Aldrich
Fos-choline-12 (FC-12)	Anatrace

Glycerol	Sigma Aldrich
Glyco-diosgenin (GDN)	Anatrace
HT 96 MemChannel Eco Screen	Molecular Dimensions
HT 96 MemGold Eco Screen	Molecular Dimensions
HT 96 MemGold2 Eco Screen	Molecular Dimensions
HT 96 MemGoldMeso Eco Screen	Molecular Dimensions
HT 96 MemMeso	Molecular Dimensions
HT 96 MemTrans Eco Screen	Molecular Dimensions
Hydrochloric acid 32%	neoFroxx GmbH / Merck (Sigma Aldrich)
Imidazole	Merck (Sigma Aldrich)
Inositol monophosphate (IP <sub>1</sub> )	Sigma Aldrich
Iodoacetamide	Sigma Aldrich
Isopropanol	Aber GmbH
Isopropyl- $\beta$ -D-1-thiogalactopyranoside (IPTG)	ThermoFisher Scientific
JBScreen LCP HTS	Jena Biosciences
Laemmli sample buffer 2x / 4x	SERVA Electrophoresis GmbH
Lauryl maltose neopentyl glycol (LMNG)	Anatrace
Leupeptin	Roche (Sigma Aldrich)
Lithium chloride	Merck (Sigma Aldrich)
M13 forward primer (100 $\mu$ M) Cat# 11-2055-1/2	MWG-Biotech AG
M13 reverse primer (100 $\mu$ M) Cat# 11-2055-2/2	MWG-Biotech AG
Magnesium chloride	Merck (Sigma Aldrich)
Magnesium chloride, 25 mM	Promega
NBT-BCIP solution	ThermoFisher Scientific
n-dodecyl- $\beta$ -D-maltopyranoside (DDM)	Anatrace
Nickel (II) sulfate	Merck (Sigma Aldrich)
Octyl glucose neopentyl glycol (OGNG)	Anatrace
Pefabloc SC (4-(2-Aminoethyl)-benzolsulfonylfluoride-hydrochloride)	Roche (Sigma Aldrich)
Perfect Protein™ Markers, 15-150 kDa #69149	Merck (Sigma Aldrich)
Ponceau S	Sigma Aldrich
Potassium chloride	Sigma Aldrich
Potassium dihydrogen phosphate	Sigma Aldrich

Potassium hydroxide	Carl Roth
Potassium phosphate dibasic	Merck (Sigma Aldrich)
Precision plus protein dual color standards, 500 $\mu$ L #1610374	BioRad
Precision plus protein™ all blue prestained protein standards #1610373	BioRad
Precision plus protein™ unstained protein standards #1610363	BioRad
Quick Coomassie stain	SERVA Electrophoresis GmbH
Sodium azide	Sigma Aldrich
Sodium bicarbonate	Merck (Sigma Aldrich)
Sodium carbonate	Merck (Sigma Aldrich)
Sodium chloride	Carl Roth
Sodium citrate	Sigma Aldrich
Sodium dihydrogen phosphate monohydrate	Carl Roth
Sodium hydroxide solution (1 N)	Carl Roth
Sodium phosphate dibasic	Sigma Aldrich
Zinc chloride	Fluka Analytical

### 3.1.7 Commercial kits

<u>Name</u>	<u>Supplier / Distributor</u>
iBlot2 nitro-cellulose transfer stacks	Novex (ThermoFisher Scientific)
IP-One Gq kit	Revvity
Plasmid DNA miniprep/ midiprep/ gigaprep kits	Qiagen

### 3.1.8 Antibodies and enzymes

<u>Name</u>	<u>Supplier / Distributor</u>
Anti-GPR68 polyclonal, produced in rabbit Cat# 720277	Invitrogen (ThermoFisher Scientific)
Anti-mouse IgG (whole molecule) – alkaline phosphatase antibody produced in rabbit Cat# A4312	Sigma Aldrich
Anti-rabbit IgG (whole molecule) – alkaline phosphatase antibody produced in goat Cat# A3687	Sigma Aldrich
DYKDDDDK tag (Anti-FLAG) antibody produced in rabbit, Cat# 600-401-383	Rockland Immunochemicals

GoTaq® G2 Flexi DNA polymerase (5 U/μL), Cat# M780B	Promega
Penta-His antibody, BSA-free produced in mouse, Cat# 34660	Qiagen

### 3.1.9 Softwares

<u>Name</u>	<u>Supplier / Distributor</u>
Adobe acrobat reader 2023.008.20555	Adobe
Adobe Illustrator 2021	Adobe
Anaconda / miniconda 23.11.0-2	Anaconda Inc.
Bridge2 (H-Bond network analysis)	[300, 301]
Chromeleon 7	ThermoFisher Scientific
Coot 0.9.8.92	[303]
CryoSPARC and CryoSPARC Live 4.4.1	Structura Biotechnology Inc.
EPU Ver 3.2.0.4766REL	ThermoFisher Scientific
ExPASy server	[304]
Graphpad Prism 10	Graphpad Software Inc.
Marvin Sketch 17.24.3	Chemaxon
Microsoft Office 365 version 2308	Microsoft
MXPro	Stratagene
Nanoscale Molecular Dynamics 2.14	[305]
Phenix 1.21	[306]
PheraStar FSX Reader Control	BMG LabTech
PR_Thermocontrol 2.1.1 EN	NanoTemper Technologies GmbH
PyMOL 2.4.1	Schrodinger
Rockmaker 3.17.9.1	Formulatrix
TALOS, TIA, FluCam Viewer	ThermoFisher Scientific
Unicorn 7.8	Cytiva
Visual Molecular Dynamics 1.9.3	[307]

### **3.1.10 Small molecule ligands**

The small molecule ligands (antagonists) used in my thesis are Boehringer Ingelheim proprietary ligands. Dr. Matthias Hoffmann (Boehringer Ingelheim, Biberach, Germany) provided me with adequate amounts of the ligands. The proprietary ligands (Ligand 1, Ligand 2, and Ligand 3) were identified by screening human GPR68. Additionally, ogremorphin (CAS# 352563-21-4), a publicly available antagonist, identified by screening zebrafish GPR68 was used in my thesis.

## 3.2 Methods

### 3.2.1 Molecular biology

#### DNA Resuspension

Lyophilized GPR68 construct DNAs (5 µg) were resuspended in 50 µL of sterile water (theoretical concentration of 100 ng/µL).

#### Transformation & Blue-white screening

Competent cells aliquots were thawed on ice. 10 ng of plasmid DNA (pFastBac1 or pcDNA5/FRT) was mixed with 50 µL of *Escherichia coli* DH10Bac (for pFastBac1) or *Escherichia coli* DH5α (for pcDNA5/FRT) competent cells in 1.5 mL Eppendorf tubes. The mixture was incubated on ice for 30 min. The mixture was exposed to heat shock in a water bath at 42°C for 30 s and immediately put on ice for 5 min. The mixture was transferred to 5 mL round bottom tubes (Falcon, Corning) and 450 µL of super optimal broth with catabolite repression medium (S.O.C medium) was added. The tubes were then incubated in the shaker at 37°C for 4 h at 250 rpm.

10-20 µL of a culture with 100 µL of S.O.C medium was plated onto a LB agar plate containing 50 µg/mL kanamycin, 7 µg/mL gentamicin, and 10 µg/mL tetracycline, 40 µg/mL IPTG and 100 µg/mL Bluo-Gal for blue-white screening of pFastBac1 recombination, whereas LB agar plates with 100 µg/mL carbenicillin were used for pcDNA5/FRT vector recombinant screening. The plates were inverted and incubated at 37°C for 48 h. The pFastBac1 plates were additionally incubated at 4°C for 24-48 h.

For the transposed pFastBac1 vector, white colonies (2-3) were picked. Each colony was inoculated in 3 mL LB medium containing 50 µg/mL kanamycin, 7 µg/mL gentamicin, and 10 µg/mL tetracycline. The inoculated medium was incubated overnight at 250 rpm and 37°C in a shaker. After the incubation, the culture was centrifuged for 15 min at 3000 × g and 4°C. The supernatant was discarded, and the pellets were stored at -80°C or directly used for plasmid DNA extraction.

For the pcDNA5/FRT plates, one isolated single colony was inoculated into 80 mL LB medium containing 100 µg/mL ampicillin in a 250 mL baffled Erlenmeyer flask and grown

overnight at 120 rpm and 37°C in a shaker. This setup is henceforth termed as pre-culture. This pre-culture was then utilized to inoculate 4 × 1 L of LB medium containing 100 µg/mL ampicillin in 5 L baffled Erlenmeyer flasks in a ratio of 1:50 i.e., using a preculture inoculum of 20 mL per liter of culture. The Erlenmeyer flasks were incubated overnight for approximately 18 h at a speed of 110 rpm and a temperature of 37°C. After the incubation, the culture was centrifuged for 30 min at 3000 × g at 4°C. The supernatant was discarded, and the pellets were weighed and stored at –80°C or directly used for plasmid DNA extraction.

#### Plasmid/Bacmid extraction

The HiSpeed Plasmid Mini/Midi/Giga prep kits from Qiagen were used to extract the bacmid or plasmid DNA from the cells. After the bacmid DNA extraction, 1 µL (out of 40 µL) was removed for polymerase chain reaction to confirm successful recombination of DNA. The remaining DNA was stored at –20°C until further use.

#### Polymerase chain reaction (PCR)- Bacmid DNA

The master mix was prepared according to the following table and 49 µL of the master mix was added to the 1 µL of bacmid DNA:

<b>Components</b>	<b>1x</b>
Bacmid-DNA	1 µL
5x PCR Buffer	10 µL
Deoxynucleotide triphosphate (dNTP) mixture (10 mM each)	1 µL
Taq Polymerase (5 U/µL)	0.5 µL
MgCl <sub>2</sub> (25 mM)	3 µL
H <sub>2</sub> O	32.5 µL
M13 forward primer (20 µM)	1 µL
M13 reverse primer (20 µM)	1 µL
<b>Total</b>	<b>50 µL</b>

M13 forward primer sequence: 5'-CCCAGTCACGACGTTGTAAAACG-3'

M13 reverse primer sequence: 5'-AGCGGATAACAATTCACACAGG-3'

PCR Setup		
Steps	Time and Temperature	Cycles
Initial denaturation	5 min at 94°C	1x
Denaturation	45 s at 94°C	35x
Annealing	45 s at 55°C	
Extension	5 min at 72°C	
Final extension	7 min at 72°C	1x
Hold	at 4°C	-

After the PCR, the samples were confirmed for successful recombination using agarose gel electrophoresis using the E-gel system from Invitrogen. The PCR mixture (50  $\mu$ L) was mixed with 5  $\mu$ L of 10x loading buffer (50% glycerol, 10 mM EDTA, bromophenol blue). 15  $\mu$ L of the sample mixture was loaded in per well on the gel (E-gel 1.2% agarose - Invitrogen), together with the 100 bp DNA Molecular Weight Marker XIV (Roche) which was used as a standard. The gel was run for 30 min. The non-recombinant DNA was identified by the presence of a band at 300 bp which is the size of the bacmid. The bacmid which is transposed with the target of interest in the pFastBac1 vector has a size of ~2300 bp + size of target of interest.

### 3.2.2 Protein expression and production

#### *Cell maintenance and counting*

The cell maintenance was performed twice a week by laboratory scientists for Sf9 cells, High Five<sup>TM</sup> cells, and HEK293F cells. The Sf9 cells were cultured in ESF-900 III SFM medium (Gibco), High Five<sup>TM</sup> cells (Hi5) in Insect-XPRESS<sup>TM</sup> medium (Lonza), and HEK293F cells in Expi293<sup>TM</sup>, GlutaMAX<sup>TM</sup> (Gibco). The insect and mammalian cells were always examined for contamination, shape and size under a microscope and subsequently counted using a CASY cell counter (Innovatis). All the cell culture work was performed in laminar air flow under sterile conditions. The cell culture medium was always pre-warmed at 27°C either in the incubator or water bath at 27°C.

### Insect cells

#### Transfection

Sf9 cells were seeded at  $0.8 \times 10^6$  cells/well in a 6-well plate. The 6-well plate was incubated at 27°C for 30-45 min to allow the cells to adhere to the bottom of the well. The following solutions were prepared for 2 wells:

The Solution A comprised of 2 µg of GPR68 bacmid DNA (10 µl) in 200 µl of ESF-900 III SFM medium (Gibco), and the Solution B comprised of 16 µl of Cellfectin II reagent (Gibco) in 200 µl ESF-900 III SFM medium (Gibco). Solution B was added to the Solution A and this mixture was incubated for 30 min at room temperature. 213 µl of this DNA-Cellfectin mixture was added to each well and the 6-well plate was incubated for 4 h at 27°C. Then the medium was removed carefully from the surface of the cells, ensuring that the cells remain attached. 2 mL of fresh ESF-900 III SFM medium was added gently to each well and the 6-well plate was incubated for 72 h at 27°C. Then, the media was removed from the surface of the cells and centrifuged for 5 min at  $200 \times g$  at room temperature. The supernatant (P0 virus) was stored at +4°C until further use.

#### Baculovirus amplification

The first amplification (P1 virus) was prepared by seeding  $6 \times 10^6$  Sf9 cells in a petri dish (150 mm  $\times$  25 mm). The cells were incubated for 30 min at 27°C to allow the cells to adhere. The media was removed from the surface of the cells and replaced with 2 mL of fresh medium along with 1 mL of P0 virus. The petri dish was then incubated for 1 h. After the incubation, 7 mL of fresh medium was added gently from the walls of the petri dish and the petri dish was incubated for 48 or 72 h at 27°C. After incubation, the media was removed and centrifuged for 5 min at  $500 \times g$  at room temperature. The supernatant (P1 virus) was stored at +4°C until further use.

The second amplification (P2 virus) was prepared by mixing 42.7 mL of Sf9 cells at a concentration of  $1 \times 10^6$  cells/mL cells in a 250 mL Erlenmeyer flask along with 8.3 mL of P1 virus. The flask was incubated for 48 h at 27°C and 85 rpm. After the incubation, the suspension culture was centrifuged for 10 min at  $1200 \times g$  at room temperature. The supernatant (P2 virus) was stored at +4°C until further use.

The third amplification (P3 virus) was prepared exactly as for the second amplification. The P1 or P3 virus preparation incubation times varied between 48 to 72 h depending on the day of the week when the virus amplification was started.

Test expression of GPR68 constructs was performed using the following virus titer: 100  $\mu\text{L}$ / 500 mL, 300  $\mu\text{L}$ / 500 mL, 1 mL/ 500 mL, 3 mL/ 500 mL and 10 mL/ 500 mL in High Five<sup>TM</sup> cells. The experimental procedures involved doing the test expressions for freshly prepared virus in a volume of 20 mL, with the viral quantities being changed accordingly. The test expressions were done in 50 mL falcons at 27°C and 140 rpm for 72 h. After the incubation of 72 h in a shaker, the cells were counted and 1 mL aliquots were taken from all virus titer expressions, which were centrifuged at 16000  $\times$  g for 5 min at room temperature. The cell pellets were resuspended in 350  $\mu\text{L}$  distilled water based on for every  $1 \times 10^6$  cell count.

After resuspension of the pellet, 20  $\mu\text{L}$  of the sample were mixed with 20  $\mu\text{L}$  of 2 $\times$  Laemmli buffer and loaded onto NuPAGE<sup>TM</sup> Novex bis-tris gels in NuPAGE<sup>TM</sup> MES SDS running buffer for 35 min at 200 V. Proteins were transferred to the nitrocellulose membrane using the iBlot2 system for Western blot analysis. As primary antibody the penta-His antibody produced in mouse (Sigma Aldrich) was used, and as a secondary antibody Anti-mouse IgG (whole molecule) – alkaline phosphatase conjugated antibody produced in rabbit (Sigma Aldrich) was used. The expression titer showing the highest expression judged from band intensity was used as optimal virus amount for GPR68 protein expression. This varied typically from construct to construct. However, the final optimal titers were usually in the range of 100  $\mu\text{L}$  - 1 mL/500 mL expression medium. For approximately 30 constructs out of all the GPR68 constructs designed for insect cell expression, the preparation of the amplified baculovirus was performed by Adelheid Loehle, Heike Rapp, Jessi Bretzel and Julian Friedl of the PEP-2 lab, Structural Research Department (Boehringer Ingelheim, Biberach, Germany).

GPR68 protein expression in the insect cells was always performed using P3 virus. The protein expression was performed in a 1.8 L Fernbach flask filled with a volume of 500 mL culture medium per flask at 27°C at 60 rpm for 72 h. The cells were harvested by centrifugation at 2500  $\times$  g for 30 min at 4°C. The cell pellets were stored at -80°C until further use.

### Mammalian cells

HEK293F cells were transiently transfected as suspension cultures for the expression of the target receptor. The cells were split at a density of  $1 \times 10^6$  cells/mL in a volume of 1 L and allowed to grow for 24 h at 37°C, 8% CO<sub>2</sub>, and 120 rpm in 850 cm<sup>3</sup> roller bottles with a maximum capacity of 2 L. The next day, the cells were transiently transfected using the following solutions, where compositions are described for 1 L expression cultures:

Solution 1: 50 mL of Opti-Mem<sup>TM</sup> + 1 mg of GPR68 construct DNA.

Solution 2: 47 mL of Opti-Mem<sup>TM</sup> + 3 mL of 1 mg/mL of linear polyethylenimine (PEI), (molecular weight - 40,000). The stock of PEI was prepared in sterile water.

First, the two solutions were individually mixed and incubate for 5 min each at room temperature. Then, Solution 2 was mixed with Solution 1, and subsequently incubated for a duration of 30 minutes at room temperature. After the incubation, 100 mL of this mixture was added to 1 L of HEK293F cells, followed by addition of 10 mM sodium butyrate to the cells. The expression was done for 72 h at 37°C, 8% CO<sub>2</sub>, and 120 rpm. The cells were harvested by centrifugation at  $900 \times g$  for 30 min at 4°C. The cell pellets were stored at – 80°C until further use.

### **3.2.3 Protein purification**

#### Cell Lysis and Solubilization

The frozen cell pellet from a 500 mL expression was resuspended in 100 mL of hypertonic buffer – 50 mM HEPES pH 7.4, 800 mM NaCl, 10 mM MgCl<sub>2</sub>, 10% glycerol and 1% LMNG (or DDM), 0.2% cholesteryl hemisuccinate (CHS) and EDTA-free cOmplete<sup>TM</sup> protease inhibitor cocktail (1 tablet / 100mL buffer). The solubilization in the hypertonic buffer with detergents was carried out for 1.5 h. The solubilized membranes were then centrifuged at  $50,000 \times g$  for 40 min at 4°C.

### Affinity purification

One of the following methods were opted for the first purification step.

#### 1. Immobilized metal affinity

The supernatant after centrifugation was supplemented with 15 mM buffered imidazole, pH 7.5, and incubated with TALON Superflow Metal Affinity Resin (Clontech) (1.5 ml bed volume per 500 mL of original culture volume), for 2 h or overnight at 4°C. After incubation, the solubilized protein mixture in the presence of the resin was allowed to flow through a glass column for packing or was centrifuged at  $100 \times g$  for 5 min at 4°C and the supernatant was removed using a pipet carefully. Washing was performed with 20 column volumes (CV) of wash buffer – 50 mM HEPES pH 7.4, 500 mM NaCl, 10 mM MgCl<sub>2</sub>, 5% glycerol, 40 mM imidazole and 0.01% LMNG/0.002% CHS or 0.1% DDM/ 0.02% CHS. Elutions were carried with 5 CV in steps of 4, where volume was equal to 1.25 CV and incubation time with the elution buffer was 5-10 min for every step. The elution buffer comprised of 50 mM HEPES pH 7.4, 500 mM NaCl, 250 mM imidazole and 0.003% LMNG/0.0006% CHS or 0.02% DDM/ 0.004% CHS. The eluates were pooled and concentrated using a 50 kDa molecular weight cut-off concentrator (Merck) to approximate volumes of 500-700  $\mu$ L in preparation for the second purification step.

#### 2. Antibody affinity

The supernatant after centrifugation was incubated with FLAG M2 resin (Sigma Aldrich) (2 mL bed volume per 500 mL of original culture volume), for 2 h or overnight at 4°C. After incubation, the solubilized protein mixture in presence of the resin was allowed to flow through a glass column for packing or was centrifuged at  $100 \times g$  for 5 min at 4°C and the flowthrough was removed carefully using a pipet. Washing was performed with 20 CV of wash buffer – 50 mM HEPES pH 7.4, 500 mM NaCl, 10 mM MgCl<sub>2</sub>, 5% glycerol, and 0.01% LMNG/0.002% CHS or 0.1% DDM/ 0.02% CHS. Elutions were carried out for 5 CV in steps of 4, where volume was equal to 1.25 CV and incubation time with the elution buffer was 25-30 min for every step. The elution buffer comprised of 50 mM HEPES pH 7.4, 500 mM NaCl, 0.1 mg/mL 3 $\times$  FLAG peptide (NeoBiotech) and 0.003% LMNG/0.0006% CHS or 0.02% DDM/ 0.004% CHS. The eluates were pooled and

concentrated using a 50 kDa molecular weight cut-off concentrator (Merck) to approximate volumes of 500-700  $\mu$ L in preparation for the second purification step.

### 3. Recombinant protein affinity

The supernatant after centrifugation was incubated with Step-tactin XT 4Flow (IBA Lifesciences) (1.5 mL bed volume per 500 mL of original culture volume), for 2 h at 4°C. After incubation, the solubilized protein mixture in presence of the resin was centrifuged at 100  $\times$  g for 5 min at 4°C and the flowthrough was removed using a pipet carefully. Washing was performed with 10 CV of wash buffer 1 – 50 mM HEPES pH 7.4, 500 mM NaCl, 10 mM ATP, 20 mM MgCl<sub>2</sub>, 5% glycerol, and 0.1% LMNG/0.02% CHS or 0.1% DDM/ 0.02% CHS and EDTA-free cOmplete™ protease inhibitor cocktail (1 tablet / 100mL buffer), followed by washing with 10 CV of wash buffer 2 – 50 mM HEPES pH 7.4, 500 mM NaCl, 5% glycerol, and 0.01% LMNG/0.002% CHS or 0.1% DDM/ 0.02% CHS and EDTA-free cOmplete™ protease inhibitor cocktail (1 tablet / 100 mL buffer). Elutions were carried with 5 CV of elution buffer in steps of 5, where volume was equal to 1 CV and incubation time with the elution buffer was 5 min for every step. The elution buffer comprised of 50 mM HEPES pH 7.4, 150 mM NaCl, freshly added 50 mM biotin (Sigma Aldrich), 50  $\mu$ g/mL Pefabloc SC (Sigma Aldrich), 5  $\mu$ g/mL leupeptin (Roche) and 0.003% LMNG/0.0006% CHS or 0.02% DDM/ 0.004% CHS. The eluates were pooled and concentrated using a 50 kDa molecular weight cut-off concentrator (Merck) to approximate volumes of 500-700  $\mu$ L in preparation for size exclusion chromatography.

#### Size-exclusion chromatography

Size exclusion chromatography (SEC) was used as a second and final step for the protein purification process. It was performed using Superdex 200 Increase 10/300 GL or Superose 6 Increase 10/300 GL. The SEC buffer comprised of – 50 mM HEPES pH 7.4, 150 mM or 500 mM NaCl, 50  $\mu$ g/mL Pefabloc SC (Sigma Aldrich), 5  $\mu$ g/mL leupeptin (Roche) and 0.003% LMNG/0.0006% CHS or 0.02% DDM/ 0.004% CHS. The SEC eluates from the peak were pooled together and concentrated using a 50 kDa molecular weight cut-off concentrator (Merck) or if Anti-BRIL FAB complexation was performed, the 100 kDa molecular weight cut-off concentrator (Merck) was used.

### Anti-BRIL FAB complexation

The Anti-BRIL FAB (FAB) and Anti-BRIL FAB nanobody (NB) were added to GPR68 (ICL3 BRIL constructs) in the ratio of 1:1.2:1.2 or 1:1.5:2. The complexation components were purified by Julian Friedel and Adelheid Loehle according to the protocols previously published: Anti-BRIL FAB <sup>[276, 308]</sup> and Anti-BRIL FAB nanobody <sup>[309]</sup>.

### Purification in presence of ligands

When a ligand was used to enhance the stability of the GPR68 construct, it was introduced into all buffers during the protein purification process at a final concentration of 10  $\mu$ M of small molecules or 100  $\mu$ M of metal ions.

### Miscellaneous

In case a purification was performed at pH 7.8, the buffering component used was HEPES and for purifications performed at pH 6.0 the buffering component used was MES. The ÄKTA Explorer or Pure was equipped with a UV cell size of 2 mm or 10 mm. All the chromatograms reported in this thesis are normalized for 10 mm UV cell. All the purification steps were carried out at 4°C. Stocks were prepared for a maximum storage for up to 3 months. Buffer pH was adjusted at room temperature using NaOH or HCl.

#### Membrane preparation (optional):

The frozen cells pellet from a 500 mL expression was resuspended in 100 mL of hypotonic buffer (20 mM HEPES pH of 7.4, 10 mM MgCl<sub>2</sub>, 20 mM KCl, DNase, and a tablet of EDTA-free cOmplete<sup>TM</sup> protease inhibitor cocktail (1 tablet / 100 mL buffer). Membrane preparation was performed using the cell disruption vessel or the nitrogen bomb (Parr Instrument Company) at 50 bar for 30 min. The cell lysate was then centrifuged at 50,000  $\times$  g for 40 min at 4°C.

#### Alkylation (optional):

Certain protocols involved alkylation of resuspended lysed pellet before solubilization in the dark, where 2 mg/mL iodoacetamide was used to block free cysteines irreversibly. Then the normal steps starting from solubilization were carried out, as described above.

### 3.2.4 Protein analysis

#### Protein concentration measurement

Estimations of concentrations for the purified protein samples or extracted plasmid DNA was obtained using NanoDrop™ One (ThermoFisher Scientific) at 280 nm and 260 nm respectively. The concentrations were calculated based on the Beer-Lamberts law. According to the Beer-Lambert law, the absorbance,  $A$ , is directly proportional to the molar concentration ( $c$ ). The equation  $A = \epsilon \times l \times c$ , represents the relationship between the molar absorption coefficient ( $\epsilon$ ) and the cell path length ( $l$ ). The molar extinction coefficient ( $\epsilon$ ) quantifies the intensity of light absorption by a chemical substance at a specific wavelength. For proteins this is mainly derived by the number of specific amino acids in a protein which act as chromophores like tryptophan, tyrosine, and cysteine [310]. The extinction coefficient of DNA and RNA is influenced by certain functional groups present in the nucleotide bases, such as the amino and keto groups. Purine bases, namely adenine and guanine, exhibit a higher capacity for UV light absorption compared to pyrimidine bases, namely cytosine, thymine, and uracil. This disparity can be attributed to the presence of an additional double bond inside their ring structure [311].

#### SDS-PAGE

Sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) was first described in 1970 when the head of bacteriophage T4 was being investigated [312]. SDS-PAGE is often used to analyze the purity of the protein sample based on separation dependent on the molecular weight of the protein. PAGE is a biochemical method that enables the separation of proteins based on their electrophoretic mobility, which refers to their ability to move rapidly in an electric field. Briefly, SDS is a harsh anionic detergent that denatures the protein making it lose its higher order structure and forcing them to become linear and effectively coating the denatured protein in a net negative charge. Thus, after the exposure to SDS, all the proteins in a sample have a net negative charge and a similar net charge-to-mass ratio. Therefore, in the context of SDS-PAGE, the separation of molecules is based on their mass, rather than their charge [313].

Increased aggregation of membrane proteins has been observed after heating or boiling of samples. Thus, SDS-PAGE samples prepared in this study were not subjected to any prior

heating or boiling before loading them on the gel <sup>[314]</sup>. Also, it has been reported previously that membrane proteins when analyzed via SDS-PAGE show differences between the theoretical molecular weights and observed molecular weights on the SDS-PAGE gel <sup>[315]</sup>.

The samples prepared from pellets, supernatant and flowthroughs from the first step of purification (affinity-based) were mixed with 2× Laemmli sample buffer (Serva Electrophoresis GmbH), whereas any other sample was mixed with 4× Laemmli sample buffer. The sample buffers were additionally mixed with dithiothreitol (DTT) to obtain a final concentration of 50 mM. DTT was added to reduce the disulfide bonds in the protein sample. The protein ladder used for the gels was Precision plus protein™ all blue prestained protein standards #1610373 (BioRad).

The NuPAGE® 4-12% Bis-Tris precast Mini Gels (ThermoFisher Scientific), were used to perform SDS-PAGE of protein samples where the running buffer used was NuPAGE® MES SDS running buffer (ThermoFisher Scientific). The electrophoresis was carried out in a Xcell Surelock™ Mini-Cell chamber (ThermoFisher Scientific). The SDS-PAGE gels were run at 200 V for a duration of 35 min. The gels were stained overnight by Quick Coomassie Stain (Serva Electrophoresis GmbH) for visualization of the proteins, while the background was de-stained using distilled water overnight.

### Western Blot

For Western blotting or immunoblotting, the gels after SDS-PAGE were not stained and subsequently transferred onto a nitrocellulose membrane using the iBlot2 system and the nitrocellulose transfer stacks (Invitrogen). To facilitate the transfer, a voltage of 20 V was applied for a duration of 7 min. Then, Ponceau S (Sigma Aldrich) solution was used to stain the nitrocellulose membrane. The membrane was blocked with a 3% BSA solution in Tris-buffered saline (TBS) with the composition: 50 mM Tris, 150 mM NaCl, pH 7.5 for a duration of 1 h. Subsequently the membrane was washed with TBS-T (50 mM Tris, 150 mM NaCl, pH 7.5, 0.05% Tween 20) three times for 5 min. After the washing, the membrane was incubated with the primary antibody (diluted in TBS – 1:1000) for 1 h. After the incubation, the membrane was washed again with TBS-T three times for 5 min. An alkaline phosphatase conjugated secondary antibody (diluted in TBS – 1:30,000) was incubated with the membrane for 40 min. The last cycle of washing was performed after

the incubation with TBS-T three times for 5 min. The protein bands were stained using the 1-Step™ NBT/BCIP solution (ThermoFisher Scientific), an alkaline phosphatase substrate. The reaction was stopped to avoid high background by washing the membrane with distilled water.

### Mass Spectrometry

In-gel protein samples were analyzed by cutting out gel pieces of the protein bands of interest from Coomassie stained gels and were handed over to Karen Köhler (Boehringer Ingelheim, Biberach, Germany) for analysis. An endopeptidase, trypsin, was used for protein digestion to yield peptides. Matrix-assisted laser desorption ionization – time of flight (MALDI-TOF) mass spectrometry was performed with digested samples and the results were analyzed using Mascot search engine software (Matrix Science).

### Differential Scanning Fluorimetry (DSF)

DSF also known as thermal shift assay is a method which utilizes a fluorescent dye which binds to the hydrophobic regions as the protein is subjected to a temperature gradient. The relationship between the fluorescence signal and temperature is graphed, yielding a sigmoidal curve. The point of inflection in this curve corresponds to the melting temperature ( $T_m$ ), at which 50% of the protein is unfolded. Typically used fluorescent dyes that bind to hydrophobic regions do not work well with detergent solubilized membrane proteins. Thus, the thiol-specific dye like 7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM) is a commonly used dye. However, more recent label-free techniques like NanoDSF are also effective in determining the melting temperatures of GPCRs.

### NanoDSF

NanoDSF is a highly sensitive and label-free technique that monitors the changes in intrinsic fluorescence originating from the inherent tryptophan in proteins, as a function of temperature. The fluorescence of tryptophan was quantified at two specific wavelengths, 330 nm, 350 nm, and the ratio of 350 nm/330 nm<sup>[316]</sup>. A common finding pertaining to the protein unfolding phenomenon is the occurrence of a red shift or a blue shift in the fluorescence emission maxima, which is towards a longer emission wavelength or a short

emission wavelength depending on the local environment changes of the tryptophans as the protein unfolds [317, 318].

The melting points in this study were determined with the Prometheus NT.48 (NanoTemper). The purified protein sample (10  $\mu$ L) was loaded in a high sensitivity glass capillary (NanoTemper). The discovery scan function was used to estimate the optimal excitation level (usually in the range of 50-100%). The melting curve was generated over a temperature range of 20-95°C, in incremental steps of 1°C per minute. Data analysis was done using the PR.ThermControl V2.1.1 software. The experiments were performed with a minimum of two replicates.

### 3.2.5 X-ray crystallography

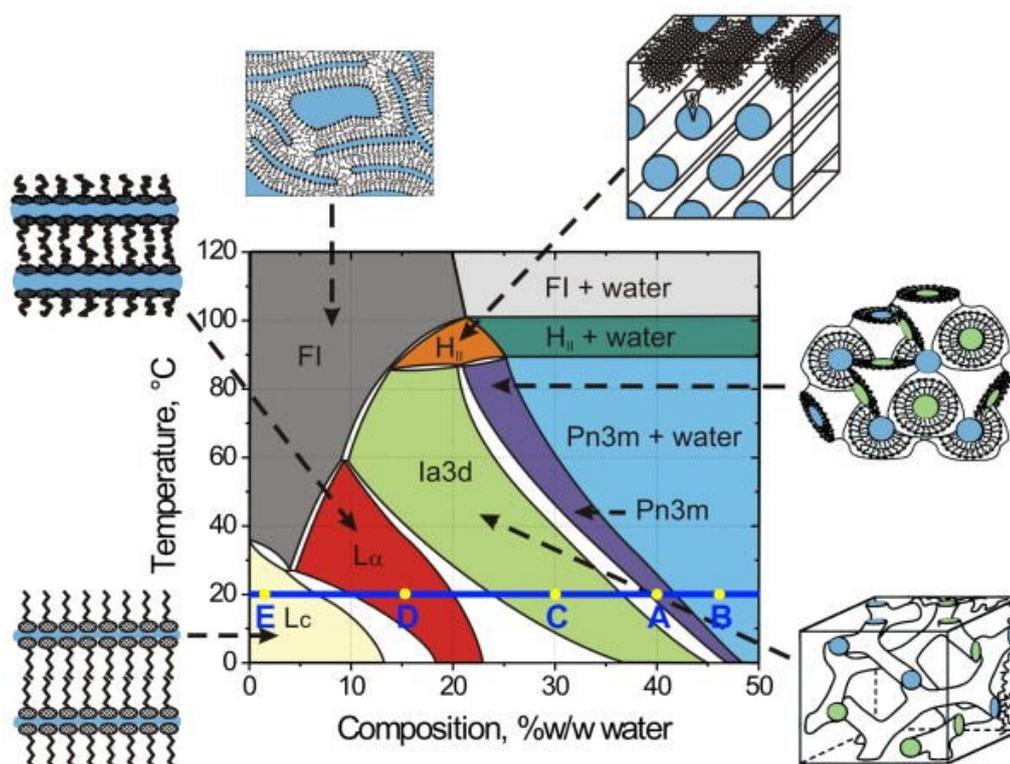


Figure 14: Phase diagram of the monoolein/water system determined at 20 °C. Ia3d and Pn3m represent cubic phases. The figure has been reused [319].

The phenomenon of X-ray diffraction by crystals was initially observed by Max Laue and colleagues in 1912 [320, 321]. The first protein structure elucidated using X-ray crystallography was myoglobin [263]. The process of a protein crystal structure elucidation involves crystallizing a concentrated protein sample and subjecting the resulting crystals to an X-ray beam. A few methods to obtain protein crystals are vapor drop diffusion method

(sitting drop or hanging drop) <sup>[322]</sup>, counter-diffusion method <sup>[323]</sup>, dialysis method <sup>[324]</sup>, batch method <sup>[325, 326]</sup> and lipidic cubic phase crystallization <sup>[267, 319]</sup>.

The obtained diffraction patterns are analyzed to determine the crystal packing symmetry and the dimensions of the repeating unit that constitutes the crystal. This information is essentially derived from the arrangement of the diffraction spots <sup>[327, 328]</sup>. There are 230 space groups categorized within 7 lattice types, however not all of them are suitable for proteins due to the chirality of amino acids. Thus, only 65 space groups are applicable for proteins <sup>[329]</sup>. The magnitude of each diffracted spot is determined by the diffracted waves that reach the detector. The intensity of the waves will be dictated by the amplitude and phase difference, which is given as an angle. The amplitude estimation is possible from the intensity as they are proportional, but the phase information cannot be determined directly. Thus, for protein structure determination the phase problem has been solved via different methods like molecular replacement <sup>[330]</sup>, isomorphous replacement <sup>[331]</sup>, multiple wavelength anomalous diffraction <sup>[332]</sup>, and single wavelength anomalous diffraction <sup>[333]</sup>. Now, structure factors can be calculated using the Fast Fourier Transform method <sup>[334]</sup> resulting in a three-dimensional electron density map, which is then used for refinement and model building to finally obtain the coordinates of the protein structure <sup>[327]</sup>.

The most widely used methods to crystallize membrane proteins are *in surf*o (in detergent) crystallization and lipidic cubic phase (LCP) crystallization or *in meso* crystallization. LCP crystallization occurs in a bicontinuous lipidic mesophase. These mesophases are very sensitive to changes in the environment (e.g., salt, pH, temperature, other additives) and composition (e.g., type of lipid used and the ratio of lipid to water content). The cubic phase is essential formed for the monoolein/water system at room temperature (~20°C) with an overall composition of approximately 40% (wt/wt) water (Figure 14). Excess amount of water would result in formation of a second phase and a two-phase system <sup>[319]</sup>. This behavior of formation of a good cubic phase is for monoolein/water system is better represented by the phase diagram (Figure 14). So far, LCP crystallization has been very successful in structure determination of GPCRs <sup>[335]</sup>.

### LCP setup

In my thesis, freshly purified GPR68 was used to perform LCP crystallization. The process of lipidic cubic phase (LCP) crystallization involved combining the protein with a lipid mixture consisting of monoolein and cholesterol in a ratio of 9:1 (w/w). The protein to lipid ratio used was 2:3.4. The LCP setup using the mosquito LCP (SPT Labtech) with a solution ratio of 40:800 nL (bolus: precipitant). The crystals were collected using MicroLoops (mitogen) and rapidly frozen using liquid nitrogen. The crystals were sent for data collection at PXI beamline of the Swiss Light Source (SLS).

Please note that the harvesting of crystals was performed by Dr. Dietmar Weichert (Boehringer Ingelheim, Biberach, Germany) and the data collection was setup by Dr Dirk Reinert (Boehringer Ingelheim, Biberach, Germany).

### **3.2.6 Single particle cryogenic electron microscopy**

Single particle cryogenic electron microscopy utilizes electrons as radiation, in a high vacuum environment to examine samples. These electrons are then accelerated along the microscope column typically using voltages ranging from 100 to 300 kV. Scattered electrons, after passing through the object, are concentrated by the electromagnetic lenses of the microscope. A higher resolution can be achieved with shorter wavelengths <sup>[336]</sup>. Vitrified protein samples are obtained after being plunge frozen into liquid ethane which are later stored in liquid nitrogen and then used for imaging. The thickness of the ice layer obtained can be quite variable. Preferably, one would observe identical copies of the protein in different orientations within the vitreous ice (Figure 15). Many select 2D images with various orientations can be combined to obtain a 3D reconstruction of the structure <sup>[336]</sup>. Unlike X-ray crystallography, this method works in the real space thus there are problems associated to the phase problem. Cryogenic temperatures help reduce radiation damage <sup>[337]</sup>, that preferably includes exact replicas of the intricate structure in various orientations.

### Sample preparation

GPR68 constructs with ICL3 BRIL construct (e.g., C99) was complexed to Anti-BRIL-FAB and Anti-FAB nanobody in the ratio of 1:1.5:2. These samples were either purified via Superose 6 Increase 10/300 GL column (Cytiva). The samples were always freshly

prepared and immediately subjected to plunge freezing. A volume of 3  $\mu\text{L}$  of protein sample was applied onto glow discharged QF-1.2/1.3-300-Au and QF-0.6/1.0-300-Cu grids. The grids were then rapidly frozen in liquid ethane using a plunge freezer, Vitrobot Mark IV (ThermoFisher). Then the grids were transferred to grid boxes and stored in liquid nitrogen.

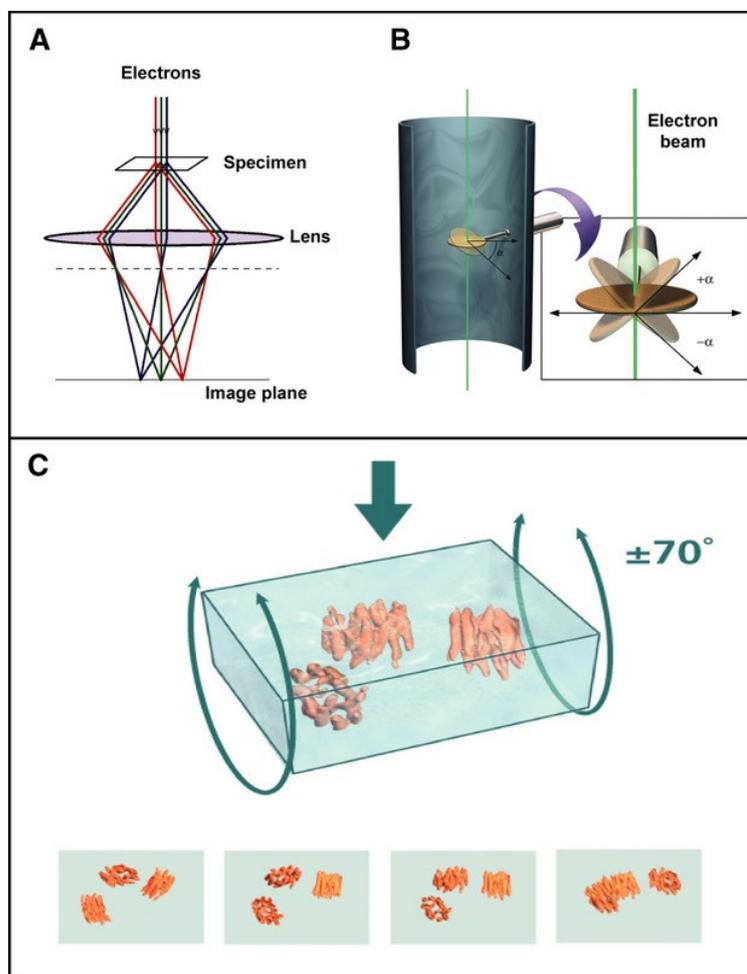


Figure 15: Working of single particle electron microscopy. (A) The figure depicts the process of image formation in an electron microscope. (B) Schematic depicting the fundamental process of data gathering for electron tomography. When the specimen is inclined in relation to the electron beam, several images are captured of the identical area. (C) Selected projection views of the protein sample in vitrified film which were used to rebuild the three-dimensional volume by using a collection of projection images that can be used to create back-projection profiles. The figure has been reused with permission from <sup>[336]</sup>.

### Data Acquisition and Processing of EM Datasets

The micrographs were automatically captured using the EPU software on a Titan Krios microscope (Thermo Fisher) operating at 300 kV. The microscope was equipped with a K2 direct electron detector in electron counting mode. The images were taken at a nominal magnification of 130,000 $\times$ , which corresponds to a calibrated pixel size of 1.077  $\text{\AA}$ . A series of 8-second videos consisting of 40 frames each were captured at an electron dose of

55 e/Å<sup>2</sup>, and the defocus values ranged from 0.5 μm to 3.0 μm. The data acquisition was performed by Dr. Rebecca Ebenhoch.

A total of 18,591 micrographs were collected and were motion corrected by patch motion correction and the CTF parameters were estimated by patch CTF estimation. Duplicate particles were eliminated, and several iterations of 2D classification were performed to generate 2D classes that clearly displayed the fiducial marker protruding. All the data processing steps were performed in CryoSPARC v4.4.1 [338].

### 3.2.7 Molecular dynamics simulations

Please note that the method section for the molecular dynamics simulations in this thesis has been originally described in a publication [339].

#### *Structural modeling of GPR68*

The structural models for GPR68 were generated using two distinct structural modeling methodologies. The initial approach entailed utilizing the webserver GPCR Online Modeling and Docking (GOMoDo) [286], along with its recently launched iteration pyGOMoDo [212]. We used these tools to choose templates for GPR68 homology modeling, perform multiple sequence alignments (MSAs) using the search-and-align technique HHblits (profile hidden Markov Models (HMM)), and do lightning-fast iterative sequence search based on HMM-HMM). A further comparison was made using GPCRdb [6].

The N-terminus and the disulfide bonds were refined using MODELLER v10.1 [340]. Class A GPCRs typically have a conserved disulfide bridge between C3.25 and C45.50 [341]. In the case of GPR68, an additional disulfide bond might connect C13<sup>1.24</sup> to C258<sup>ECL3</sup> [171]. Two separate MODELLER calculations were performed, one calculation with a restraint imposed on C94<sup>3.25</sup>-C172<sup>45.50</sup> bridge and, another with restraints imposed on C94<sup>3.25</sup>-C172<sup>45.50</sup> and C13<sup>1.24</sup>-C258<sup>ECL3</sup> disulfide bridges. For each of the described set, 1000 GPR68 models were generated, for a total of 6000 models, which were then ranked according to their calculated DOPE score (Shen and Sali, 2006). After visual inspection of the ten highest-scoring models, models with unstructured regions or other structure artifacts were discarded, and the structure that had both the highest DOPE score and lacked structure artifacts was kept. The N-terminal segment (first 18 amino acids) of all structural templates

were modeled and refined with MODELLER. Thus, we kept in total six models from the first strategy that includes the usage of GOMoDo and MODELLER. The structural models based on the first strategy were prepared by Filippo Baldessari and Prof. Alejandro Giorgetti (University of Verona).

The utilization of the more recent artificial intelligence (AI) based structural modeling software AlphaFold <sup>[213]</sup> constituted the second approach. The GPR68 model was obtained from alphafold.ebi.ac.uk and the AlphaFold script was utilized to determine the underlying templates used by AlphaFold for modeling. The AlphaFold algorithm uses a confidence metric known as the predicted local distance difference test (pLDDT) on a per-residue basis. The AlphaFold GPR68 model had an overall pLDDT score of 79.69, which falls within the range of scores between 70 and 90, indicating a high level of confidence in the model's overall quality. The confidence scores of termini and loop regions were lower as these regions are generally flexible. Additionally, I prepared a bovine rhodopsin-based model of GPR68 for comparison with other models generated in the study and with Ludwig and colleagues <sup>[112]</sup>. This model was not used for simulations. The in-detail protocol for model preparation can be found here <sup>[339]</sup>.

Model#	Proteins used as template (s)	Sequence Identity (%)	Disulfide bridges	Sequence range
<i>Structural models prepared with GOMoDo</i>				
1	Protease Activated Receptor 2 (5NJ6 - Inactive)	17	C94-C172	1-299
2			C94-C172, C13-C258	
3	Angiotensin II Type 1 Receptor (4ZUD - Inactive)	20	C94-C172	12-299
4			C94-C172, C13-C258	1-299
5	Angiotensin II Type 1 Receptor (6DO1- Active)	20	C94-C172	12-290
6			C94-C172, C13-C258	1-290
<i>AlphaFold AI-based model</i>				
7	Multiple templates	-	C94-C172, C13-C258	1-317
8		-	C94-C172, C13-C258	1-317 C304, C310 - palmitoylated

Table 1: Structural models of GPR68 used in this study.

A longer amphipathic helix 8 was observed in the AlphaFold model GPR68. The palmitoylation of the C-terminus of GPR68 may occur at two specific locations, namely 304 and 310, where two Cys groups are present. Thus, to assess the possibility of palmitoylation, we used two AlphaFold models, where one model was palmitoylated at the two sites – C304 and C310 and the other model was not palmitoylated.

### Simulation Setup

Using the website Orientations of Proteins in Membranes (OPM) <sup>[342]</sup> each GPR68 model was oriented in the membrane. The GPR68 protein was incorporated into hydrated bilayers of POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) lipids using the CHARMM-GUI <sup>[343]</sup>, with the addition of a neutralizing 0.15 M sodium chloride salt. In a distinct simulation, we examined the behavior of GPR68 when incorporated into a POPC membrane containing 10% cholesterol (Sim #9 in Table 2). This cholesterol content falls within the range typically observed in plasma membranes <sup>[344]</sup>. Although, the Sim #9 has not been described in my thesis but has been discussed in my previously published work <sup>[339]</sup>. The force-field parameters of CHARMM36m <sup>[345]</sup> were used in conjunction with the TIP3P model for water molecules <sup>[346]</sup>, and NAMD v2.14 <sup>[305, 347, 348]</sup>. The standard CHARMM-GUI protocol was utilized to conduct the initial equilibration and geometry optimization of all systems, along with velocity rescaling and constraints. Production runs were performed in the NPT ensemble (constant number of particles N, constant pressure P = 1 bar, and temperature T = 303.15 K) using a Langevin dynamics scheme with a collision frequency of 5 ps<sup>-1</sup> and a Nosé-Hoover Langevin piston <sup>[349–351]</sup>. Coulomb interactions were calculated using the smooth-particle mesh Ewald method and for real-space interactions we used a switching function between 10 and 12 Å. Lengths of bonds involving hydrogen atoms were fixed. Heating, equilibration and the first 1 ns of the production run were performed with an integration step of 1 fs, and the remaining of each of the productions run, with a step of 2 fs.

For all simulations performed to identify the best compatible structural model of GPR68 (Simulations #1 to #8 as described in Table 2) standard protonation states were used for all titratable sidechains, i.e., positively charged arginine and lysine, and negatively charged aspartate and glutamate, and histidine singly protonated on Nε2. We used the best-ranked GPR68 model (Model #5) (according to the compatibility of MD-equilibrated models with

known structural features of class A GPCRs, see ranking protocol below) to understand how protonation states impact conformational dynamics of GPR68. I performed independent simulations (Sim #10 - Sim #18 in Table 2) with different protonation states of selected Asp, Glu, and His sidechains (Table 2).

Each simulation was prolonged to 0.6-1  $\mu$ s, and coordinates were saved each 10 ps. We computed H-bond graphs and average values from the last 600 ns of Simulation #5, and the last 300 ns of Simulations #9-18 (Table 2).

Protocol to rank structural models of GPR68

As summarized above, six GOMoDo-generated and two AlphaFold-generated structural models were selected for MD simulations (Table 1). To understand the relevance of the equilibrated structures obtained from the eight independent MD simulations performed, we first monitored, for each simulation, percentage of alpha-helical content (Figure S1), the structural stability of the TM helical region according to root-mean-squared distance (rmsd) values relative to the starting coordinates (Figure S2). Separately, each of the eight equilibrated structural models were closely inspected for its compatibility with generally accepted class A GPCRs.

Sim#	Model#	Protonation	Length ( $\mu$ s)	Lipid Bilayer Composition
1	1	Standard	1	POPC
2	2		1	
3	3		1	
4	4		1	
5	5		1	
6	6		1	
7	7		1	
8	8		1	
9	5		0.6	POPC:cholesterol
10	5	E149-O $\epsilon$ 1	0.6	POPC
11		E149-O $\epsilon$ 2	0.6	
12		H269	0.6	

13		E103-O $\epsilon$ 1	0.6	
14		E103-O $\epsilon$ 2	0.6	
15		D282-O $\delta$ 1	0.6	
16		D67-O $\delta$ 2, E149-O $\epsilon$ 1, D282-O $\delta$ 2	0.6	
17		D67-O $\delta$ 2, E149-O $\epsilon$ 2, D282-O $\delta$ 2	0.6	
18		D67-O $\delta$ 2, E103-O $\epsilon$ 1, E149-O $\epsilon$ 1, D282-O $\delta$ 2	0.6	

Table 2: Summary of all MD simulations performed in this thesis. The designation 'Standard' denotes that all titratable sidechains possess standard protonation, as previously described. Simulations 10-18 involved the inclusion of titratable sidechains that exhibited non-standard protonation. Specifically, the protonated oxygen atom of the neutral Asp/Glu in the simulations 10, 11, 13-18, while simulation 12 had a positively charged His sidechain (H269). Lipid bilayers were comprised of POPC, except for Sim #9, where GPR68 with standard protonation was embedded within a POPC membrane supplemented with 10% cholesterol.

### Graph computations of dynamic H-bond networks of GPR68

Dynamic H-Bond Networks and water wires were calculated using Bridge2 [300, 301]. Briefly, Bridge2 is a set of graph-based algorithms with graphical user interface for highly efficient computations of protein-water H-bond networks. A graph of H-bonds consists of nodes, which in our case are the H-bonding sidechains of GPR68, and edges, which are direct or water-mediated H-bonds between sidechains (on average, may contain up to 3 water molecules in one bridge). H-bonds are computed using geometric criteria; here, H-bond computations for the protein-water H-bond were performed with a distance criterion of  $\leq 3.5$  Å between the H-bond donor and acceptor hetero-atoms, and an H-bond angle of  $\leq 60^\circ$ . An H-bond cluster consists of nodes and edges that are all inter-connected to each other. An H-bond path between two nodes  $i$  and  $j$  of the graph consists of the nodes and edges that interconnect  $i$  and  $j$ . The joint occupancy (JO) of an H-bond path reports the percentage of coordinate sets from the trajectory segment used for analyses in which all intermediate path segments are sampled simultaneously. For clarity, unless otherwise specified, we show H-bond graphs with H-bond occupancies of at least 30%.

### Time series and average number of internal water molecules

The number of internal water molecules in structural models of GPR68 simulations (Figure S3) were monitored by the membrane region according to the average location of the phosphate atoms of the lipid bilayer within 10 Å of the protein and monitored the number of water oxygen atoms within  $\pm 5$  Å of these membrane boundaries. The calculations were performed based on the scripts from a previous study on opioid receptor [163].

#### Inter-helical distances in GPR68 models and static structures of GPCRs

A geometrical evaluation was introduced in Dalton et al., 2015 to compare inactive vs. active structures of GPCRs. This geometrical evaluation included a set of inter-helical distances measured between the C $\alpha$  atoms of G35<sup>1.46</sup>, D67<sup>2.50</sup>, V109<sup>3.40</sup>, W146<sup>4.50</sup>, P197<sup>5.50</sup>, F237<sup>6.44</sup> and C279<sup>7.46</sup>; They used a dataset of 25 static GPCR structures, of which 7 were presented for active conformations, and 18, for inactive conformations [352]. We monitored the averages of the time series of the distances along simulations #1-#8 (Figure S4) and compared it to the Dalton and colleague dataset, along with Hauser and colleague dataset [104]. Molecular graphics were prepared with Visual Molecular Dynamics, VMD [307].

### **3.2.8 Functional assay**

#### Principle

I used the previously described homogenous time-resolved fluorescence (HTRF) IP-One assay to track GPCR activation via G $\alpha_{q/11}$  pathway [302]. This assay is typically performed at physiological pH to gain insights into small molecule ligands in a high throughput manner [353], therefore the assay had to be modified so that it is suitable to be tested for GPR68, where the receptor is activated by acidic pH. Upon exposure to an increased concentration of agonist (in my case, protons) the receptor (in my case, GPR68) would couple to G $\alpha_{q/11}$ , which would then activate the enzyme phospholipase C to produce the secondary messengers IP<sub>3</sub> and diacyl glycerol (DAG). However, IP<sub>3</sub> is not stable and gets converted to inositol diphosphate (IP<sub>2</sub>), inositol monophosphate (IP<sub>1</sub>) and lastly to d-myo-inositol due to the presence of inositol phosphatases that degrades IP<sub>3</sub> into the previously described products [354]. It is possible to block the activity of inositol monophosphates by addition of LiCl to prevent further degradation of IP<sub>1</sub> to d-myo-inositol [355]. This leads to the accumulation of IP<sub>1</sub> in the cells upon receptor activation in presence of LiCl.

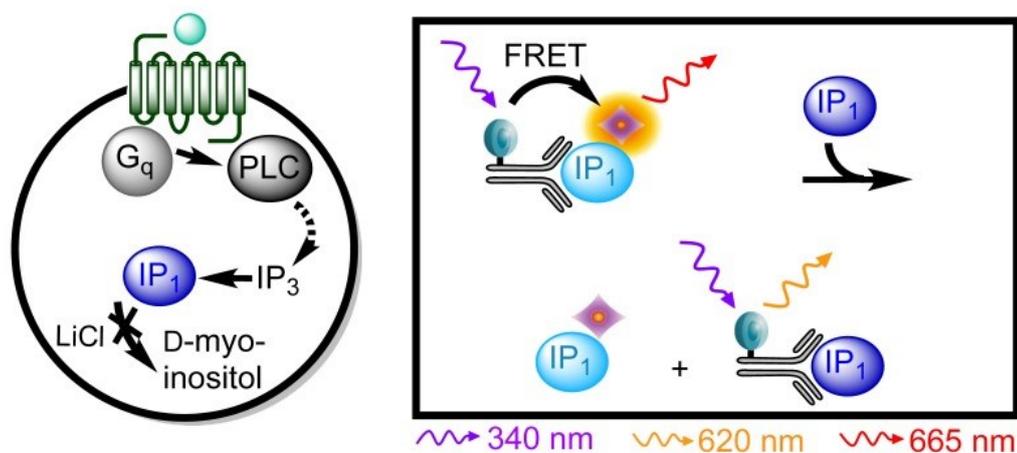


Figure 16: HTRF IP-One assay principle. The snake diagram in the circle represents a GPCR and the cyan sphere represents the cognate agonist for the receptor. The abbreviations used in the figure are – PLC (phospholipase C), IP<sub>3</sub> (inositol, triphosphate), IP<sub>1</sub> (inositol monophosphate), LiCl (lithium chloride), and FRET (Förster resonance energy transfer). The figure has been reused <sup>[353]</sup> under Creative Commons Attribution license.

The HTRF assay is competitive binding assay where after incubation with the agonist, as a final step known amounts of d2 (acceptor) labelled IP<sub>1</sub> and Terbium cryptate (donor) labeled IP<sub>1</sub> specific antibody are added. These components are prepared in lysis buffer to lyse the cells and stop further accumulation of IP<sub>1</sub> produced due to receptor activation. This endogenously produced IP<sub>1</sub> competes with the labelled IP<sub>1</sub> to stop the FRET reaction (Figure 16). The excitation wavelength is 337 nm and the emission wavelength for the donor is 620 nm and for the acceptor it is 665 nm. The signal is measured as the ratio of acceptor to donor emission. This signal is further multiplied by a factor of 10,000 which is then referred to as the HTRF ratio and forms the result <sup>[302, 353]</sup>. As the amount of endogenous IP<sub>1</sub> is increased due to receptor activation the HTRF ratio decreases.

The time-resolved nature of the HTRF assay mitigates the background fluorescence that arises from the buffer, media components etc. The time delay introduced between the excitation of the donor molecule and the measurement of the emission of the acceptor molecule is referred to as the time-resolution. This would require a fluorophore that emits for a longer duration e.g., millisecond timescale. Therefore, Terbium cryptate is used as a donor as it can emit for a longer duration. The photobleaching would occur in the timescale

of a few minutes and therefore it does not affect the HTRF measurement as it is done within half a millisecond (Figure 17) [302, 353].

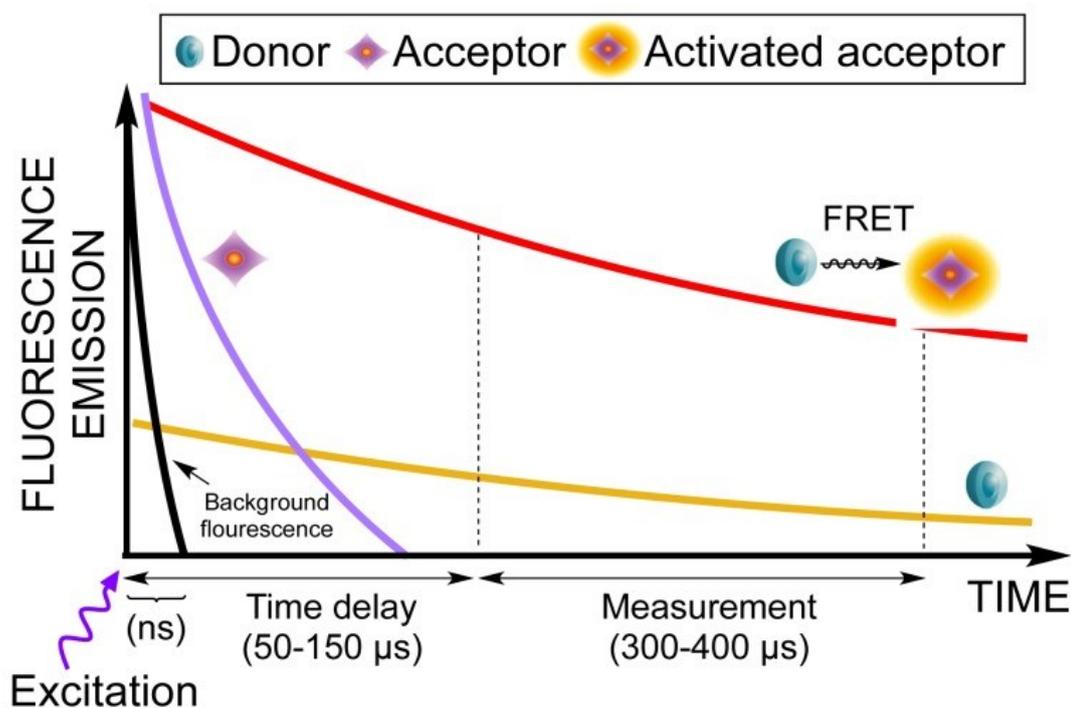


Figure 17: Principle for time-resolved fluorescence. The figure has been reused [353] under Creative Commons Attribution license.

### Cell culture

HEK293H (Gibco), human embryonic kidney 293 cells were cultured in Dulbecco's modified eagle medium (DMEM) growth media supplemented with GlutaMax™ (ThermoFisher Scientific), 10% heat inactivated fetal bovine serum (Gibco), and 100 U/mL of penicillin-streptomycin (Gibco). The HEK293H cells were used as they offer faster growth, high transfection efficiency, and better adherence. The cells culture maintenance was done biweekly, and the cells were grown adherently in high humidity at 37°C and 5% CO<sub>2</sub>.

### Transfection and preparation for cellular assay

The HEK293H cells were seeded in a 6-well plate (Corning, Falcon) with a cell density of  $0.5 \times 10^6$  cells/well (2 mL volume/well) overnight to achieve 80-90% confluency. Transient transfections were performed for cells in each well by addition of an appropriate GPR68

construct-DNA (2500 ng DNA) with 6  $\mu$ L of Lipofectamine reagent (Invitrogen) mixed in Opti-MEM™ I reduced serum medium (Gibco). After 24 h of transfection, cells were detached with TrypLE™ express enzyme (Gibco), counted and resuspended in fresh growth medium. Subsequently, the cells were plated into Poly-D-Lysine 384-well solid white opaque flat-bottom plates (Corning) at a cell density of 15,000 cells/well in a volume of 50  $\mu$ L/well. The cells were then allowed to adhere and recover for 20-24 h before further experimentation.

### Buffers and pH

Experiments were carried out using Hank's balanced salt solution (HBSS) buffer without calcium or magnesium salts and without phenol red. HBSS was further modified by addition of 8mM each of MES (2-(N-morpholino) ethanesulfonic acid), HEPES 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid and Bicine for ideal buffering capacity to cover the entire pH range used during the cellular assay. The pH adjusted buffers were prepared at room temperature ( $\sim$ 20-22°C) and the pH was adjusted using KOH or HCl. The modified HBSS buffer used for washing the cells was adjusted to pH 7.4 and is henceforth termed as 'wash buffer'. The cellular assay involved screening of the receptor behavior at eight different pH points: 5.2, 5.8, 6.4, 6.8, 7.2, 7.6, 8.0 and 8.4. On the assay day, 50mM lithium chloride (LiCl) and 0.1% bovine serum albumin (BSA) were added freshly to these eight different buffers. Lithium chloride inhibits the inositol-monophosphatases which prevents degradation of inositol-1-phosphate (IP<sub>1</sub>) to myo-inositol.

### IP<sub>1</sub> formation assay

The IP-One Homogenous Time Resolved Fluorescence (HTRF) assay kit (Revvity) was used to assess GPR68 coupling to G $\alpha_{q/11}$  indirectly by tracking the accumulation of the secondary messenger, IP<sub>1</sub>. The cellular assay involved washing the cells three times with 80  $\mu$ L of the wash buffer using an automated washer (Bio-Tek ELx405 HT). The automated washer was programmed to remove all the buffer from each well after the washing was completed. Subsequently, a multidrop dispenser was used to dispense 20  $\mu$ L of different pH buffers in the 384-well plate and the plate was incubated at 37°C for 30 min. After the incubation was completed, the detection reagents (5  $\mu$ L each of the IP<sub>1</sub>-d2 and Anti-IP<sub>1</sub>-

cryptate prepared in lysis buffer from the kit) were added to the each well and the plate was then incubated at room temperature in the dark for 60 min before final measurement.

### Data acquisition

The final measurement was carried out using PHERAstar FSX (BMG Labtech) with the two optical modules set to excitation and emission filters of 337 nm/665 nm and 337 nm/620 nm. The data was collected using 5 flashes/well with an integration time of 400  $\mu$ s and an integration delay of 60  $\mu$ s at a focal height of 11.8 mm. The final raw data illustrates the ratio of the emission wavelengths 665 nm/620 nm (acceptor/donor) with a multiplication factor of 10,000 which collectively is called as HTRF ratio.

### Data analysis

GraphPad Prism 10 was used for data analysis. The pH titration data from wild-type GPR68 (WT) and mutants acquired in terms of the HTRF ratio was interpolated or normalized against the standard curves obtained for the eight different pH points at which the assay was performed to get the accumulated IP<sub>1</sub> values. This interpolated data was baseline corrected by the interpolated data obtained from non-transfected cells. Afterwards, a non-linear fit was performed followed by normalization against wild type (WT) efficacy maximum averaged over the three independent datasets. The negative logarithm of half-maximum effective proton concentration has been reported as pH<sub>50</sub> values and the efficacy maximum is denoted as E<sub>max</sub>. The normalized data points are represented as mean  $\pm$  standard error mean for three independent experiments performed in duplicates, where every data point in an independent experiment consisted of duplicates.

I have used the following frequently used terms to describe the results of my functional assay. “The basal activity of the receptor, and its ability to activate intracellular signaling pathways, is defined by the probability that a fraction of the receptor adopts the active state in the absence of ligand” [356]. Efficacy maximum is defined as the maximal effect of the agonist (in this case, protons) or maximum signaling response [24]. It has been denoted as E<sub>max</sub>. Generally, potency is defined as the ligand concentration required to produce a given effect [357]. In our case as the agonist are protons, pH<sub>50</sub> can be described as the negative logarithm of the half-maximum effective concentration [134].

## 4. Results

### 4.1 Expression, Purification and Structure Determination Trials of GPR68

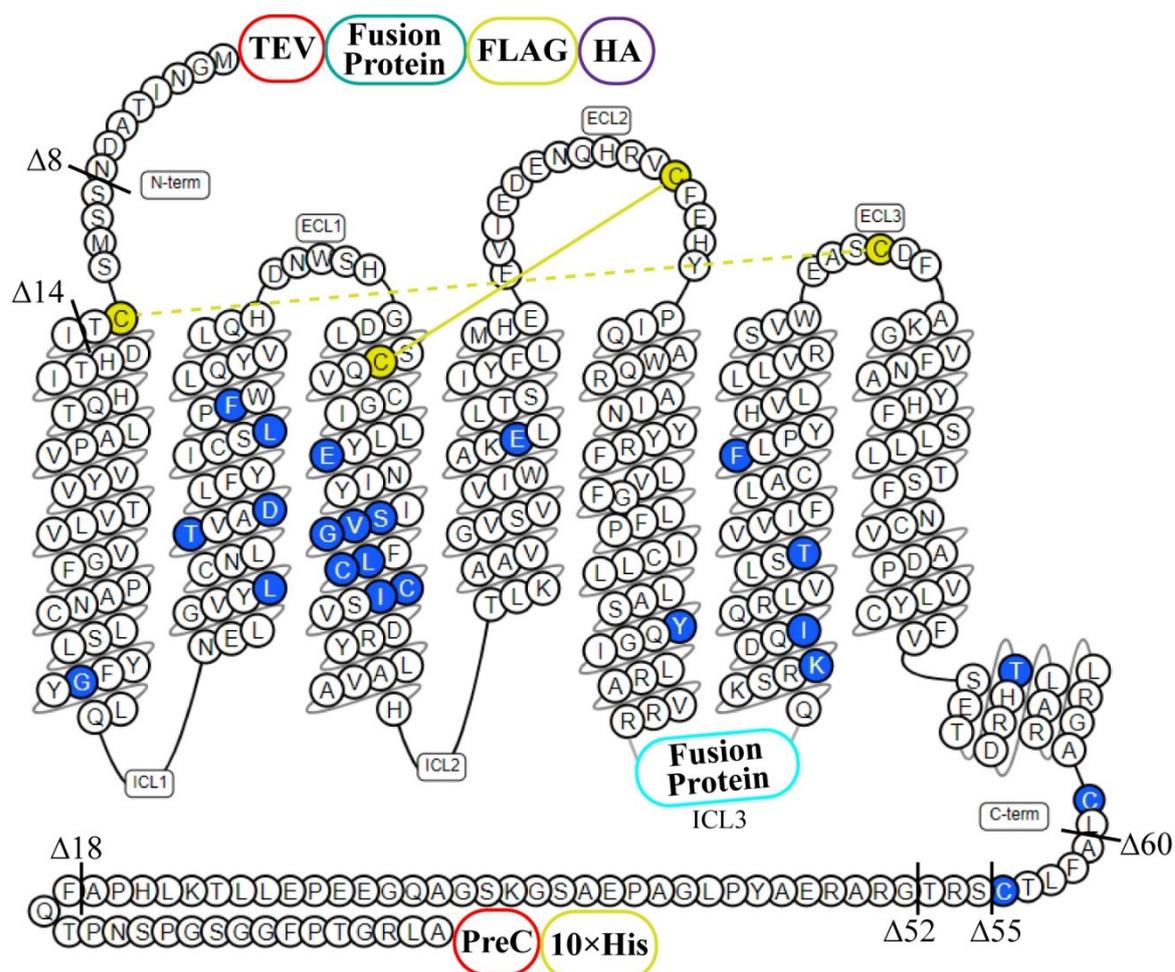


Figure 18: A generic representative construct design strategy for human GPR68 used in this thesis. The conserved class A GPCR disulfide bond has been shown in yellow solid line and the putative disulfide bond has been shown in dotted yellow line. Additionally, the cysteines involved in disulfide bond formation are highlighted in yellow. The commonly used truncations throughout the thesis at *N*-terminus or *C*-terminus are marked with a black solid line along with deletion of amino acids represented with Greek letter ‘ $\Delta$ ’ followed by number of amino acids. The mutated amino acids described in this section have been highlighted in blue. Certain representative modifications in the constructs were insertion of fusion proteins (either at *N*-terminus or ICL3), tobacco etch virus (TEV) protease cleavage site or PreScission (PreC) protease cleavage site, purification tags like deca-Histidine tag (10 $\times$  His) or FLAG tag, and haemagglutinin signal peptide (HA).

The primary objective was to express, purify and optimize different GPR68 constructs to identify a construct suitable for inactive state structure determination via X-ray crystallography or single particle cryogenic electron microscopy in presence of small molecule ligands (antagonists). There was no previous GPR68 construct design or purification protocol reported in the literature when I started my thesis. This chapter of my thesis focuses on the different strategies undertaken for the optimization of the GPR68 constructs for structure determination. Briefly, the strategies undertaken encompass usage of protease cleavage sites (TEV or PreScission), purification of tags at both terminus (10× His, FLAG, or Twin-Strep tag), hemagglutinin signal peptide (HA) for targeted expression of GPR68 to the plasma membrane, *N* and *C* terminal truncations, mutations, screening and insertion of various fusion proteins in ICLs or *N*-terminus, and varying insertion regions of fusion proteins. A representative construct design has been shown in Figure 18 and specific construct details segregated based on strategy used has been summarized in Table 3. For a more elaborate description of these strategies and their effectiveness in different structure determination techniques, please refer to the sections 1.4.1 and 1.4.2. The GPR68 constructs details, the fusion acronyms or any other descriptive legends can be found in the Appendix 1. These acronyms or legends are used throughout this section. Furthermore, the absence of a specified species for the GPR68 construct implies that the receptor being referred to is of human origin.

GPR68 purified in presence of detergents on the size exclusion chromatography (SEC) columns will run at a higher molecular weight as it is surrounded by a detergent micelle and the additional increase in molecular weight is due the molecular weight of the detergent micelle itself. The molecular weight of the detergents mainly used in my thesis are – DDM (~70 kDa) [358], LMNG (variable MW, as it has been suggested that membrane proteins are not truly embedded in detergent micelle and are rather surrounded by minimal amount of detergent required to shield hydrophobic areas) [359], and GDN (~80-85 kDa, as estimated based on the fact that the hydrodynamic radius of the GDN micelle is 15-20 % more than DDM [360, 361]). Note that the micelle size can vary based on environmental factors.

The retention volume for monomeric GPR68 constructs irrespective of detergent was observed to be at an average retention volume or time of ~13.2-13.5 mL (Superdex 200 increase 10/300 GL column), or ~16.8-17.1 mL (Superose 6 increase 10/300 GL column) or ~13.8-14.1 min (Superose 6 increase 5/150 GL column) or ~18.2 min (UNIX-C SEC-

300 4.6/300 column). Slight deviations may be observed from these values due to usage of different columns throughout my thesis. The retention time/ retention volume of the GPR68 construct monomer/ complex peak was estimated by comparing with the column calibration standard run for a specific column. The criteria used to analyze the monomer peak from chromatograms were monomer peak size, monomer to aggregate peak size ratio and other characteristics of the peak itself (shape, separation from other peaks, tailing, or peak shoulders). All the purification results are normalized for a 10 mm UV cell and for a 500 mL cell pellet unless mentioned otherwise.

#### **4.1.1 Expression of constructs**

The bacmid preparation for all GPR68 constructs was confirmed using agarose gel electrophoresis after DNA extraction and PCR (section 3.2.1). A representative agarose gel electrophoresis of GPR68 C21 recombinant bacmids PCR products highlights two important scenarios. The first sample, C21-Bacmid 1, where the 300 bp band for the bacmid was not observed (Figure 19A) explains that 100% transposition had occurred for the bacmid with the pFastbac1 (2300 bp) containing the GPR68 insert. The total size of this would be the 'size of the pFastbac1' + 'size of the GPR68 C21 insert' i.e., (2300 bp + 1617 bp). The second scenario where recombinant DNA of GPR68 C21 PCR product band along with the 300 bp bacmid PCR product band are visible (Figure 19A) would represent the case, where I had incomplete transposition of the bacmid DNA. In certain scenarios, when I only had the 300 bp bacmid PCR product band would mean that either the transposition failed, or incorrect colony was picked after blue-white screening. All the GPR68 bacmid DNAs were screened as described above and only the samples containing pure recombinant bacmid DNA were used for transfection. After the transfection with the correct recombinant DNA followed by multiple virus amplification rounds, the optimal virus titers amounts were used to express each of the GPR68 constructs. The expression of different GPR68 constructs were tested using Anti-His primary antibody in Western blots using the expressed pellet samples (section 3.2.2). A representative Western blot has been shown in this section for the constructs C1-C9 (Figure 19B), which shows varying expression levels as indicated by varying intensities of bands for different GPR68 constructs. This method is also efficient for GPR68 pilot expression level screens as we can clearly observe the signal and differences in the signal due to varying molecular weight of the GPR68 constructs (Figure 19B). The overexpressed receptor construct bands are indicated by arrows (Figure

19B). It also crucial to observe that these constructs run slightly lower than their expected theoretical molecular weight, which is also very common for GPCRs and other membrane proteins. This is likely because the SDS alone is not enough to denature the proteins samples thus resulting in non-optimal charge distribution along with compact protein structures leading to differences in migration. The same result may lead to presence of a band for each construct at twice the size (Figure 19B), such oligomeric bands are often observed for GPCRs in SDS-PAGE gels and Western blots. Moreover, post-translational modifications such as N-glycosylation and palmitoylation, frequently occur in GPCRs and can result in GPCRs having bands with apparent molecular weights that are higher than expected, which adds to the complexity in interpretation of SDS-PAGE gels. Of all the GPR68 constructs (C1-C75, and C99), the overall expression levels were very low for C1 (WT-like) and C5 (mT4L) (Figure 19B), and no expression was observed for C11.

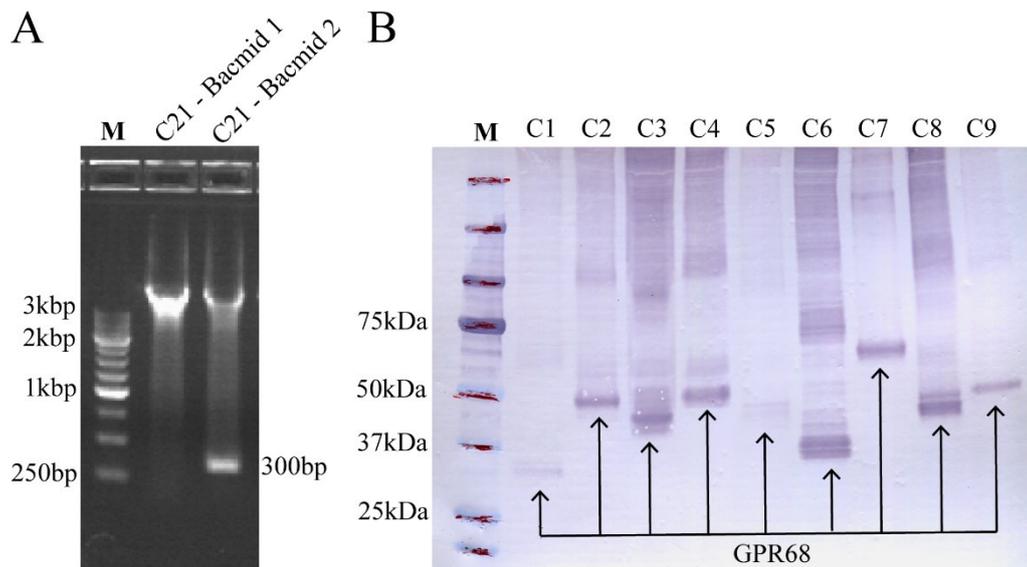


Figure 19: Representative GPR68 PCR products and the optimal baculovirus titer screening for GPR68 constructs. (A) Agarose gel electrophoresis of two bacmids from GPR68 C21, where M represents the DNA marker. (B) Anti-His tag Western blot from the SDS-PAGE of the overexpressed proteins from the cell pellet samples of GPR68 C1-C9, here M represents the protein marker.

Construct ID	N-terminus	N-terminal Fusion	GPR68	C-terminus
<b>WT-like construct</b>				
C1	FLAG	-	1-365	10× His
<b>Initial fusion (ICL3) constructs</b>				
C2	FLAG	-	216--T4L--217	10× His
C3	FLAG	-	216--BRIL--217	10× His
C4	FLAG	-	216--PGS--217	10× His
C5	FLAG	-	216--mT4L--217	10× His
C6	FLAG	-	214--RDX--220	10× His
<b>N-terminal truncation testing constructs</b>				
C9	FLAG, Δ8	-	216--T4L--217	10× His
C10	FLAG, Δ14	-	216--T4L--217	10× His
C11	FLAG, Δ20	-	216--T4L--217	10× His
<b>C-terminal truncation testing constructs</b>				
C12	FLAG	-	216--T4L--217	Δ18, 10× His
C13	FLAG	-	216--T4L--217	Δ52, 10× His
C14	FLAG	-	216--T4L--217	Δ55, 10× His
C15	FLAG	-	216--T4L--217	Δ60, 10× His
<b>N-terminal BRIL Fusion and truncation combination constructs</b>				
C8	FLAG, Δ8	BRIL	-	10× His
C27	FLAG, Δ14	BRIL	-	Δ18, 10× His
C28	FLAG, Δ14	BRIL	-	Δ52, 10× His
C29	FLAG, Δ14	BRIL	-	Δ60, 10× His
C32	FLAG	BRIL	-	10× His
C33	FLAG	BRIL	-	Δ18, 10× His
C34	FLAG	BRIL	-	Δ52, 10× His
C35	FLAG	BRIL	-	Δ60, 10× His
<b>ICL3 Fusion and truncation combination constructs</b>				
C7	FLAG, Δ8	BRIL	216--T4L--217	10× His
C18	FLAG, Δ8	-	216--PGS--217	10× His
C19	FLAG, Δ8	-	216--PGS--217, N86Q	10× His
C20	FLAG, Δ14	-	216--T4L--217	Twin-Strep
C21	FLAG, Δ14	-	216--T4L--217	Δ18, 10× His
C24	FLAG, Δ14	-	216--PGS--217	Δ18, 10× His
C66	FLAG, Δ8	BRIL	214--BLEK--221	Δ18, 10× His
C67	FLAG, Δ14	-	214--CalB--219	Δ46, CalA, 10× His
<b>ICL3 BRIL Fusion and truncation combination constructs</b>				
C16	FLAG	-	214--BRIL--219	10× His
C17	FLAG	-	211--BRIL $\chi$ --221	10× His
C30	FLAG, Δ14	-	211--BRIL $\chi$ --221	10× His
C31	FLAG	-	214--BRIL--219	Δ18, 10× His
C43	FLAG, Δ14	-	214--BRIL--219	Δ18, 10× His
C61	FLAG, Δ14	-	214--BRIL $\chi$ --221	Δ18, 10× His
C75 (HEK293)	FLAG	-	214--BRIL $\chi$ --221	Δ18, 10× His
<b>ICL1 or ICL2 BRIL Fusion and truncation combination constructs</b>				
C73	FLAG, Δ8	-	50--BRIL(ICL1)--55	Δ46, 10× His
C74	FLAG, Δ8	-	122--BRIL(ICL2)--135	Δ46, 10× His
<b>Zebrafish construct</b>				
C99	FLAG, Δ8	GFP	209--BRIL--219, N9Q	Δ36, Twin-Strep

Table 3: A summary of construct for which chromatograms are shown and described in this section. The truncation of amino acids represented with Greek letter ‘Δ’ followed by number of amino acids. The fusion protein inserted in an ICL of GPR68 is shown with the insertion region (amino acids between which fusion protein is inserted) separated by a double hyphen (--). The fusions are inserted in ICL3 unless mentioned otherwise e.g., C73. Note that other redundant details like cleavage tags, linkers and signal peptide are not

described in this table. Other specific mutation construct details have been described in Table 4. The abbreviations used and the other construct details has been described in Appendix 1.

#### 4.1.2 Screening of different protein fusions to GPR68

Low expression levels of C1 (wild type-like construct) necessitates the use of protein engineering to increase expression levels and stability of the receptor using fusion protein-based modifications which are often used for class A GPCRs (section 1.4.1). The first step was to introduce fusion proteins in the ICL3 of the receptor, as most of G proteins couple to ICL3 in Class A GPCRs for activation. Thus, introducing a fusion protein in this loop increases the likelihood of locking the receptor in an inactive conformation and can likely provide crystals contacts for LCP.

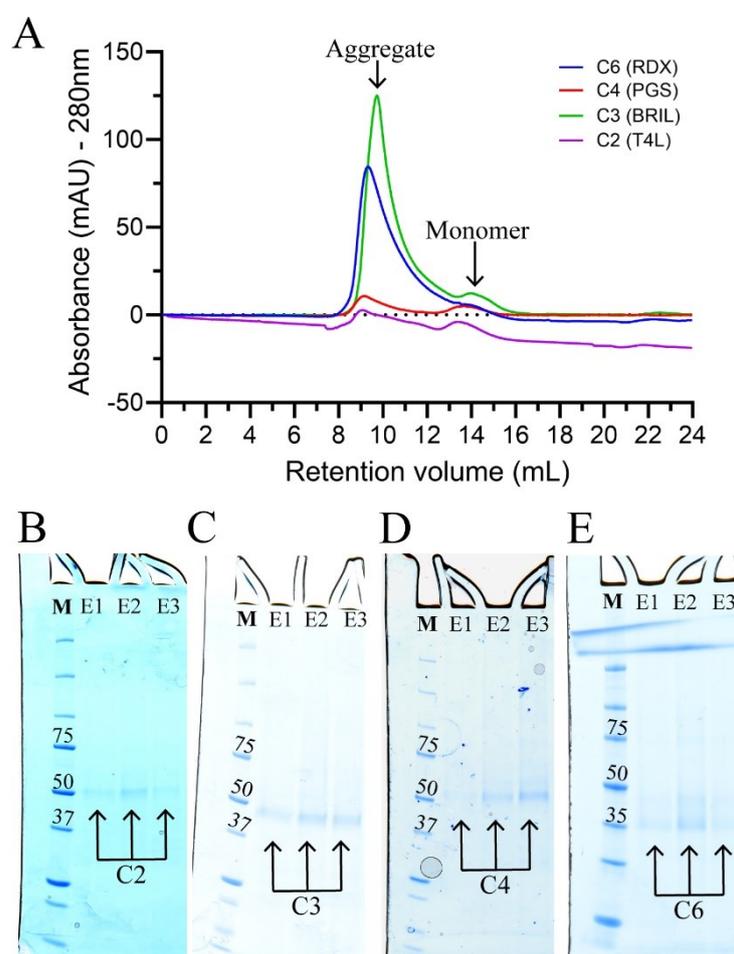


Figure 20: Screening of different fusion proteins inserted in the ICL3 of GPR68. (A) Size exclusion chromatograms from the pooled eluates after affinity purification (TALON) for C2-C4 and C6. The SEC was performed on Superdex 200 increase 10/300 GL from 500 mL cell pellet and all the purification steps were performed using DDM/CHS at pH 7.8. (B-E) SDS PAGE of the first three TALON eluates for C2 (panel B), C3 (panel C), C4 (panel D) and C6 (panel E). The 'M' represents the protein marker, also note that the marker used in panel E is different from panel B-D. Note that the numbers on the marker bands represent the corresponding molecular weight in kDa for the band.

The fusion proteins that showed improved expression of the modified GPR68 constructs were rubredoxin (RDX), *Pyrococcus abyssi* glycogen synthase (PGS), apocytochrome b562 (with M7W/H102I/R106L mutations) (BRIL) and T4 lysozyme (T4L) (Figure 19B, 20A). The expression levels between constructs are compared based on Western blot results and chromatograms (ratio of aggregate and monomer peak maximum). The overall yields were higher for BRIL (C3) and RDX (C6) based constructs in comparison to T4L (C2) and PGS (C4) fusions. Although, based on the monomer peak characteristics and monomer to aggregate peak maximum ratio for C2 and C4 was better than C3 and C6 (Figure 20A). The affinity purified proteins could also be observed on the Coomassie-stained SDS-PAGE gels at the expected molecular weights (Figures 18 B-E). The yields were not enough for the protein to be stained after SEC. Based on the parameters described above, all fusions, except for RDX, were selected for further testing. The above tested fusion protein combinations for GPR68 did not yield enough protein to pursue structural studies, therefore it was clear that more protein engineering was required to further improve overall protein yield and stability.

#### 4.1.3 Effect of termini truncations

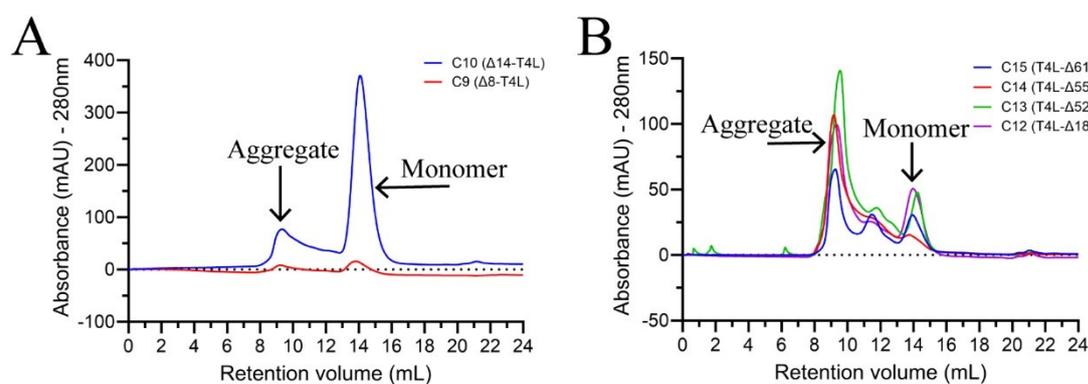


Figure 21: Screening of the T4L fusion protein inserted in the ICL3 of GPR68 with truncation at either terminus. (A) Size exclusion chromatograms from the pooled eluates after affinity purification (TALON) for the *N*-terminal truncation constructs (C9 and C10). (B) Size exclusion chromatograms from the pooled eluates after affinity purification (TALON) for the *C*-terminal truncation constructs (C12-C15). The SEC was performed on Superdex 200 increase 10/300 GL from 500 mL cell pellet and all the purification steps were performed using DDM/CHS at pH 7.8. Note that the Greek letter  $\Delta$  represents truncation, if the  $\Delta$  is before the fusion protein, it is an *N*-terminal truncation e.g., C9 ( $\Delta$ 8-T4L) means that the 8 amino acids from the *N*-terminus are truncated, whereas if  $\Delta$  is after the fusion it represents *C*-terminal truncation e.g., C12 (T4L- $\Delta$ 18) means that the 18 amino acids from the *C*-terminus are truncated.

Another often used approach involves truncating either terminus of GPCRs to enhance protein production and stability (section 1.4.1). Furthermore, removal of disordered termini may improve crystal contacts during crystallization. I used the T4L fusion in the ICL3 of

GPR68 and tested 3 *N*-terminal truncations and 4 *C*-terminal truncations. Out of the three *N*-terminal truncation constructs (C9, C10 and C11), C11 did not express, which is likely because of excessive truncation ( $\Delta 20$  amino acids) into the TM1 of GPR68, as observed from snake diagram of GPR68 in GPCRdb (Figure 18). Amongst all *N*-terminal truncation constructs a deletion of 14 amino acids (C10) was clearly better than deletion of 8 amino acids (C9) based on monomeric peak characteristics as well as monomer to aggregate peak size ratio (Figure 21A). However, the SEC profile of C10 was not reproducible and the behavior was later found to be like C9, except with more overall yields. The explanation for the lack of reproducibility of C10 chromatogram has been described later in my thesis under the section 4.1.5. Finally, amongst the *C*-terminal truncation-based constructs (deletion of 18, 52, 55, and 61 amino acids in the constructs C12, C13, C14 and C15 respectively), the monomer peak size and overall SEC profile was similar amongst all *C*-terminal truncation-based constructs. On comparing the T4L based GPR68 construct without any truncations (C2) to the overall truncation results showed that all the truncation constructs showed significant overall yield improvements (Figure 20A, 21A,B), of which the *C*-terminal truncations seem to be more effective. This is likely because the *C*-terminus is longer than the *N*-terminus in GPR68 and truncation of longer likely disordered regions would show more improvements in protein behavior.

#### 4.1.4 Combination of fusions, truncations, and variations of fusion insertion site

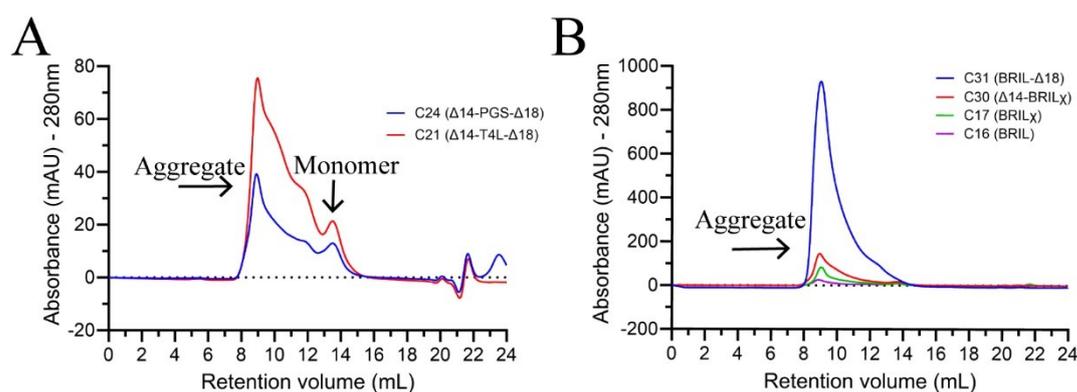


Figure 22: Screening of constructs containing a combination of a fusion protein inserted in the ICL3 of GPR68 along with a truncation at either termini or both termini. (A) Size exclusion chromatograms from the pooled eluates after affinity purification (TALON) for the GPR68 constructs which have either T4L (C21) or PGS (C24) as ICL3 fusion. (B) Size exclusion chromatograms from the pooled eluates after affinity purification (TALON) for the GPR68 constructs based on ICL3 BRIL insertion. These BRIL constructs vary in insertion points in the ICL3, addition of chimeric BRIL (BRIL $\chi$ ), or a combination with truncation. The SEC was performed on Superdex 200 increase 10/300 GL from 500 mL cell pellet and all the purification steps were performed using DDM/CHS at pH 7.8. The truncation denotation used in this figure is as formerly described in Figure 21.

The next set of GPR68 constructs involved a combination of *N*-terminal truncation like C10 ( $\Delta 14$ ) and *C*-terminal truncation like C12 ( $\Delta 18$ ) along with an ICL3 fusion – T4L (C21) and PGS (C24). This strategy seemed ineffective as the overall yields decreased and monomeric peak characteristics were worse (Figure 22A) in comparison to when the above-described strategies were implemented separately (Figure 20, 21). Additional constructs including T4L (C22 and C23) or PGS (C25 and C26) ICL3 fusions were created, differing only in their *C*-terminal truncation (Appendix 1). These constructs had comparable *C*-terminal truncations like C13 and C15 for each ICL3 fusion protein and exhibited similar behavior to C21 and C24 i.e., resulting in a reduced overall yield and poor monomer peak characteristics (Figure 22A).

I also investigated a distinct set of constructs that aimed to optimize the insertion of BRIL in ICL3 of GPR68. These constructs were designed using alternative insertion locations compared to the previously tested BRIL construct (C3). These BRIL constructs did not show any monomeric peak and were not any better than strategies previously accomplished (Figure 22B). The overall yields were also poor except for C31 (Figure 22B) which included a truncation like construct C12.

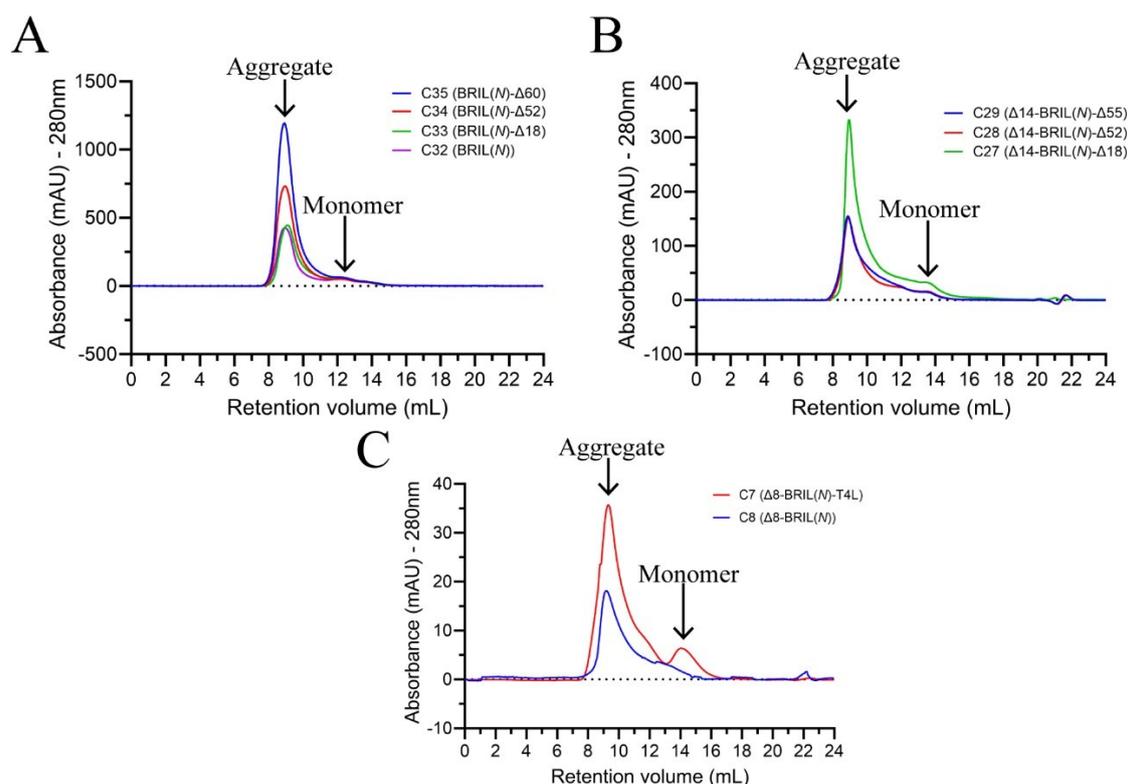


Figure 23: Screening of the combination constructs that include a BRIL fusion protein inserted at the *N*-terminus of the GPR68 constructs and it is denoted by 'BRIL(*N*)'. (A) Size exclusion chromatograms from

the pooled eluates after affinity purification (TALON) for the GPR68 constructs which have variations in the C-terminal truncation and a BRIL(*N*). (B) Size exclusion chromatograms from the pooled eluates after affinity purification (TALON) for the GPR68 constructs having BRIL(*N*) and a combination of termini truncations. (C) Size exclusion chromatograms from the pooled eluates after affinity purification (TALON) for the GPR68 constructs having BRIL(*N*) and a combination of termini truncations. The SEC was performed on Superdex 200 increase 10/300 GL from 500 mL cell pellet and all the purification steps were performed using DDM/CHS at pH 7.8. The truncation denotation used in this figure is as formerly described for Figure 21.

A similar strategy of adding *N*-terminal BRIL was implemented, but with an additional removal of the first 14 amino acids at the *N*-terminus (C27-C29), where it was observed that lesser truncation of *C*-terminus was more effective to obtain more overall yields (Figure 23B). The yields of the constructs C27-C29 (~150-400 mAU) was less than C32-C35 (~500-100 mAU). Lastly, the constructs C7 and C8 had very poor yields (Figure 23C) and form the worst set of *N*-terminal BRIL constructs.

#### 4.1.5 Effect of cobalt ions on GPR68 stability

In the section 4.1.3, I described that C10 had an ideal SEC profile (Figure 21A), but the result was not reproducible. Upon further investigation, I found out that the TALON resin was likely leaking cobalt ions as it was an older batch stored for a long time. The cobalt ions leaking from the resin were then contributing to the stability of C10. This hypothesis was confirmed by repeating the C10 purification using a new batch of TALON resin. A distinct SEC profile was found, showing a reduced amount of stable protein (Figure 24A) compared to the previous purification performed with the old batch of TALON resin that was previously described (Figure 21A). To further confirm the theory, the same purification was repeated with a fresh batch of TALON resin in presence of 1 mM EDTA in the SEC buffer, and interestingly the overall amount of protein on the SEC column was reduced drastically with barely any protein in the typical monomeric protein retention volume (Figure 24A). Additionally, a FLAG based affinity purification was performed in presence of 100  $\mu$ M cobalt chloride and a monomeric peak was observed (Figure 24A). Thus, validating my theory about the leaking metal ions. To investigate the role of the *C*-terminal 10 $\times$  His-tag in stabilizing C10 via forming a possible structured region with *C*-terminus, a new construct C20 like C10 was designed with a Twin-Strep tag substituting the 10 $\times$  His-tag (Table 3). C20 was subjected to FLAG based affinity purification in presence of 100  $\mu$ M cobalt chloride and still showed stabilization via cobalt ions (Figure 24B). Although, the yields had reduced to about 50% of C10.

The C12 was also purified via FLAG tag in presence of 100  $\mu$ M cobalt chloride and this construct was also found to be stabilized via cobalt ions. However, when the concentration of cobalt chloride was reduced to 25  $\mu$ M, no stability was seen. This suggests that the receptor's affinity for cobalt ions is likely low.

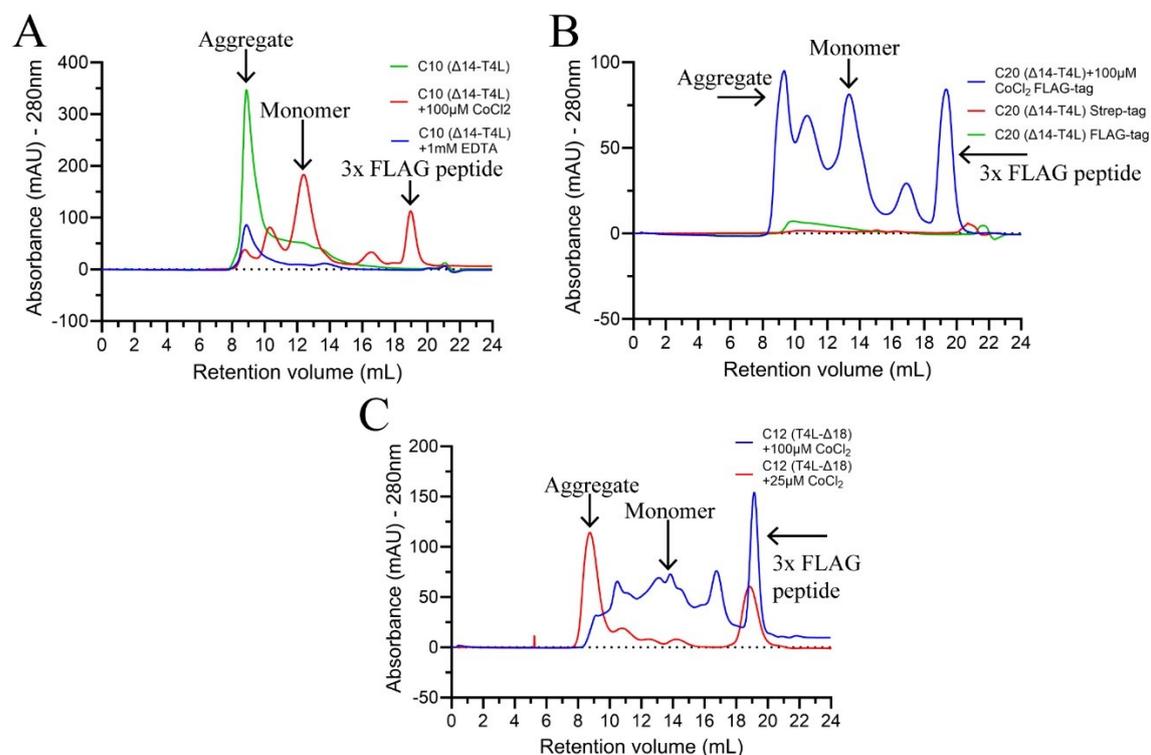


Figure 24: Screening of the combination constructs (C10, C12 and C20) in presence of cobalt chloride. Size exclusion chromatograms from the pooled eluates after affinity purification (TALON, if not mentioned otherwise) for GPR68 C10 (panel A), GPR68 C20 (panel B), and GPR68 C12 (panel C). The SEC was performed on Superdex 200 increase 10/300 GL from 500 mL cell pellet and all the purification steps were performed using DDM/CHS at pH 7.8. The truncation denotation used in this figure is as formerly described for Figure 21.

Minimal quantities of monomeric protein were detected for C7, C21, and C24 (Figure 25A), however the overall yield improved compared to the purification without cobalt ions (Figure 22A, 23C). Similar trend was observed for constructs that had a ICL3 BRIL fusion (Figure 22B, 24B), but C17 containing the chimeric BRIL (BRIL $\chi$ ) showed a distinct monomeric peak (Figure 25B).

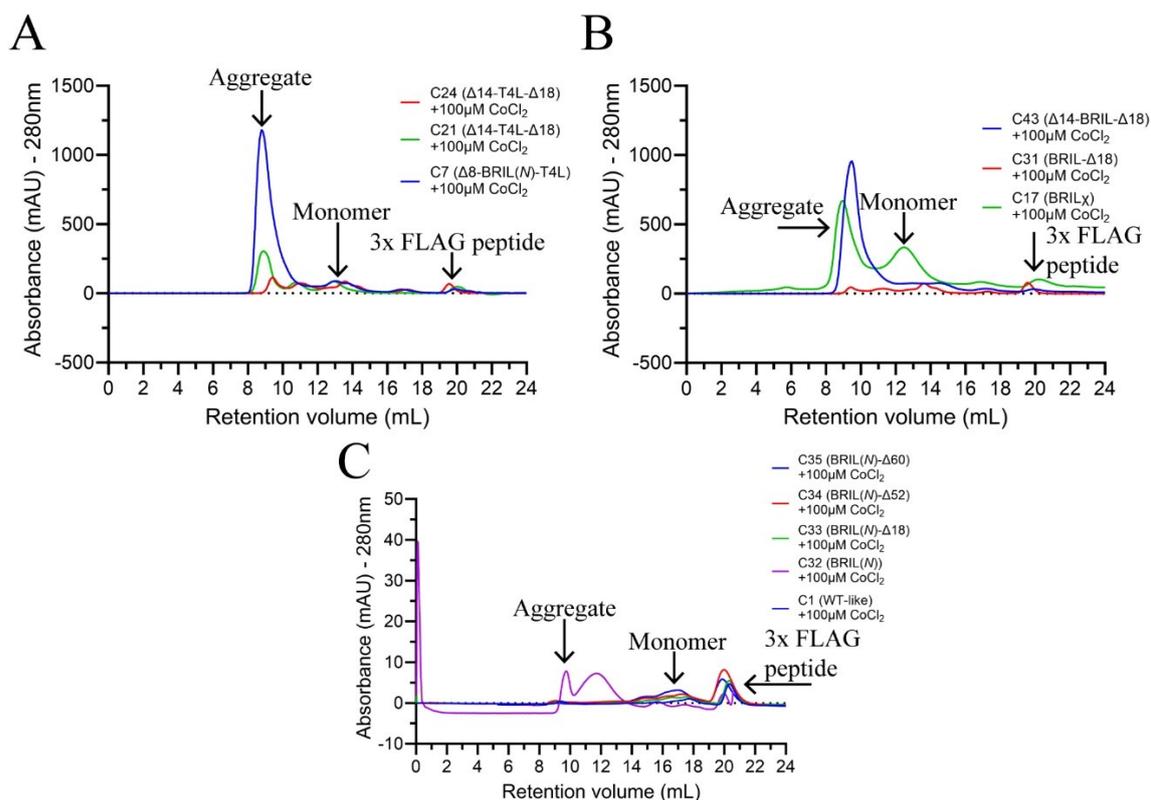


Figure 25: Screening of the various combination constructs in presence of cobalt chloride. Size exclusion chromatograms from the pooled eluates after affinity purification (FLAG) for GPR68 C7, C21 and C24 (panel A), GPR68 C17, C31 and C43 (panel B), and GPR68 C1, C32, C33, C34 and C35 (panel C). The SEC was performed on Superdex 200 increase 10/300 GL (panel A,B) and on Superose 6 increase 10/300GL (panel C) from 500 mL cell pellet and all the purification steps were performed using DDM/CHS at pH 7.8 (panel A,B) or at pH 6.8 (panel C). The truncation denotation used in this figure is as formerly described for Figure 21.

I investigated the potential of cobalt ions to enhance the stability of GPR68 constructs with a BRIL fusion at the *N*-terminus (BRIL(*N*)) and a native ICL3, thereby facilitating the required conformational change in GPR68 for the active state structure. Upon testing the idea, extremely small amounts of yield were observed for all constructs, although interestingly no aggregate peak was observed for these constructs except for C32 (Figure 24C).

#### 4.1.6 Addition of small molecules enhanced protein stability

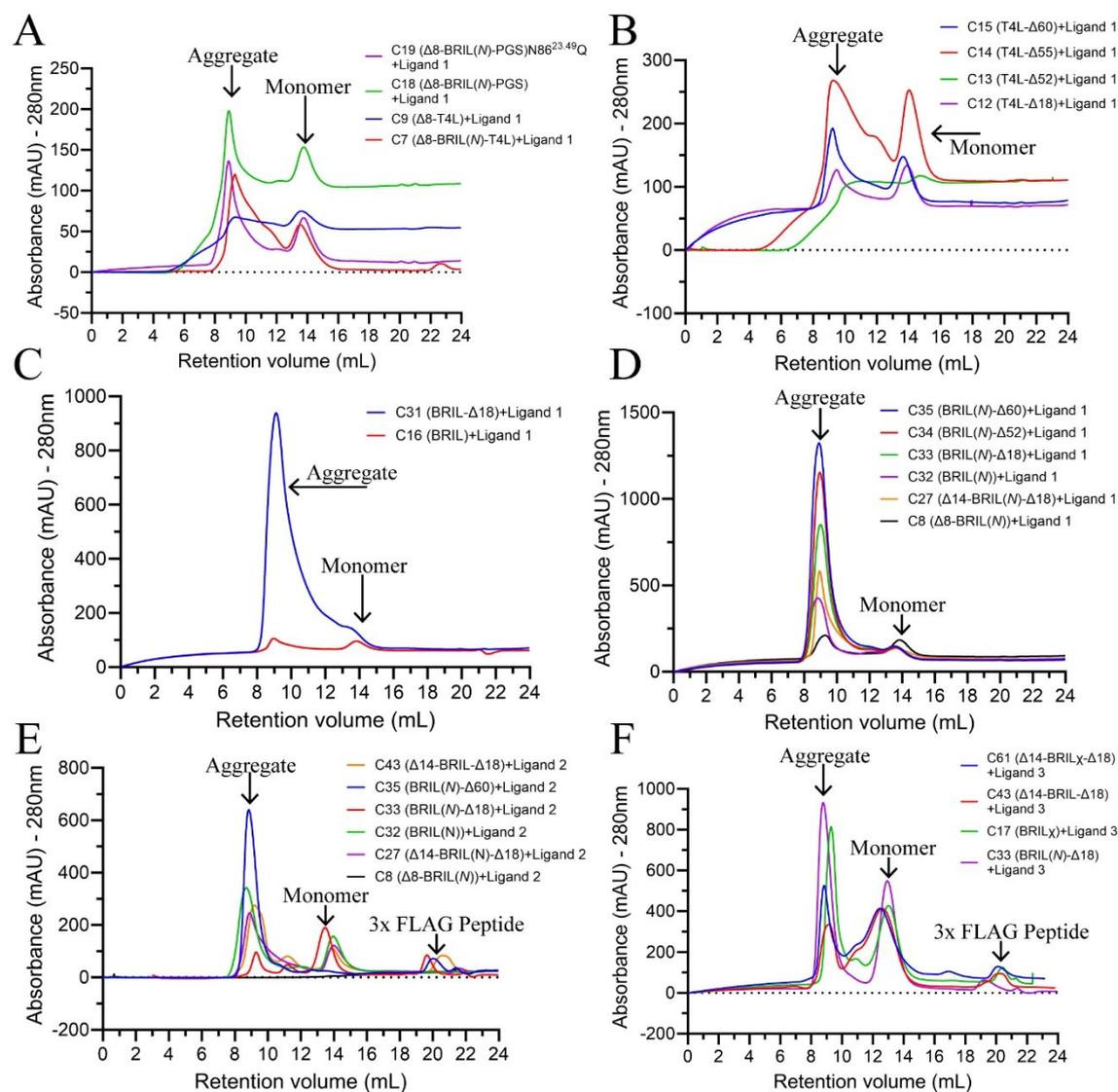


Figure 26: Screening of the GPR68 constructs in presence of proprietary ligands (Ligand 1, Ligand 2 and Ligand 3). Size exclusion chromatograms from the pooled eluates after affinity purification TALON (panel A-D) and FLAG (panel E-F) for purification of different GPR68 constructs in Ligand 1 (panel A-D), Ligand 2 (panel E) and Ligand 3 (panel F). The SEC was performed on Superdex 200 increase 10/300 GL from 500 mL cell pellet and all the purification steps were performed using DDM/CHS at pH 7.8. The truncation denotation used in this figure is as formerly described for Figure 21.

On average, the protein yields have consistently been very low for different GPR68 constructs when purified without ligands or in the presence of cobalt ions. I was provided with over a dozen proprietary ligands which were identified as antagonists against human GPR68. To identify the most effective ligands for stabilizing the human GPR68 in a non-native environment, a range of representative constructs were purified using different ligands. I only report the ligands that stabilized the human GPR68.

The initial testing of various GPR68 constructs was performed using Ligand 1, which stabilized all the different types of GPR68 constructs (Figure 26A-D) compared to apoprotein (section 4.1.2-4.1.4) based on overall yields and monomeric protein yields.

A distinct monomeric peak, with improved protein stability and overall yields were observed using Ligand 2 and Ligand 3 when various types of combination constructs containing truncations, different fusion proteins in ICL3 or *N*-terminal BRIL (Figure 26E-F) in comparison to apoprotein, or in presence of Ligand 1 or cobalt ions. Ligand 3 contributed to higher stabilization of the receptor constructs and thus was used to produce protein for structure determination.

#### 4.1.7 Effect of potential thermostabilizing mutations

Construct	Mutation	Rationale
<b>Mutations introduced into C19 ((Δ8-BRIL(N)-PGS) N86<sup>23,49</sup>Q) as base construct</b>		
C36	F241 <sup>6,48</sup> W	Restoring CWxP motif
C37	I226 <sup>6,33</sup> W	Predicted by CompoMug <sup>[246]</sup>
C38	G110 <sup>3,41</sup> W	Predicted by CompoMug <sup>[246]</sup> / GPCRdb <sup>[362]</sup>
C39	Y205 <sup>5,58</sup> A	Destabilize active state conformation stabilized via Y <sup>5,58</sup> -Y <sup>7,53</sup> of DPxxY <sup>7,53</sup> motif <sup>[363]</sup>
C40	D67 <sup>2,50</sup> N	Neutralizing Sodium pocket
C42	T293 <sup>8,50</sup> F	High conservation in Helix 8 <sup>[364]</sup>
<b>Mutations introduced into C17 (BRIL<sub>γ</sub>) as base construct</b>		
C62	S108 <sup>3,39</sup> R	Sodium pocket ionic lock (D67 <sup>2,50</sup> -S108 <sup>3,39</sup> R) <sup>[365]</sup>
<b>Mutations introduced into C33 (BRIL(N)-ΔI8) as base construct</b>		
C46	C304A, C310A	Potential palmitoylation sites
C47	G46 <sup>1,57</sup> A	Predicted by CompoMug <sup>[246]</sup>
C48	L60 <sup>2,43</sup> I	Decrease the size of hydrophobic pocket <sup>[366]</sup>
C49	L112 <sup>3,43</sup> A	Increase the size of hydrophobic pocket <sup>[366]</sup>
C50	I115 <sup>3,46</sup> A	Increase the size of hydrophobic pocket <sup>[366]</sup>
C51	K223 <sup>6,30</sup> E	Restore DR <sup>3,50</sup> Y motif potential interaction partner R119 <sup>3,50</sup> -K223 <sup>6,30</sup> E
C52	T64 <sup>2,47</sup> A	Predicted by CompoMug <sup>[246]</sup>
C53	V109 <sup>3,40</sup> A	Disrupting PV <sup>3,40</sup> F motif
C54	E103 <sup>3,34</sup> A	Potential role in proton sensing (section 4.2.4 and 4.3.3 of my thesis)
C55	E149 <sup>4,53</sup> A	Role in proton sensing <sup>[134]</sup>
C57	T233 <sup>6,40</sup> V	Restore hydrophobicity of the pocket <sup>[366]</sup>
C68	F76 <sup>2,59</sup> C	Introducing artificial disulfide bond
C70	C113 <sup>3,44</sup> L, C114 <sup>3,45</sup> L	Class conservation based on GPCRdb sequence alignment
C71	S108 <sup>3,39</sup> K	Sodium pocket ionic lock (D67 <sup>2,50</sup> -S108 <sup>3,39</sup> K) <sup>[367]</sup>
C72	L74 <sup>2,57</sup> P	<i>Amelogenesis imperfecta</i> mutant <sup>[189]</sup>

Table 4: Compilation of potential thermostabilizing mutations for human GPR68.

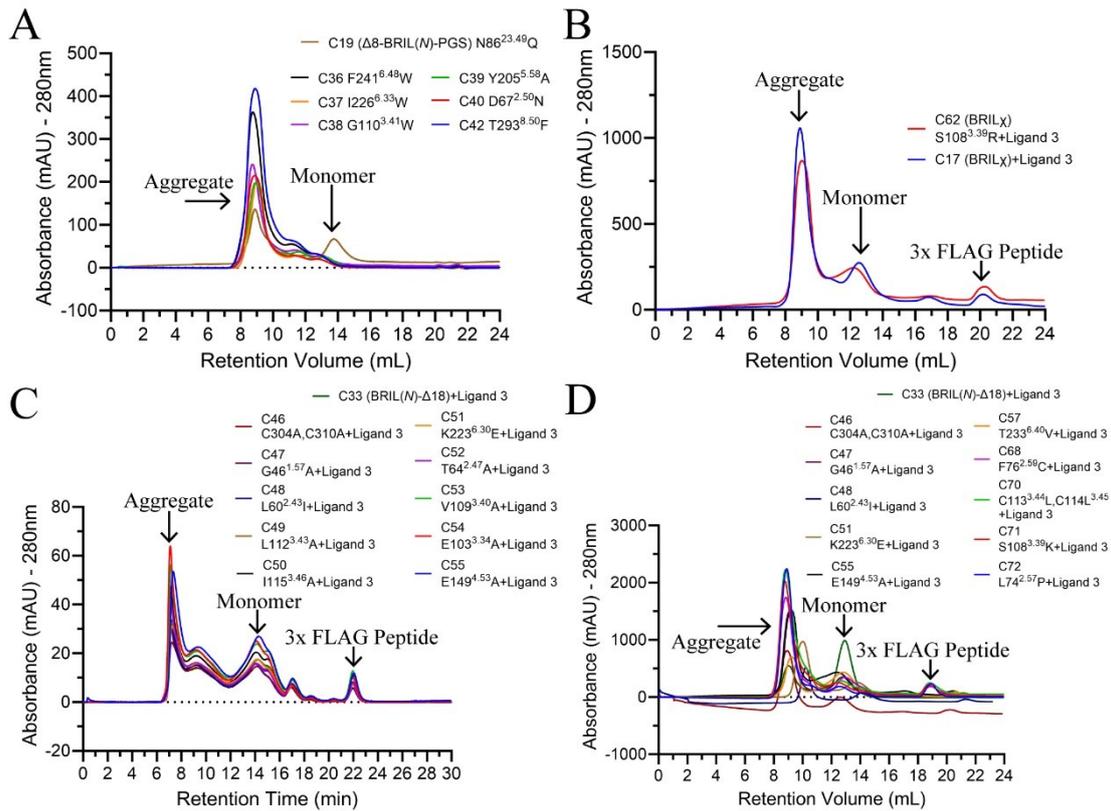


Figure 27: Screening of the GPR68 point mutations. Size exclusion chromatograms from the pooled eluates after affinity purification TALON (panel A,B) and FLAG (panel C,D) for purification of different GPR68 mutation constructs. The SEC was performed on Superdex 200 increase 10/300 GL from 500 mL cell pellet (panel A,B and D) or on Superose 6 increase 5/150 GL column (panel C) using 50 mL cell pellet and all the purification steps were performed using DDM/CHS at pH 7.8 without ligand (panel A) in presence of Ligand 3 (panel B-D). The truncation denotation used in this figure is as formerly described for Figure 21.

GPR68 mutations were tested in different groups throughout the thesis using three different constructs (C17, C19 and C33) in which mutations were designed, referred to as ‘base construct’. The mutations that were tested and the rationale for their design have been presented in Table 4. The mutations predicted by ComoMug <sup>[246]</sup> were proposed by Dr. Christofer Tautermann (Boehringer Ingelheim, Biberach, Germany).

The mutations designed using C19 as a base construct resulted in an increased protein aggregation and reduced monomer formation as inferred from chromatogram (Figure 27A). C62 was also an ineffective mutant as it had a similar chromatogram to its base construct, C17 (Figure 27B). The mutations that were inserted in C33 when tested in small scale showed extremely similar SEC profile for all the constructs tested, but a higher monomeric peak was observed for C50, C53, C54 and C55 in comparison to C33 (Figure 27C). A few mutation constructs from the previously described set (C46, C47, C48, C51) and the mutant construct that had improved peak characteristics (C55, E149A<sup>4.53</sup>) in comparison to the

base construct (C33) were purified again on a large scale (500 mL) and were observed to behave worse than the base construct (C33) in terms of monomeric peak yield and characteristics (Figure 27D). Thus, the next set of constructs were directly tested in a 500 mL scale and were also found to be worse than the base construct, C33 (Figure 27D).

Ultimately, all efforts to stabilize GPR68 by mutations were unsuccessful.

#### 4.1.8 Anti-BRIL FAB complexation trials

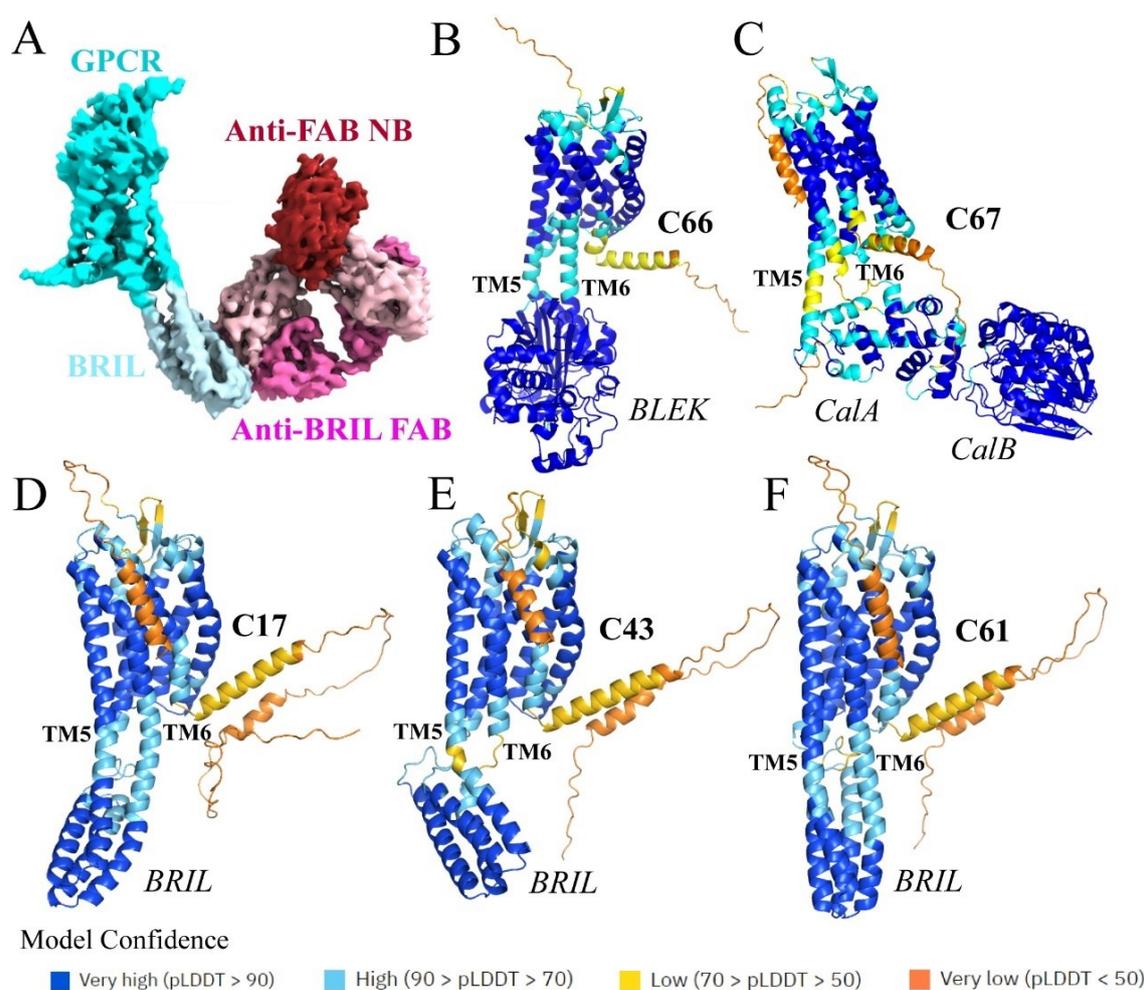


Figure 28: Construct design strategy for GPCRs using artificial intelligence based structural modeling tool, AlphaFold<sup>[213]</sup>. The acronyms used in this figure are described as follows for the fragment antigen-binding region (FAB), single monomeric variable antibody domain designed to bind Anti-BRIL FAB (NB), beta-lactamase from *Escherichia coli* K12 (BLEK), calcineurin sub-unit A (CalA), calcineurin sub-unit A (CalB) and per-residue local distance difference test (pLDDT), a global model scoring system. (A) Representative BRIL-FAB-NB complex. This panel was adapted and reused with permission under Creative Commons Attribution license<sup>[368]</sup>. Structural models generated for different GPR68 constructs containing the following fusion proteins – beta lactamase (panel B), calcineurin (panel C), BRIL fusions based (panel D-F).

Cryo-EM of class A GPCRs have been successful earlier using the BRIL-FAB strategy, where a GPCR construct with a BRIL fusion in ICL3 is purified and complexed with Anti-BRIL FAB and Anti-BRIL-FAB nanobody (NB) (Figure 28A). AlphaFold was publicly released while I pursued my thesis work, I used it to design rigid BRIL fusion based GPR68 constructs. Here, the term ‘rigid’ represents a continuous helix formation of the region where the BRIL fusion is added e.g., if BRIL is added in ICL3, then TM5 and TM6 of the GPCR would ideally form continuous helices extending into the BRIL fusion. The strategy used to design such GPR68 constructs involved usage of sequence manipulation (addition or deletion of residues) of the GPR68 sequence in the ICL3 region followed by addition of BRIL or BRIL $\chi$  sequence to this manipulated region. Many such sequences with different combinations were prepared and used as input files for the AlphaFold algorithm. This cycle was continued until a construct with a potential rigid BRIL fusion was observed from the structural models generated by AlphaFold with a consensus which can be further used for experimentation. An example of such a construct generated with high confidence score was C61 (Figure 28F). Then, I back tested some of the BRIL constructs which were design before AlphaFold was released and found that C17 (Figure 28D) was more rigid than C43 (Figure 28E).

Complex formation was tested using purified GPR68 constructs containing the ICL3 BRIL fusion constructs (C17, C43, and C61), as well as an *N*-terminal BRIL fusion (BRIL(*N*)) construct (C33), using the Anti-BRIL FAB (FAB) and/or Anti-BRIL-FAB nanobody (NB) at pH 7.8. The initial attempt was conducted qualitatively by adding an excessive amount of FAB to the purified protein. The objective was to notice a decrease in retention volume, because the size of the complex (GPR68 construct + FAB) is larger than that of the monomeric protein. However, the results indicated that complexation of FAB with both types of constructs – the ICL3 BRIL fusion constructs and the *N*-terminal BRIL fusion construct did not cause any change in retention volume due to changes in size or molecular weight (Figure 29A-D). Consequently, the experiment was conducted again with C61 quantitatively at two distinct pH values (pH 6.0 and pH 7.8), resulting in identical outcomes, i.e., no observable shift in retention volumes (Figure 29E,F). Another key observation from these set of experiments was that upon complex formation the size of the “complex peak” would reduce instead of increasing when compared to the size of monomeric peak (Figure 29E,F).

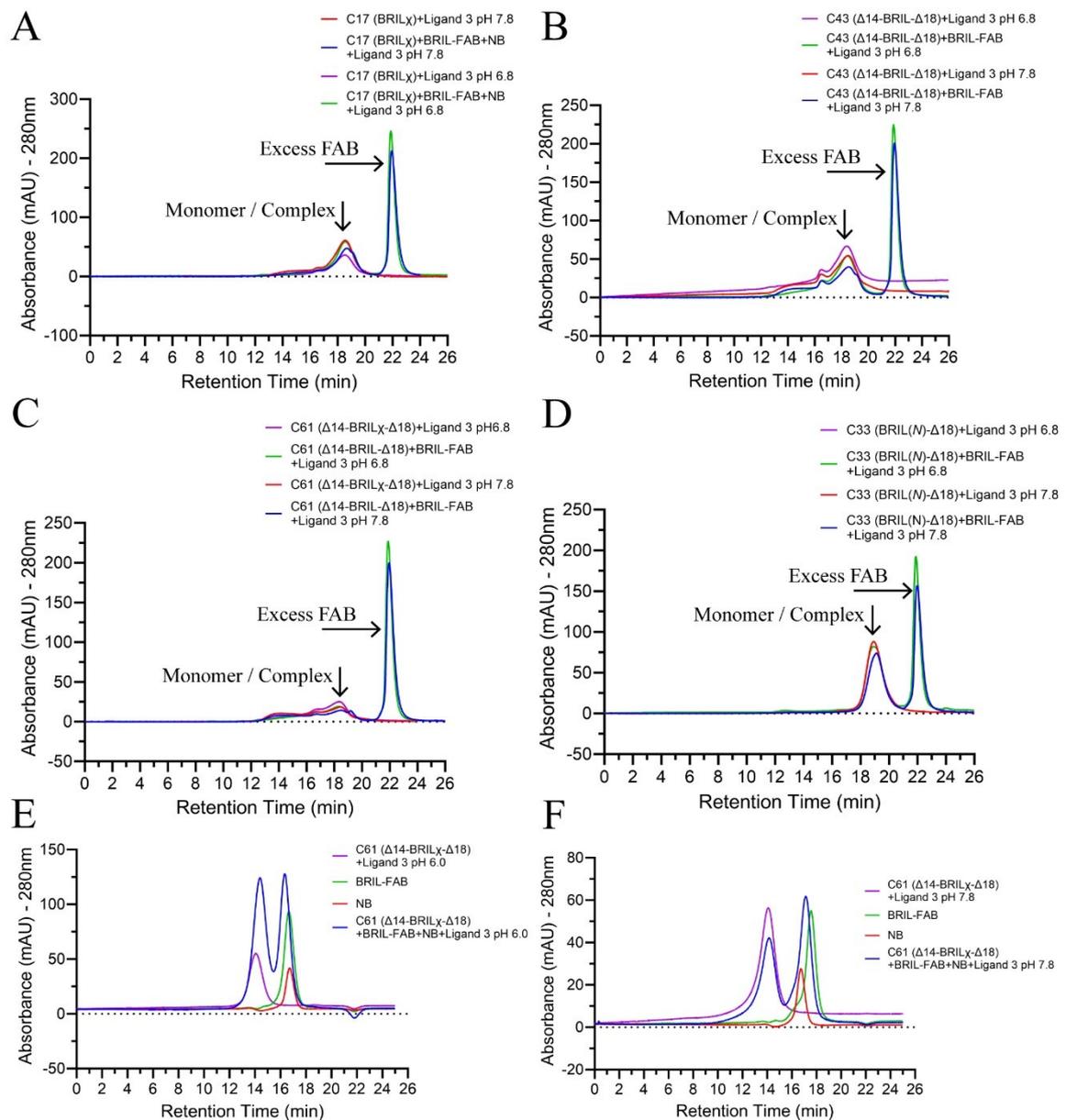


Figure 29: Screening of the GPR68 BRIL fusion constructs for complexation with Anti-BRIL FAB. Size exclusion chromatograms from the pooled eluates after affinity purification FLAG and SEC for complexation with Anti-BRIL FAB. The SEC was performed on Unix-C SEC-300 4.6/300 column for C17 (panel A), C43 (panel B), C61 (panel C) and C33 (panel D) or on Superose 6 increase 5/150 GL column for C61 (panel E,F) All the purification steps were performed using DDM/CHS at different pH values indicated in the respective figure legend. All purification were performed in presence of Ligand 3. The truncation denotation used in this figure is as formerly described for Figure 21.

In consideration of the puzzling outcomes, I decided to proceed with an alternative approach. Specifically, I designed two constructs, C73 and C74, in which a BRIL was fused to the ICL1 and ICL2 regions respectively, utilizing AlphaFold predictions. Similarly, these two constructs also exhibited findings comparable to C61 (Figure 30 A,B). As a succession, I considered changing the expression system to mammalian cells (HEK293) for testing

FAB complexation by utilizing an ICL3 BRIL construct, C75. However, this approach yielded comparable outcomes as previously described (Figure 30C).

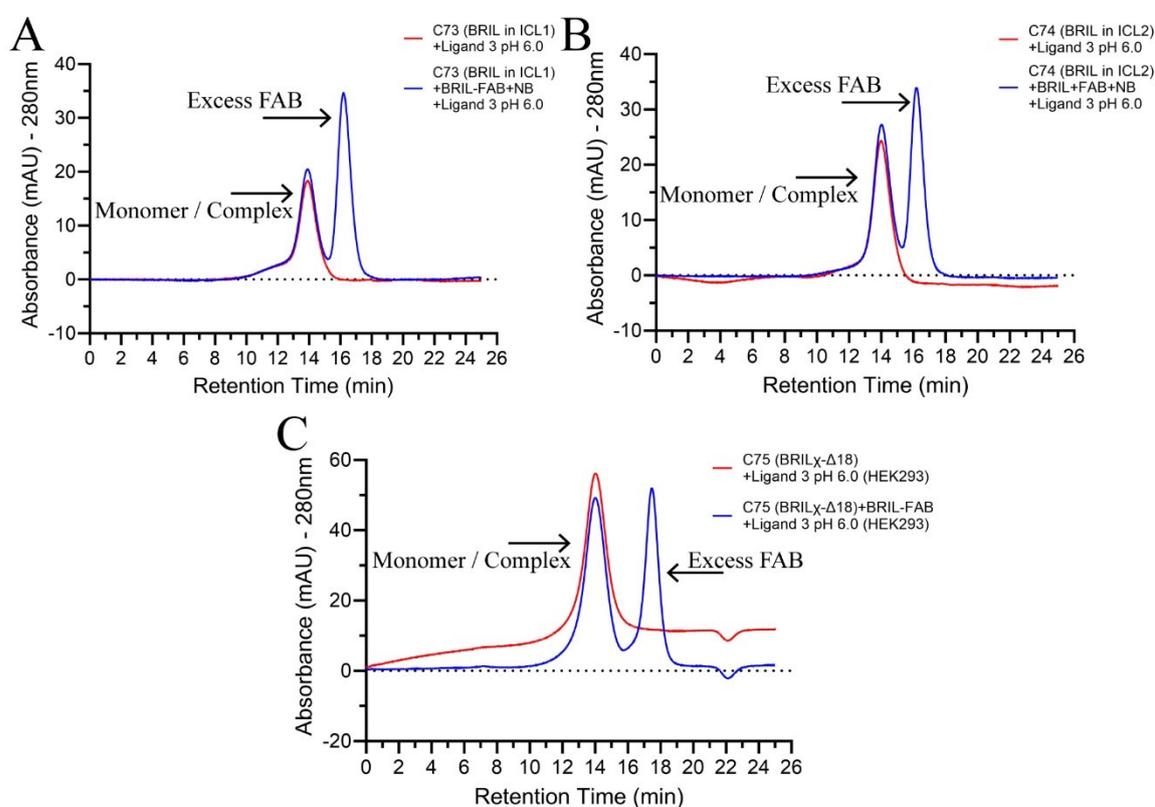


Figure 30: Screening of the GPR68 BRIL fusion constructs for complexation with Anti-BRIL FAB based on new strategies. Size exclusion chromatograms from the pooled eluates after affinity purification, FLAG and SEC for complexation with Anti-BRIL FAB (BRIL-FAB) and or nanobody (NB). The SEC was performed on Superose 6 increase 5/150 GL column for testing complexation with ICL1 BRIL fusion construct C73 (panel A), ICL2 BRIL fusion construct C74 (panel B) and ICL3 BRIL fusion construct C75, expressed in HEK293 cells (panel C). All the purification steps were performed using DDM/CHS at different pH 6.0. All purification were performed in presence of Ligand 3. The truncation denotation used in this figure is as formerly described for Figure 21.

Alternatively, a completely different construct design approach was followed after the initial failure with ICL3 BRIL based constructs, where the first strategy involved fusion of beta-lactamase from *Escherichia coli* K12 (BLEK) to the ICL3 in GPR68 and the second strategy involved using calcineurin (CalA and CalB) as a fusion protein (Figure 28C) as described before <sup>[238]</sup>. These constructs were mainly designed for the use in cryo-EM as the BLEK fusion and the calcineurin fusion would increase the size of the constructs to ~83 kDa and ~98 kDa respectively and could act as a fiducial marker.

#### 4.1.9 X-ray crystallography trials with GPR68

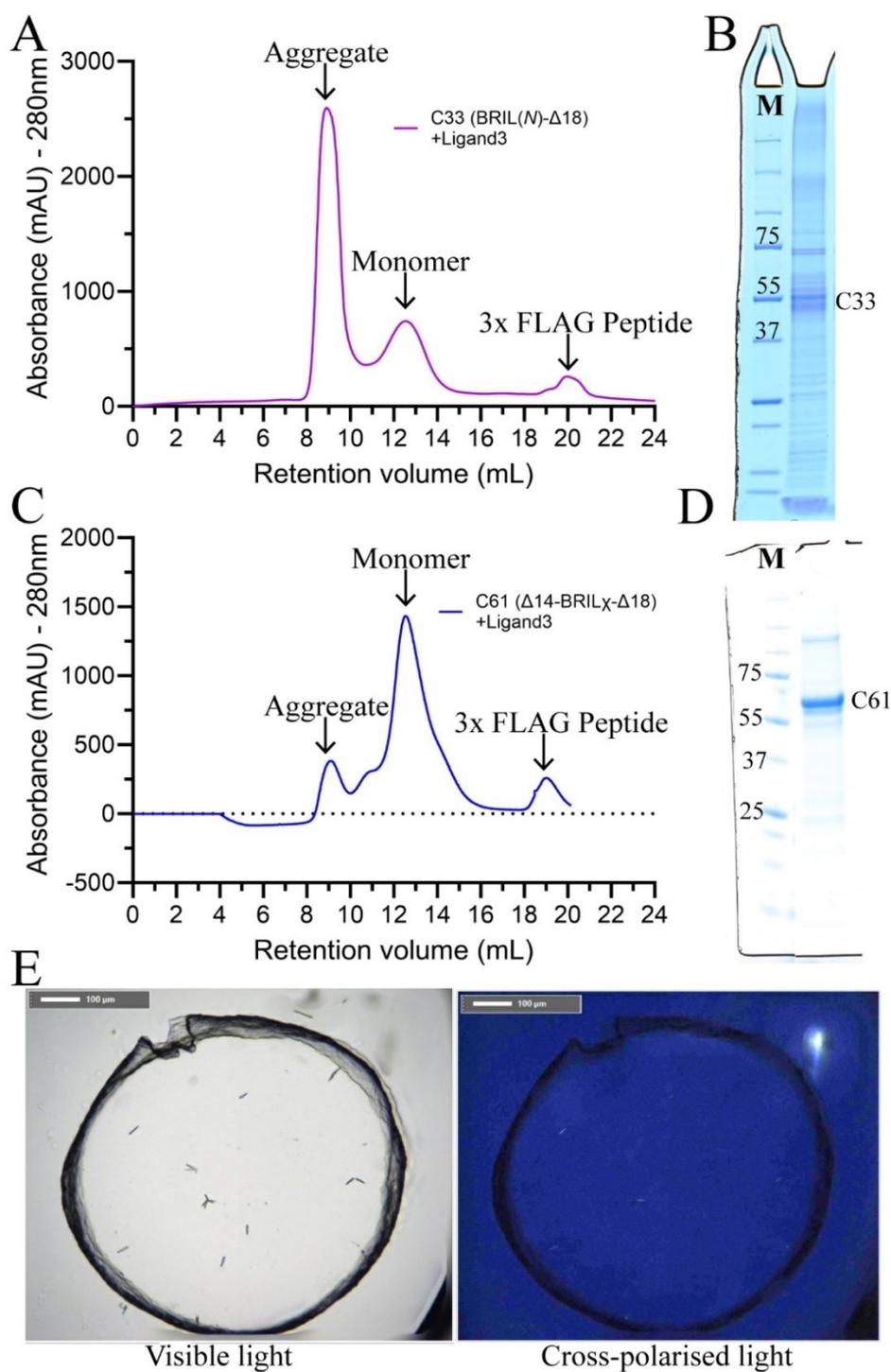


Figure 31: Summary of the main X-ray crystallography trials undertaken using human GPR68 constructs. Size exclusion chromatograms from the pooled eluates after affinity purification (FLAG) from a 2L expression of C33 purified using a Superdex 200 increase 10/300 GL column in the presence of DDM/CHS at pH 7.8 (panel A) and its respective SDS-PAGE gel (panel B). Size exclusion chromatograms from the pooled eluates after affinity purification (FLAG) from a 2L expression of C61 purified using a Superdex 200 increase 10/300 GL column in the presence of DDM/CHS at pH 6.0 (panel C) and its respective SDS-PAGE gel (panel D). The crystals from the crystallization of C61 has been shown after maximum growth under visible and cross-polarized light (panel E). All purifications were performed in the presence of Ligand 3. The

truncation denotation used in this figure is as formerly described for Figure 21. The 'M' represents the protein marker, and the numbers on the marker bands represent the corresponding molecular weight in kDa for the band.

Two human GPR68 constructs (C33 and C61), were used to attempt lipidic cubic phase crystallization to obtain protein crystals for structure determination via X-ray crystallography. The purifications of the constructs C33 and C61 yielded distinct monomeric peaks (Figure 31A,C), which were concentrated to ~32 mg/mL and 28 mg/mL respectively for setting up LCP trials. The purity of the C33 was poor (Figure 31B), and on the contrary C61 had a higher purity (Figure 31D).

After setting up a series of commercial crystallization screens, some tiny needle shaped crystals were identified when screening the C61 via LCP trials under the following conditions – 100 mM MES pH 6.5, 100 mM (NH<sub>4</sub>)<sub>2</sub>HPO<sub>4</sub> and 20% V/V PEG 400. These crystals started growing within the first 24 h and grew to the maximum size within 5 days (Figure 31E). Considering the crystallization condition, I was unsure if these were protein or salt crystals and because these crystals also glowed slightly under the cross-polarized light (Figure 31E). Thus, they were harvested for X-ray diffraction trials. Unfortunately, the X-ray diffraction did not yield a diffraction pattern.

Later, no more attempts were made for X-ray crystallography after suspicions of poor protein quality and after observing the puzzling Anti-BRIL FAB complexation behavior of C61 as described in the section 4.1.8.

#### 4.1.10 Single particle cryogenic electron microscopy trials with human GPR68

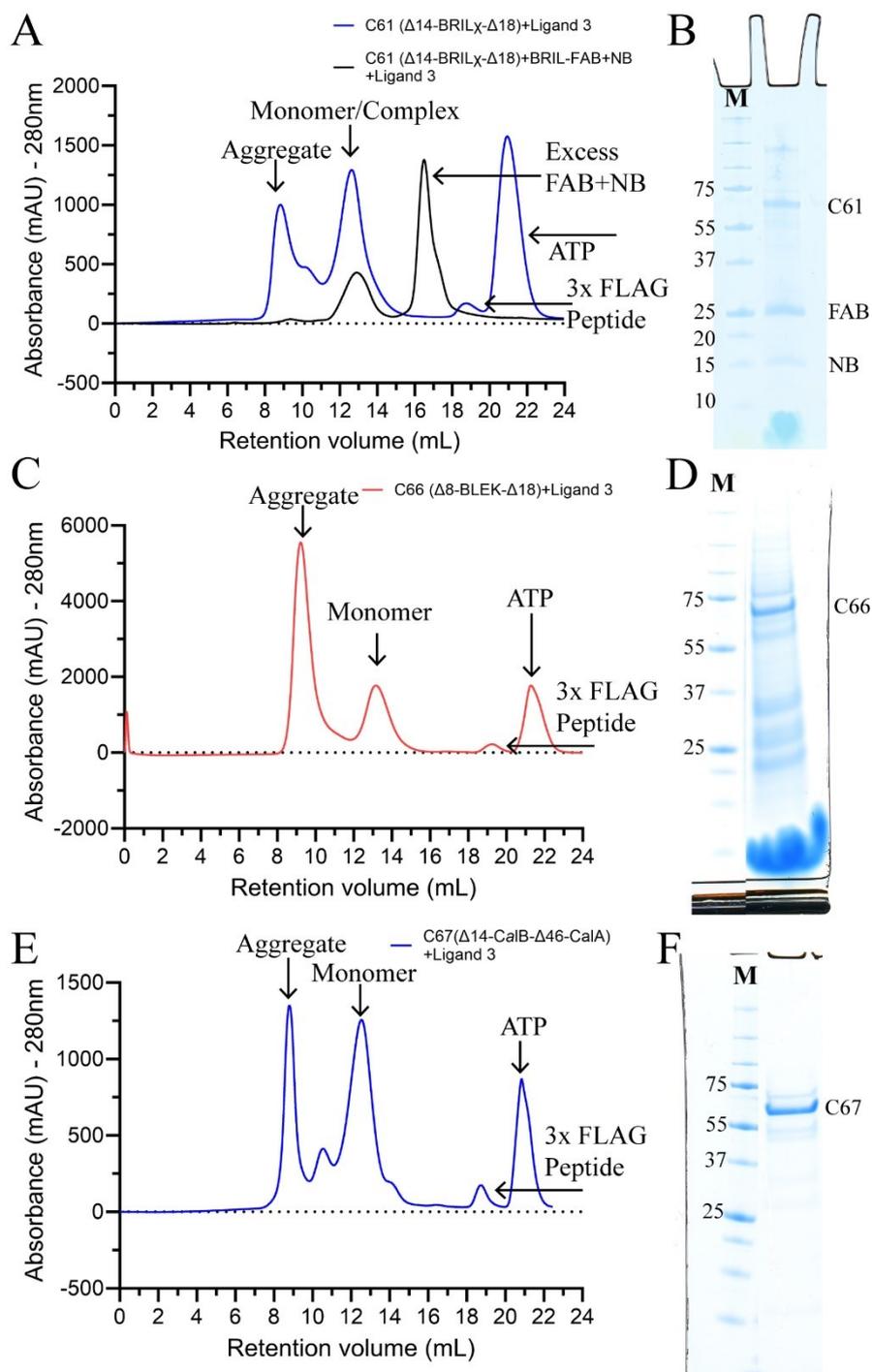


Figure 32: Summary of the representative cryo-EM trials undertaken using human GPR68 constructs. Size exclusion chromatograms from the pooled eluates after affinity purification (FLAG) from a 2L expression and SDS-PAGE gels of the final after SEC has been shown for C61 (panel A,B), C66 (panel C,D) and C67 (panel E,F). These constructs were purified using a Superdex 200 increase 10/300 GL column in the presence of DDM/CHS at pH 6.0. All purifications were performed in the presence of Ligand 3. The truncation denotation used in this figure is as formerly described for Figure 21. The 'M' represents the protein marker, and the numbers on the marker bands represent the corresponding molecular weight in kDa for the band.

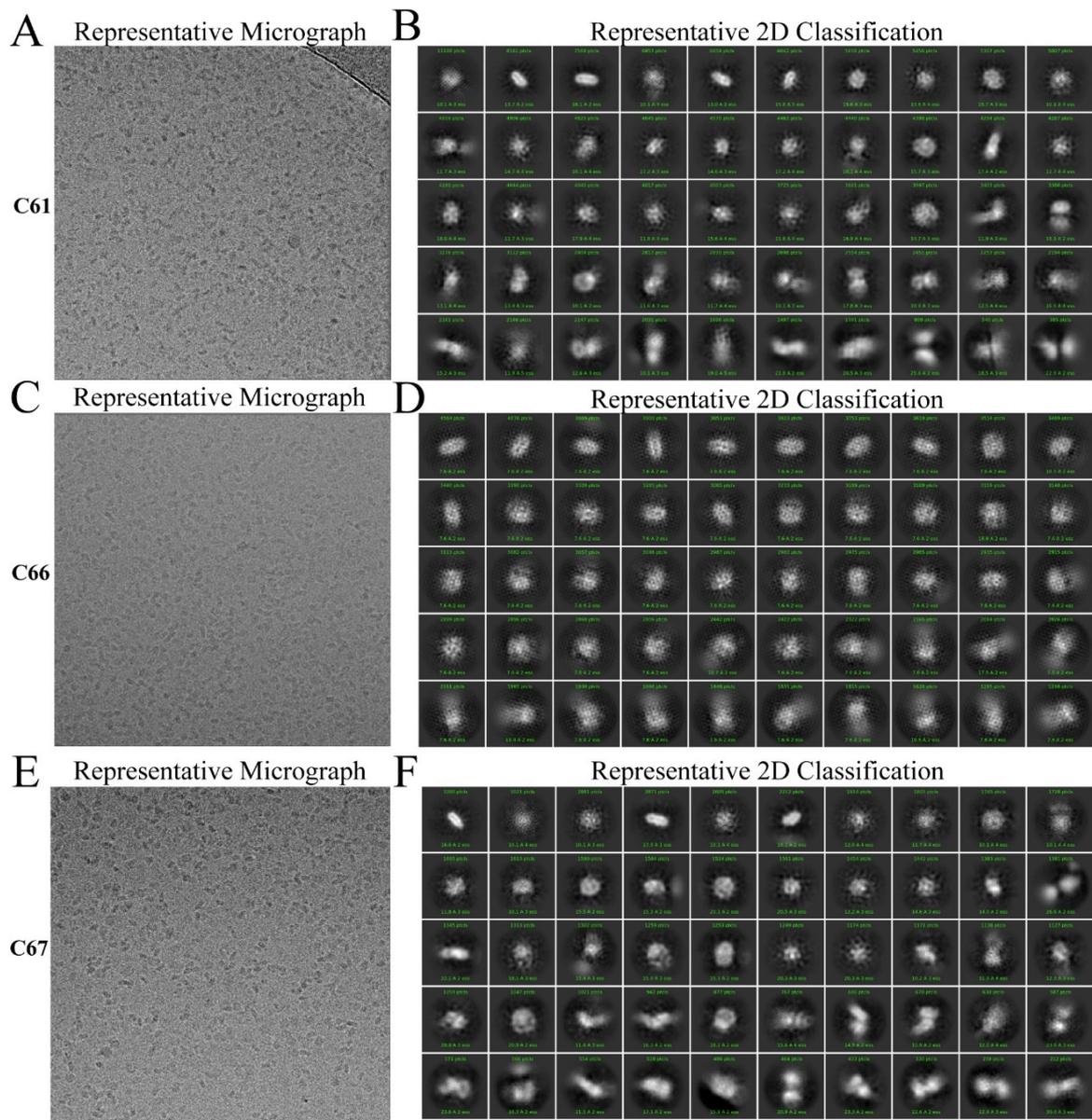


Figure 33: Summary of the representative micrographs and 2D classification of the cryo-EM trials undertaken using human GPR68 constructs – C61 (panel A,B), C66 (panel C,D) and C67 (panel E,F). Note that the representative 2D classification has been shown after a few iterations.

Even though I had described earlier in the section 4.1.8 that the Anti-BRIL FAB complexation of C61 had failed with respect to retention volume/time shift, I observed the necessary complexation partners on the SDS-PAGE gels for a similar purification (Figure 32A). The receptor was observed between the 75 kDa and 55 kDa molecular weight marker, FAB at about 25 kDa marker and NB at about 15 kDa (Figure 32B), which are approximately at the expected molecular weights for the components. Similarly large-scale purifications of C66 and C67 from 2L expression were performed for cryo-EM trials in absence of FAB and NB as these constructs had other fusions instead of BRIL. The

purifications of all the three constructs yielded distinct monomeric peaks (Figure 32A,C,E) and the C61 and C67 were pure samples (Figure 32B,F) whereas the C66 was not a pure sample (Figure 32D). Upon careful observation, I noticed that even though the molecular weights for the three constructs C61 (~54 kDa), C66 (~95 kDa) and C67 (~98 kDa) were different, still they were running at similar molecular weights on the SDS-PAGE. This further increased my suspicions of poor protein quality which is supported by the data analyzed from cryo-EM as described below.

The representative micrographs of all the three constructs showed similar observations, where ellipsoidal particles were visible with an adequate contrast along with reasonable particle distribution and uniformity (Figure 33A,C,E). Despite the decent micrographs, the 2D classification was very poor, no fiducial marker (fusion protein) was prominently observed from the detergent micelle and only fuzzy micelle was observed for all the constructs. Additional protein purification attempts, and protocol modifications were undertaken to improve 2D classification of particles; these efforts are outlined in Table 5 below.

<b>Variation Type</b>	<b>Description</b>
Constructs	C61, C66 and C67
Grid Type	Gold, copper and graphene oxide coated
Grid Mesh Dimensions	0.6/1.0 and 1.2/1.3
Detergents	DDM/CHS, LMNG/CHS and GDN
Additives	Fluorinated Fos-choline 8 and fluorinated octyl maltoside
Protein concentrations	2.5 mg/mL to 16 mg/mL
Salt concentration	Rapid sample dilution to obtain a final NaCl concentration in the range of 150mM NaCl to 500mM
Plunge Freezer Settings	Blotting force, blotting time, and humidity

Table 5: Summary of the variations attempted for the cryo-EM trials during the whole duration of my thesis.

All the variations I tested for the cryo-EM trials with the human GPR68 resulted in fuzzy micelles even after multiple iterations of 2D classification and different strategies of particle picking. Thus, all the attempts made inadequate to pursue structure determination using the human GPR68 constructs.

#### 4.1.11 Single particle cryogenic electron microscopy trials with zebrafish GPR68

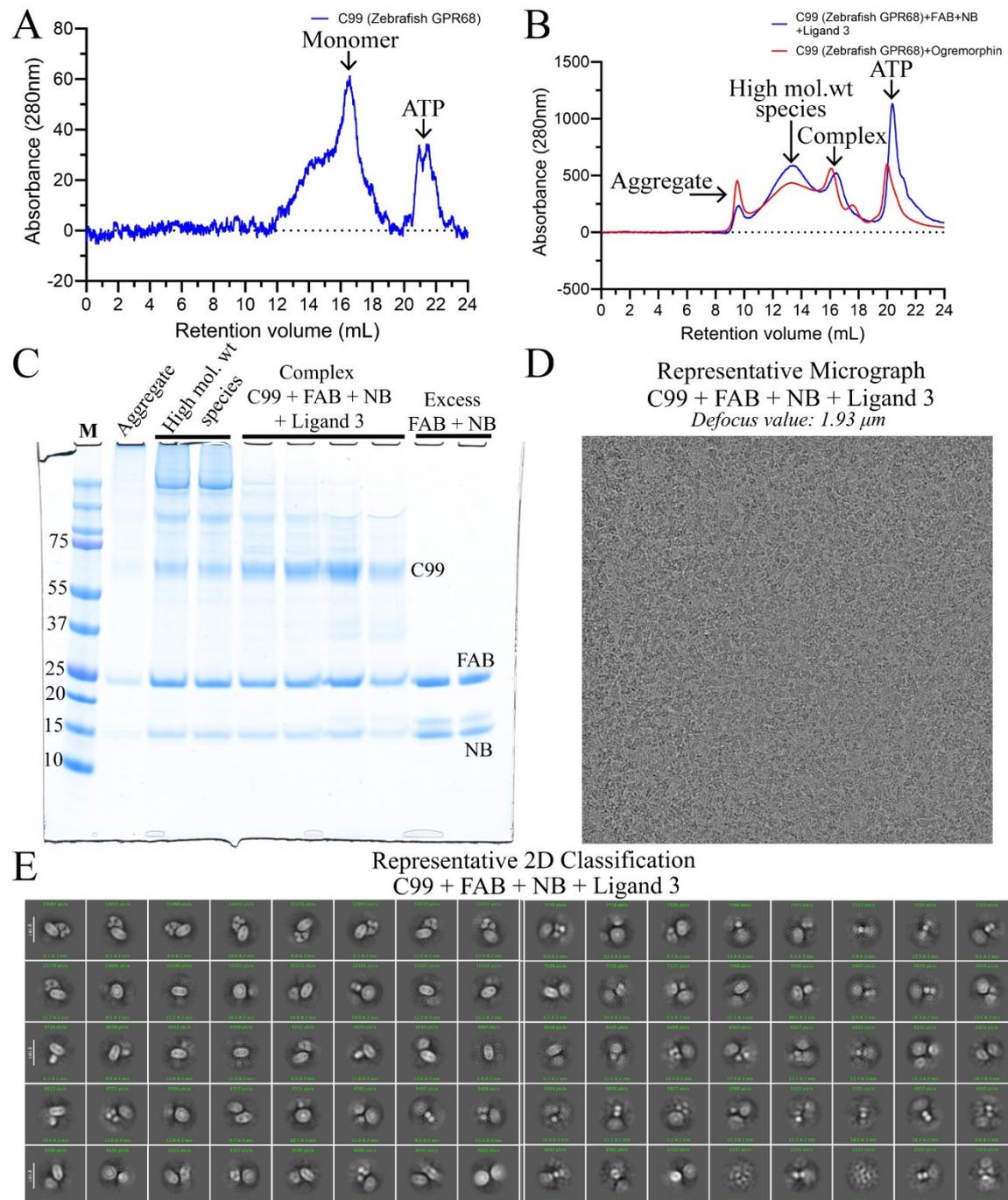


Figure 34: Summary of the representative cryo-EM trials undertaken using zebrafish GPR68 construct C99. Size exclusion chromatograms from the pooled eluates after affinity purification (Strep-tag) from a 1.5L expression and SDS-PAGE gels of the final after SEC has been shown for apoprotein C99 (panel A), and a complex of C99 with FAB and NB in presence of Ligand 3 or in the presence of Opremorphin (panel B). The representative SDS-PAGE gel (panel C), micrograph (panel D) and 2D classification (panel E) highlight that the zebrafish GPR68 C99 was complexed successfully with FAB and nanobody in the presence of Ligand 3. The constructs were purified using a Superose 6 increase 10/300 GL column in the presence of LMNG/CHS at pH 7.4. The truncation denotation used in this figure is as formerly described for Figure 21. The ‘M’ represents the protein marker, and the numbers on the marker bands represent the corresponding molecular weight in kDa for the band.

A final attempt at the inactive state structure of GPR68 was made using the zebrafish GPR68 construct (C99) was used. It was designed with an ICL3 BRIL fusion to allow for complexation with FAB and NB. The apoprotein obtained for this construct was stable but poor yields were obtained (Figure 34A), but in the presence of Ligand 3 and ogremorphin had a distinct monomeric peak (Figure 34B) and enough yields (~25  $\mu$ L with a concentration of ~7-9 mg/mL complex) were obtained to attempt cryo-EM. Cryo-EM grids for the C99+FAB+NB complex in presence of Ligand 3 were frozen at ~9.2 mg/mL. Moreover, the prepared sample was also pure (Figure 34C). The particles in the micrographs were crowded yet were uniform and homogenous with a decent contrast (Figure 34D). However, there were a lot of snake-like and circular detergent artifacts which led to poor particle picking. Nevertheless, after multiple rounds of particle filtering via 2D classification led to particles which showed the fiducial marker or the FAB and NB sticking out of the micelles. For now, the noise in the dataset remains a bottleneck to obtain the structure of the zebrafish GPR68 in presence of Ligand 3.

## 4.2 Molecular Dynamics Simulations

### 4.2.1 Sequence and structural features of GPR68

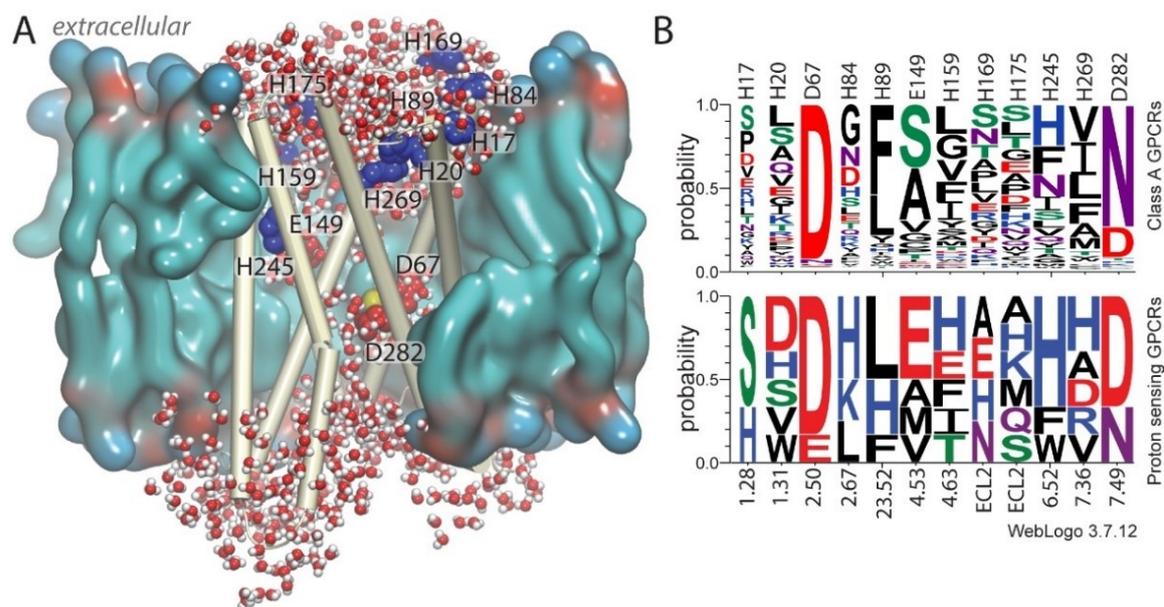


Figure 35: Distinct characteristics of human GPR68 based on sequence and structural model. (A) Cut-away view of membrane-embedded GPR68, based on the standard protonation simulation. Lipid and water molecules shown are within 10 Å and 3 Å of the protein respectively. Titratable residue sidechains thought important for proton sensing are shown. (B) Sequence conservation of titratable residue sidechains thought important for proton sensing in GPR68 with other class A GPCR members and with other proposed proton sensing GPCRs (GPR4, GPR65, GPR132, GPR31 and GPR151). The amino acid residue in GPR68 is shown at the top of the panel, while the corresponding number in the BW scheme is shown at the bottom of the panel, if it exists. The height of each amino acid residue symbol indicates its conservation among the class A GPCR sequences aligned. The thinner one-letter symbols for positions 1.28, 1.31 and 2.67 indicate gaps in some of the aligned GPCR sequences. The sequence alignment was based on GPCRdb <sup>[51]</sup> and the sequence conservation calculations were done by WebLogo server <sup>[369]</sup>. The figure has been reused and adapted <sup>[339]</sup>.

The extracellular and transmembrane histidines proposed for the pH sensing function of GPR68 are not highly conserved residue positions based on BW numbering scheme within class A GPCRs, except for the transmembrane histidine, H245<sup>6,52</sup> which is about 30% conserved (Figure 35B). Although, when compared to other proton sensing GPCRs (GPR4, GPR65, GPR132, GPR31 and GPR151), most positions have a histidine or a high probability of some other polar residue which could contribute to hydrogen bonding (Figure 35B). Single histidine mutants of GPR68 do not completely abolish receptor function and witness partial restoration of the function at more acidic pH <sup>[112]</sup>. This suggests that the proton sensitivity is perhaps less dependent on extracellular histidines, which was later supported by a study that highlighted the importance of buried acidic triad (D67<sup>2,50</sup>, E149<sup>4,53</sup> and D282<sup>7,49</sup>) in the proton sensing function of GPR68 <sup>[134]</sup>.

Evolutionary analysis based on number and location of extracellular histidines and buried acidic triad showed that GPR4, GPR68 and GPR65 are the closest relatives <sup>[134]</sup>. GPR4 is closest homolog to GPR68 with about 39% sequence identity (Table 6). The buried acidic triad residues are highly conserved within proton sensing GPCRs, except for E4.53 position (Figure 35B) which is not conserved in the distant relatives (GPR132, GPR31 and GPR151) which have limited data on their proton sensing function. Note that D67<sup>2.50</sup> forms a part of the allosteric sodium binding pocket and D282<sup>7.49</sup> forms a part of the DPxxY motif.

<b>GPCR</b>	<b>Sequence identity (%)</b>	<b>Sequence similarity (%)</b>
<b>GPR68</b>	-	-
GPR4	39	48
GPR65	27	43
GPR132	21	38
GPR31	19	34
GPR151	12	24

Table 6: Sequence identity and similarity between proton sensing GPCRs. The values were compiled based on the information from GPCRdb. The table has been reused <sup>[339]</sup>.

GPR68 consists of all the conserved class A GPCR motifs (1.1.3 and Figure 6), but with variation. Instead of the typical CWxP motif, PIF motif and NPxxY motif, GPR68 has CFxP motif, PVF motif and DPxxY motif (Figure 36A). Such variations are not uncommon in class A GPCRs, where other highly similar residues participate in the conserved motif formation. Based on GPCRdb alignment and conservation analysis of class A GPCRs, it was observed that for the CW/FxP motif, the F6.48 has 16% conservation, whereas W6.48 has 62% conservation. Similarly, for the PI/VF motif, the V3.40 has 22% conservation, whereas I.340 has 36% conservation. Additionally, in the N/DPxxY motif, the D7.49 has 18% and N7.49 has 79% conservation.

In addition to extracellular histidines, GPR68 also has other titratable residues like the surface aspartates and glutamates. There are a total of 6 aspartates and 6 glutamates located at the extracellular loops or tips of helices of GPR68 (Figure 36B), which may play a role in proton capturing as these residues are negatively charged. Furthermore, another buried residue E103<sup>3.34</sup> proximal to the buried acidic triad was observed in the core of GPR68 (Figure 36B). Comparison of the residue position, 3.34, with other class A GPCRs revealed

that 78% of the receptors have a hydrophobic amino acid, with about only 1% having a negatively charged residue, e.g., GPR68 – E103<sup>3.34</sup>. Interestingly, the other proton sensing GPCRs do not have a negatively charged residue at this position. Instead, GPR4, GPR65, GPR132, GPR31 and GPR151 have a T3.34, M3.34, C3.34, L3.34 and T3.34 respectively. Thus, the buried residue, E103<sup>3.34</sup> and surface aspartates and glutamates were investigated closely in the molecular dynamics study of GPR68 in my thesis along with the previously described regions which are crucial for the proton sensing function.

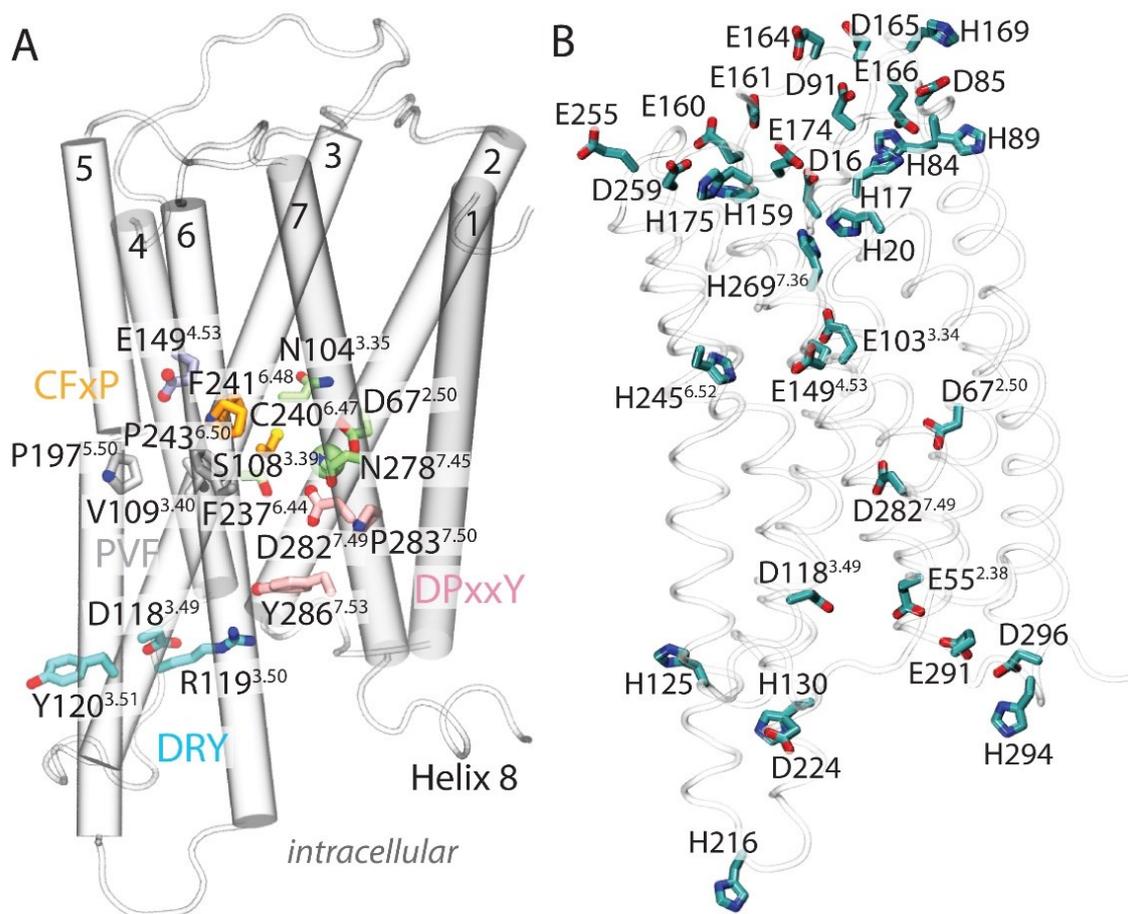


Figure 36: The position of the conserved functional motifs and the His, Asp, and Glu sidechains of GPR68 are illustrated. (A) The amino acid residue sidechains of the conserved class A GPCR motifs in GPR68. The sidechains of the residues in each motif are distinctly colored – The CFXP toggle switch (orange), PVF motif (gray), sodium pocket (green), where the Na<sup>+</sup> is shown as a green sphere, DPxxY motif (pink) and DRY motif (cyan). (B) Spatial positioning of the His, Asp, and Glu sidechains in GPR68. The figure has been reused [339].

## 4.2.2 Structural model validation

There were 8 models generated for the investigation of GPR68 via MD simulations in my thesis, which are described in the section 3.2.7 of this thesis. The in-detail model validation process has been covered in my previously published work <sup>[339]</sup>.

This section summarizes only the details of the most optimal structural model of GPR68 which was further used to perform specific protonation simulations.

Model#	RMSD (Å)	Observations
1	2.6 ±0.2	Y205 <sup>5.58</sup> faces the membrane instead of the interior of the protein.
2	2.5 ±0.2	Y205 <sup>5.58</sup> faces the membrane, PVF motif distorted, helix 8 unstructured.
3	2.4 ±0.2	Y205 <sup>5.58</sup> faces the membrane, TM3 partially unstructured, three sodium ions enter the receptor.
4	3.0 ±0.3	TM4 partially unstructured, one sodium ion at the binding pocket, two more sodium ions visit, PVF motif distorted.
5	2.3 ±0.1	Good structural stability, one sodium ion bound, correct orientation of Y205 <sup>5.58</sup> .
6	2.3 ±0.2	2 sodium ions interact with D67 <sup>2.50</sup> and D282 <sup>7.49</sup> , TM4 partially unstructured.
7	2.1 ±0.2	Y205 <sup>5.58</sup> faces the membrane. TM7 partially unstructured, helix 8 tilted, D282 <sup>7.49</sup> interacts with R301.
8	1.6 ±0.2	Y205 <sup>5.58</sup> faces the membrane, D282 <sup>7.49</sup> interacts with R301 (Helix 8).

Table 7: Summary of structural parameters used to evaluate the conformational dynamics of the GPR68 homology models with standard protonation states for all titratable sidechains. We report, for each model, the average RMSD values of the TM helices computed every 10 ps from the last 600 ns of each simulation. The table has been reused <sup>[339]</sup>.

Briefly, Model #5 had good overall stability with an rmsd value of 2.3 ±0.1 Å (Table 7) with an  $\alpha$ -helical content of ~80% (Figure S1) throughout the simulation, inter-helical distances overall consistent with corresponding values reported for static structure of class A GPCRs, correct orientation of Y205<sup>5.58</sup> and conserved class A GPCR motifs, and a stably bound single sodium ion that enters the protein from the bulk and then it remains bound at the sodium pocket throughout the entire simulation (Figure 37). The quantity of water molecules within the membrane embedded helical region achieves a stable value at approximately 100 water molecules (Figure S3). The results discussed here demonstrated

that the MD simulation of Model #5 had reached convergence. Furthermore, Model #5 aligns with the established structural parameters of class A GPCRs, making it a suitable choice for further exploration with select residue protonation simulation. For simplicity, I refer to Model #5 as GPR68.

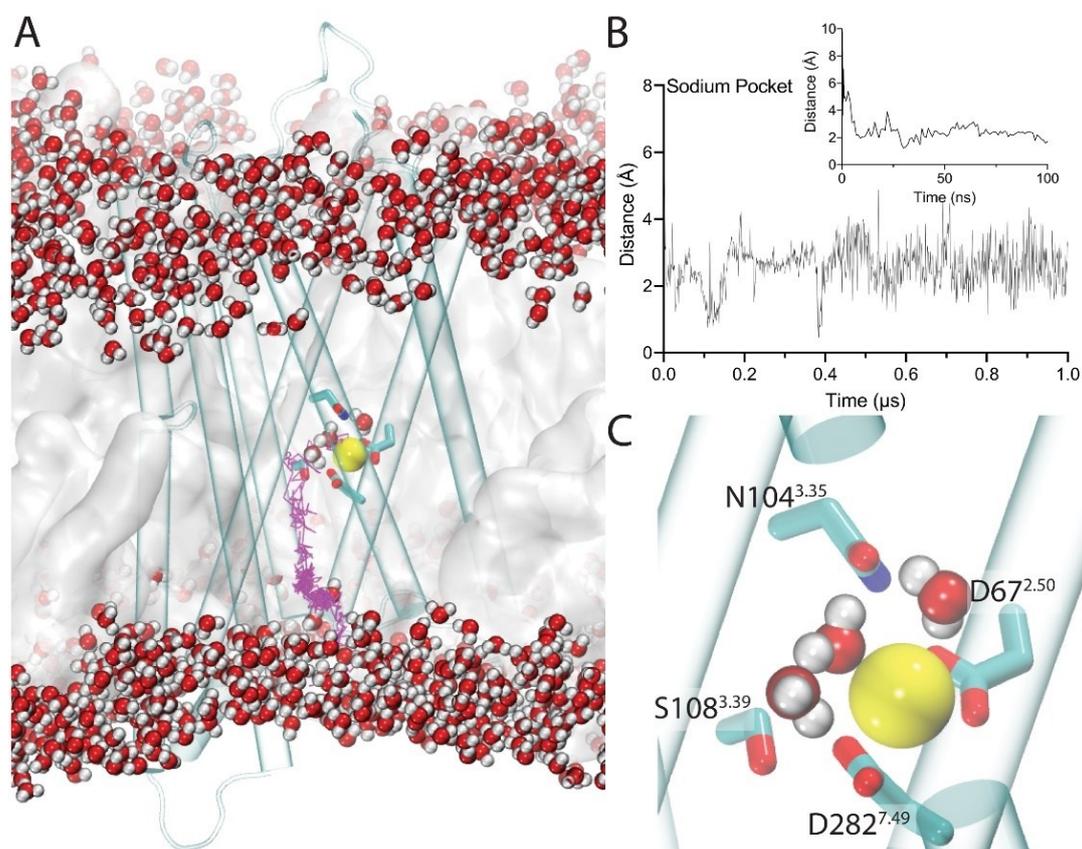


Figure 37: Dynamics of the sodium entry in the sodium-binding pocket of GPR68 with standard protonation states (Sim #5). (A) The trajectory of the sodium ion entry in the sodium pocket of GPR68 Model #5 is shown in purple. Internal water molecules are shown within 3.5 Å of the Na<sup>+</sup> (yellow sphere). Time series of the sodium ion entry to the sodium pocket, with the inset figure highlighting the first 100 ns of the trajectory. The distance (Å) reported is between the geometric center of the Ca coordinates of D67<sup>2.50</sup>, N104<sup>3.35</sup>, S108<sup>3.39</sup>, and D282<sup>7.49</sup>. (C) Close-up snapshot of the allosteric sodium ion pocket illustrating the waters and amino acid sidechains near the sodium ion. The figure has been reused <sup>[339]</sup>.

#### 4.2.3 Insights from standard protonation simulation

GPR68 has a dynamic and extensive water mediated H-bond network that inter-connects the surface histidine and carboxylate clusters to the internal carboxylic triad via other intermediate polar side chains (Figure 38A). In the last 600 ns of the standard protonation simulation of GPR68, about 25 direct H-bonds and 150 water mediated H-bonds consisting of up to 3 water molecules were found between residue sidechains (Figure 38C). Most of these water wires and direct H-bond are very dynamic, with occupancies <10% (Figure

38B). The water molecules that interact with the membrane embedded inter-helical region of GPR68 can engage in dynamic hydrogen bond networks that connect distant regions of the receptor.

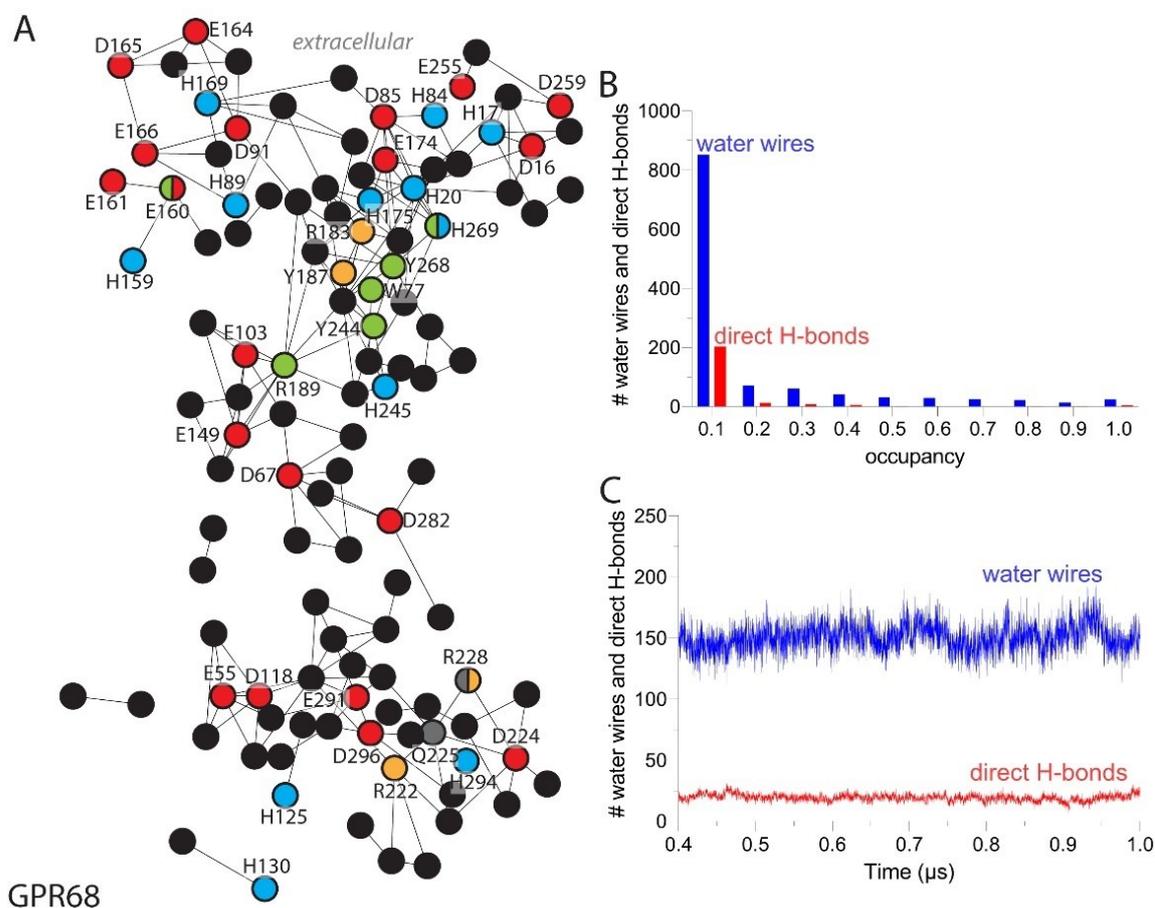


Figure 38: Dynamic water mediated protein-water H-bond network of GPR68. (A) H-bond network of GPR68. The graph nodes are assigned colors to enhance clarity and visualization – Asp/Glu sidechains (red) His sidechains (blue), amino acid sidechains that form the suggested positive allosteric binding pocket (green) and other H-bonding sidechains (black). Note that H269<sup>7,36</sup> and E160<sup>4,64</sup> are also a part of the allosteric binding site and are thus shown in two colors. For clarity, only selected nodes are labeled. The minimum H-bond occupancy shown is 30%. (B) Histogram of the average occupancy of all H-bond connections. (C) Time series of all unique H-bond connections sampled at least once during the trajectory segment used for data analyses. The figure has been reused <sup>[339]</sup>.

Further dissection of the extensive H-bond network of GPR68 via connected component analysis of GPR68 based on E149<sup>4,53</sup>, which would comprise of that part of the H-bond network which is directly or indirectly connected to E149<sup>4,53</sup>. This yields a graph shown in Figure 39A. Analysis of the local clusters using connected component analysis shows that E149<sup>4,53</sup> is very well connected to the different regions of the receptor – to the predicted allosteric pocket (Figure 39B), other surface histidine (H175<sup>ECL2</sup>) and transmembrane histidines (H245<sup>6,52</sup>, H269<sup>7,36</sup>) (Figure 39C), the other buried carboxylates (D67<sup>2,50</sup>,

E103<sup>3,34</sup> and D282<sup>7,49</sup>) (Figure 39D), surface carboxylates of extracellular loop 1 (ECL1) (D85<sup>ECL1</sup>, D91<sup>3,22</sup>) (Figure 39E,F), and the surface carboxylates of ECL2 (E164<sup>ECL2</sup>, D165<sup>ECL2</sup>, E166<sup>ECL2</sup>, E174<sup>45,52</sup>) (Figure 39F). Additionally, the transmembrane H269<sup>7,36</sup> is well connected to the allosteric pocket (Figure 39G), to the other surface histidines (Figure 39H), and to the other surface carboxylates (Figure 39I, J).

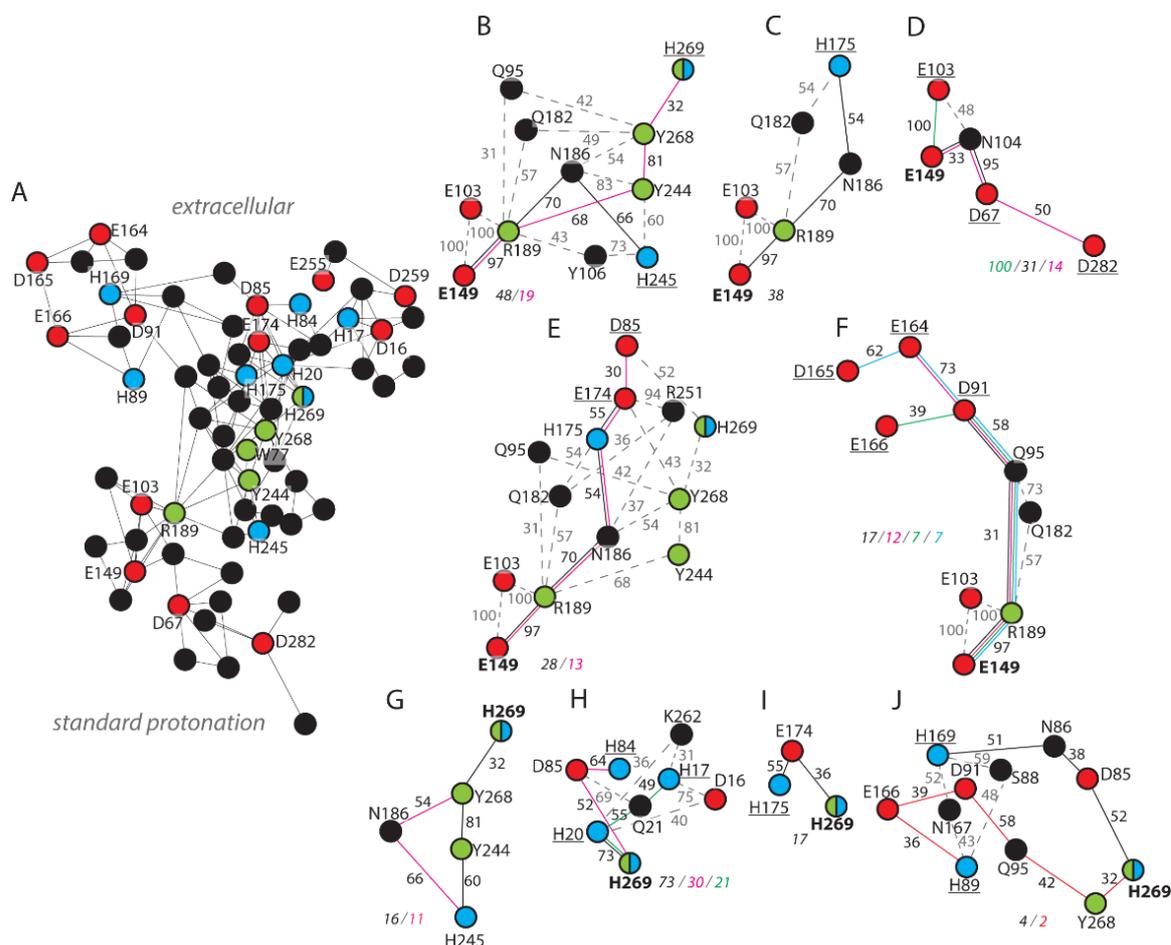


Figure 39: The water-mediated hydrogen bond network of GPR68. The graph computation and color codes are identical to those described in detail for Figure 38. Only specific nodes are labeled for the purpose of enhancing clarity. (A) The hydrogen bond cluster of E149 extracted from larger network of GPR68. (B-J) Local hydrogen bond clusters that connect specific histidine, aspartic acid, and glutamic acid side chains, where the end nodes are highlighted in bold. The color-coded numbers in italics represent the joint occupancy (JO) values of the relevant hydrogen bond pathways. The figure has been reused <sup>[339]</sup>.

I have referred to these regions of proteins as ‘well-connected’ based on significant joint occupancy values. The connections from E149<sup>4,53</sup> to different regions of the protein were sampled very frequently as can be seen from the joint occupancy values ranging from 7-48% (Figure 39B-F, Table S1). These values are quite significant especially H-bond paths with end node connections such as E149<sup>4,53</sup>-D85<sup>ECL1</sup>, E149<sup>4,53</sup>-H269<sup>7,36</sup>, and E149<sup>4,53</sup>-E174<sup>45,52</sup> which have simultaneous sampling of the intermediate polar groups over distance

of 18-28Å. These connections often have 3-5 intermediate polar sidechains as connections which may collectively have about 10-15 water molecules in the long-distance water wires (Figure S5). Certain connections in the H-bond network are dynamic and are sampled <50% of the time, but certain connections like E149<sup>4.53</sup>-E103<sup>3.34</sup>, E103<sup>3.34</sup>-R189<sup>5.42</sup> are sampled almost 100% of the time.

In summary, the connections observed for the hydrogen bonds between two residues were generally facilitated by 2-3 water molecules. There were only a few connections that were facilitated by a single water molecule on average (Figure S5). Considerable sampling times were observed for the dynamic long-distance connections between the extracellular histidine-carboxylic cluster and internal carboxylic groups, which were mediated through multiple polar sidechains, including crucial regions like the proposed positive allosteric modulator pocket (residues marked in green – Figure 38, 39).

#### 4.2.4 Insights from select residue protonation simulations

Protonation of the internal acidic triad impacts the H-bond connections of the extracellular histidine-carboxylic cluster. Comparison of standard protonation simulation with the protonation simulation was done based on the core network of GPR68 which would involve Class A GPCR conserved motifs, buried carboxylic groups, allosteric pocket, TM histidines, and a representative surface carboxylate (E174<sup>45.52</sup>). GPR68 with standard protonation states for all titratable sidechains lacks H-bond connections between the DRY motif and the core H-bond network (Figure 38A, 39A). By contrast, when both E149<sup>4.53</sup> and E103<sup>3.34</sup> are protonated, the core network H-bond network of GPR68 includes shortest-distance H-bond paths that inter-connect TM histidines (H245<sup>6.52</sup> and H269<sup>7.36</sup>) to sodium pocket (D67<sup>2.50</sup> and S108<sup>3.39</sup>), to DPxxY (D282<sup>7.49</sup> and Y286<sup>7.53</sup>), and to DRY motif (R119<sup>3.50</sup>). But these long-distance interactions are sampled only when E103<sup>3.34</sup> and are protonated at –Oε1 and not –Oε2 (Figure 40A-D). The preferential protonation state may be due to their neighbouring residues K148<sup>4.52</sup> and R189<sup>5.42</sup> potentially forming ion-pair interactions with E149<sup>4.53</sup> and E103<sup>3.34</sup> respectively (Figure 38A, 39A, S6).

When D282<sup>7.49</sup> is neutral, its local H-bond cluster with D67<sup>2.50</sup>, S108<sup>3.39</sup> and H245<sup>6.52</sup> connects to the extracellular side but lacks bridges to the cytoplasmic DRY motif (Figure

40E). Although, this is not the case when H269<sup>7,36</sup> is protonated and all the core areas of GPR68 are connected including the DRY motif (Figure 40F).

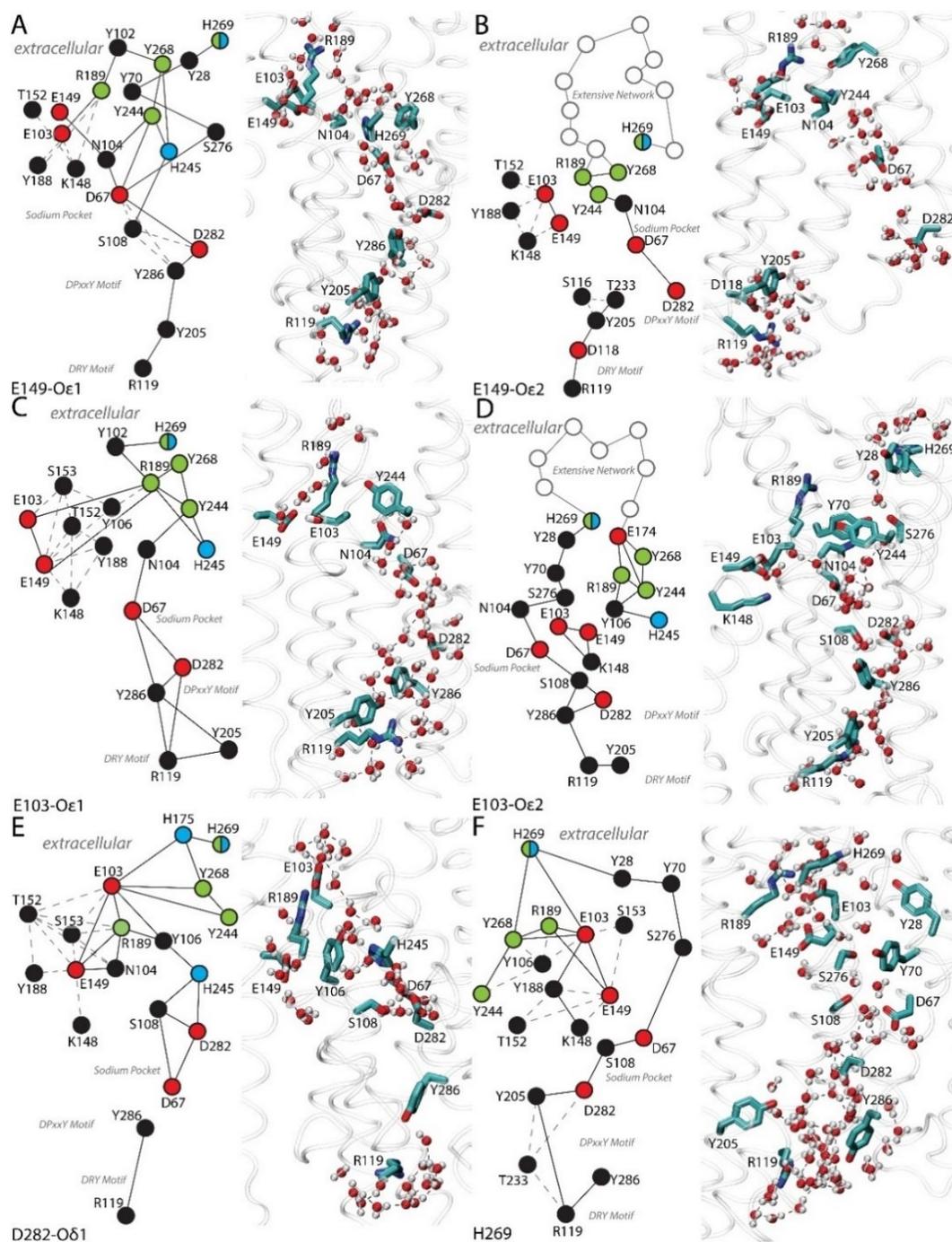


Figure 40: Protonation-coupled dynamics of the H-bond network of GPR68. (A-F) Core network of GPR68 with protonated E149-O $\epsilon$ 1 (panel A), protonated E149-O $\epsilon$ 2 (panel B), protonated E103-O $\epsilon$ 1 (panel C), protonated E103-O $\epsilon$ 2 (panel D), protonated D282-O $\delta$ 1 (panel E), and positively charged H269 (panel F). Note that these simulations lack transient sodium ion binding at the sodium-binding site. The figure has been reused <sup>[339]</sup>.

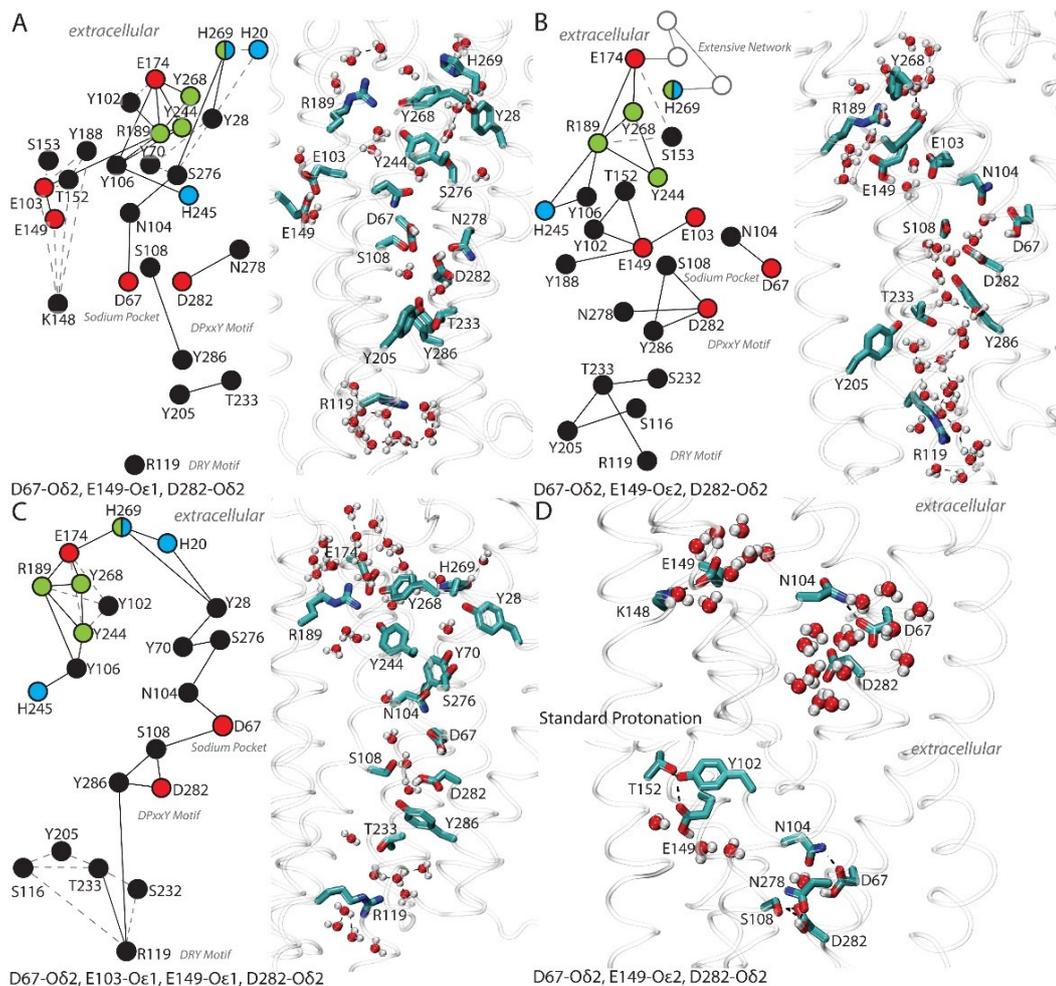


Figure 41: Protonation of all three sidechains of the internal carboxylic triad associates with perturbed local interactions. (A-D) Core protein-water H-bond network of GPR68 with protonated D67-O $\delta$ 2, E149-O $\epsilon$ 1, D282-O $\delta$ 2 (panel A), protonated D67-O $\delta$ 2, E149-O $\epsilon$ 2, D282-O $\delta$ 2 (panel B), protonated D67-O $\delta$ 2, E149-O $\epsilon$ 1, D282-O $\delta$ 2 (panel C). (D) Direct-H bonding partners and waters within 3.5 Å of the internal carboxylate triad (D67, E149, D282), when the receptor has standard protonation and when the triad is D67-O $\delta$ 2, E149-O $\epsilon$ 2, D282-O $\delta$ 2 protonated. The figure has been reused [339].

When all three carboxylic sidechains of the internal triad are protonated (D67<sup>2.50</sup>, E149<sup>4.53</sup>, D282<sup>7.49</sup>), with different protonation state for E149<sup>4.53</sup> i.e. -O $\epsilon$ 1 or -O $\epsilon$ 2, the core network of GPR68 is broken into small local clusters around DRY motif, DPxxY motif, and sodium pocket, which suggests that such protonation state appears incompatible with extensive local connections involving all three carboxylic sidechains (Figure 41A,B). A long-distance H-bond path extends from the DRY motif to the extracellular histidine-carboxyl cluster via DPxxY motif and sodium pocket, when E103<sup>3.34</sup> is protonated in conjunction with the D67<sup>2.50</sup>, E149<sup>4.53</sup> and D282<sup>7.49</sup>. Although, E103<sup>3.34</sup> and E149<sup>4.53</sup> are not a part of the common core H-bond network (Figure 41C).

## 4.2.5 Insights into crucial residues orientations

Previous studies suggested that the residue H245<sup>6.52</sup> faces the lipid core and thus is likely to be involved in the structural integrity of the receptor [112]. To validate whether this inference is influenced by the choice of template utilized. I created a homology model based on the same study using bovine rhodopsin structure as a template and observed that in this model H245<sup>6.52</sup> is positioned towards the membrane (Figure 42A). Although, the best compatible model (Model #5) from my work shows that the H245 is positioned towards the buried core of the receptor. On further investigations of the GPR68 model based on the bovine rhodopsin structure the residues W77<sup>2.60</sup>, Y205<sup>5.58</sup>, and Y268<sup>7.35</sup> are positioned towards the membrane rather than the interior of the receptor (Figure 42A). This conclusion is supported by docking and site-directed mutagenesis studies for W77<sup>2.60</sup> and Y268<sup>7.35</sup> [175], as well as an analysis of previously solved Class A GPCR structures shows that Y205<sup>5.58</sup> also faces in the protein core [96]. Furthermore, in the homology model derived from bovine rhodopsin, both E103<sup>3.34</sup> and E149<sup>4.53</sup> are positioned towards the central region of the membrane.

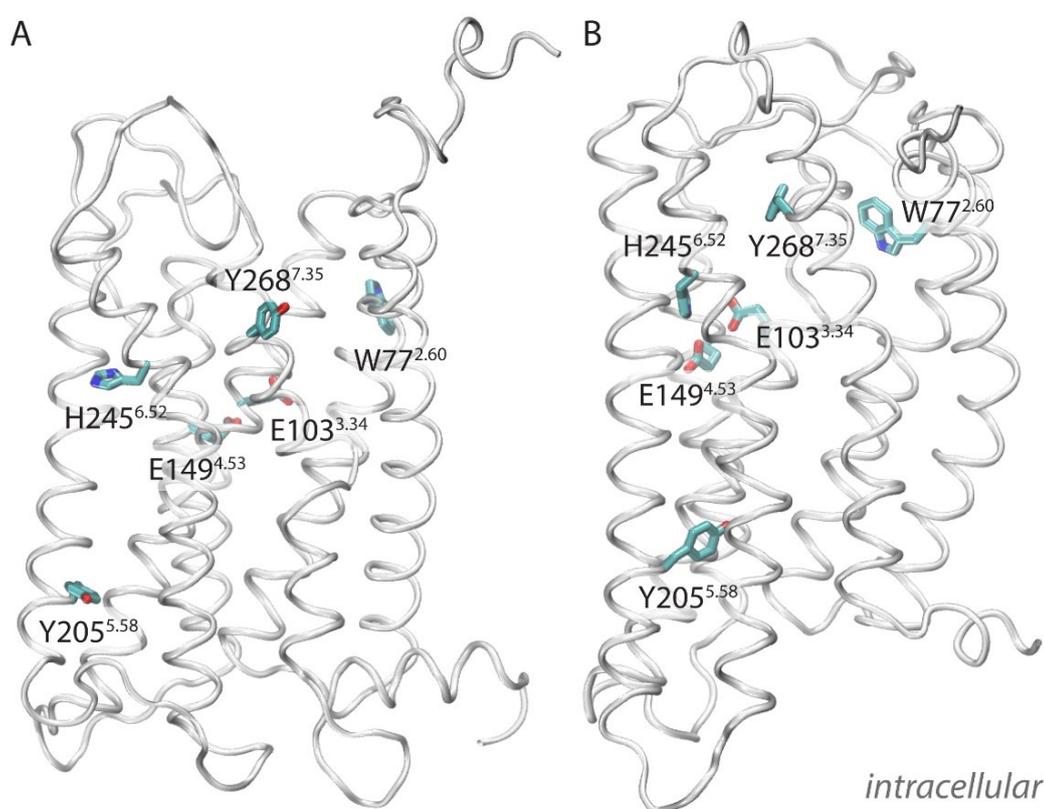


Figure 42: Orientation of crucial select residues in GPR68 models prepared from different templates (A) GPR68 Model based on Bovine Rhodopsin template (PDB ID: 1F88). (B) GPR68 Model #5. The figure has been reused and adapted [339].

## 4.3 Functional Assay

### 4.3.1 Homogenous time-resolved fluorescence assay

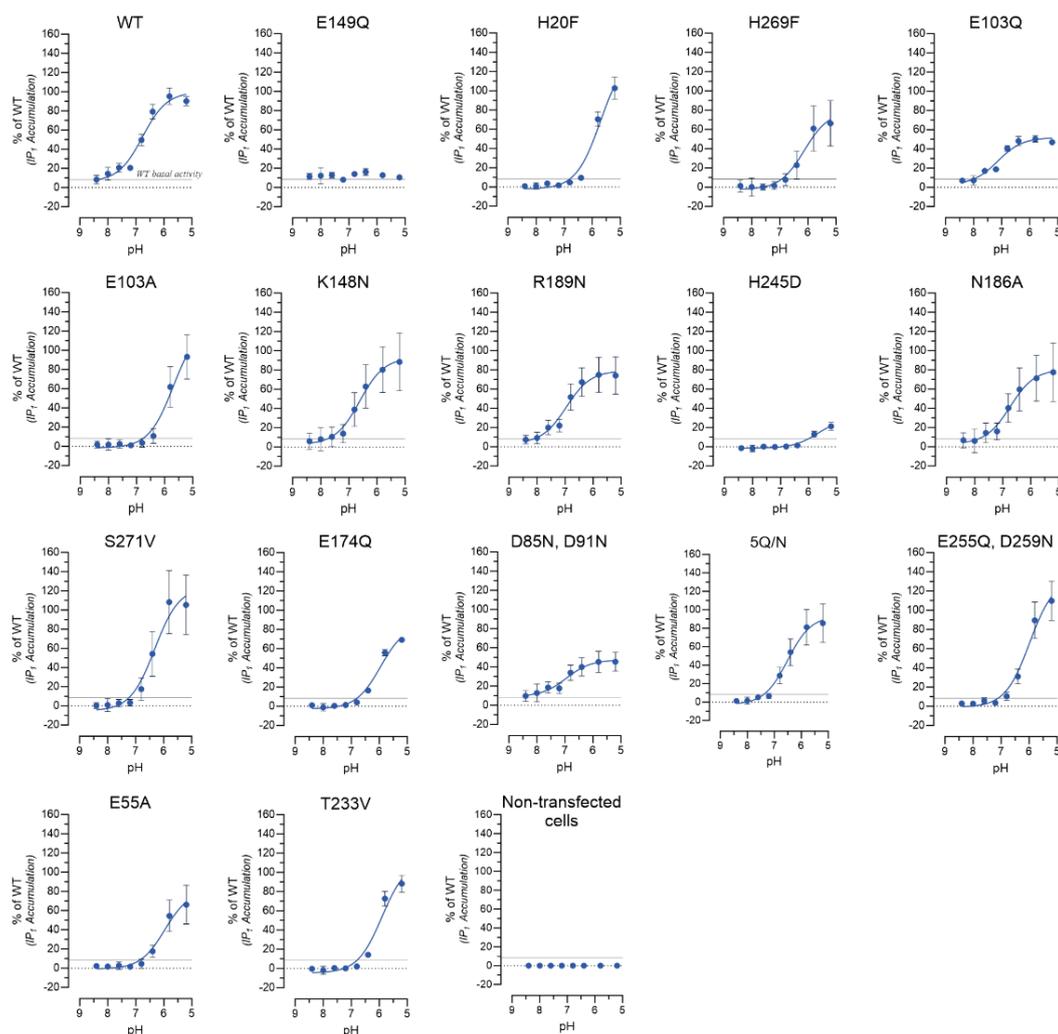


Figure 43: Homogenous time-resolved fluorescence (HTRF) assay for  $G\alpha_{q/11}$  coupling of GPR68. HTRF ratio was measured at the indicated pHs in the presence of 50 mM LiCl and was later normalized from standard curves to derive the amount of IP<sub>1</sub> accumulation. Cellular assay was performed for non-transfected cells, wild-type (WT) GPR68, and 16 mutants. Note that the mutant ‘5Q/N’ represents the mutant – E160Q, E161Q, E164Q, D165N and E166Q. The basal activity of WT GPR68 has been marked with a gray solid line. The entire dataset was normalized to the E<sub>max</sub> of WT as 100% and baseline corrected against non-transfected cells (control). The data is a collection of three independently performed experiments. The data is represented as mean±s.e.m. The pH values at which half-maximal activation is observed (pH<sub>50</sub>) are reported separately in Table 8.

Construct	pH <sub>50</sub> Values			Mean pH <sub>50</sub>	S.E.M	Mean Difference	Adjusted P Value	Summary
	Dataset 1	Dataset 2	Dataset 3					
WT	6.81	6.62	6.82	6.75	0.06	0.00	-	-
E149Q	<i>n.d.</i>	<i>n.d.</i>	<i>n.d.</i>	<i>n.d.</i>	<i>n.d.</i>	<i>n.d.</i>	<i>n.d.</i>	-
H20F	5.64	5.68	5.79	5.70	0.05	-1.05	<0.0001	****
H269F	6.15	6.14	6.05	6.11	0.03	-0.64	<0.0001	****
E103Q	7.08	7.20	7.15	7.14	0.04	0.39	0.0181	*
E103A	5.55	5.80	5.62	5.65	0.07	-1.10	<0.0001	****
K148N	6.55	6.62	6.75	6.64	0.06	-0.11	0.9775	<i>ns</i>
R189N	6.86	6.92	7.12	6.97	0.08	0.22	0.4345	<i>ns</i>
H245D	5.32	5.50	6.02	5.62	0.21	-1.13	<0.0001	****
N186A	6.64	6.67	7.08	6.79	0.14	0.04	>0.9999	<i>ns</i>
S271V	6.12	6.39	6.37	6.29	0.09	-0.46	0.0044	**
E174Q	5.92	5.99	5.97	5.96	0.02	-0.79	<0.0001	****
D85N, D91N	6.89	6.90	7.12	6.97	0.08	0.22	0.4106	<i>ns</i>
E160Q, E161Q, E164Q, D165N, E166Q	6.48	6.54	6.45	6.49	0.03	-0.26	0.2428	<i>ns</i>
E255Q, D259N	5.92	6.02	5.99	5.98	0.03	-0.77	<0.0001	****
E55A	5.95	5.95	6.02	5.97	0.03	-0.78	<0.0001	****
T233V	5.87	5.90	5.99	5.92	0.04	-0.83	<0.0001	****

Table 8: Summary of all the pH<sub>50</sub> values for WT-GPR68 and mutants tested via HTRF assay for Gα<sub>q/11</sub> signaling pathway. The significance of the mean pH<sub>50</sub> values was evaluated using one-way ANOVA analysis in GraphPad Prism 10. The words *n.d.* and *ns* stand for ‘not defined’ and ‘not significant’. The P-value scale: *ns* (P>0.05), \* (P≤0.05), \*\* (P≤0.01), \*\*\* (P≤0.001), and \*\*\*\* (P≤0.0001).

Wild-type (WT)-GPR68 showed a pH sensing range of about ~7.6 (fully inactive) to ~6.0 (fully active), with a half-maximal activation of  $6.75 \pm 0.06$  (mean±s.e.m) in the HTRF Gα<sub>q/11</sub>-coupled assay using transiently transfected HEK293H cells (Figure 43, Table 8). A basal activity of ~8.5% was observed for the WT-GPR68, when considering non-transfected cells as the baseline.

#### 4.3.2 Insights into extracellular and transmembrane histidines

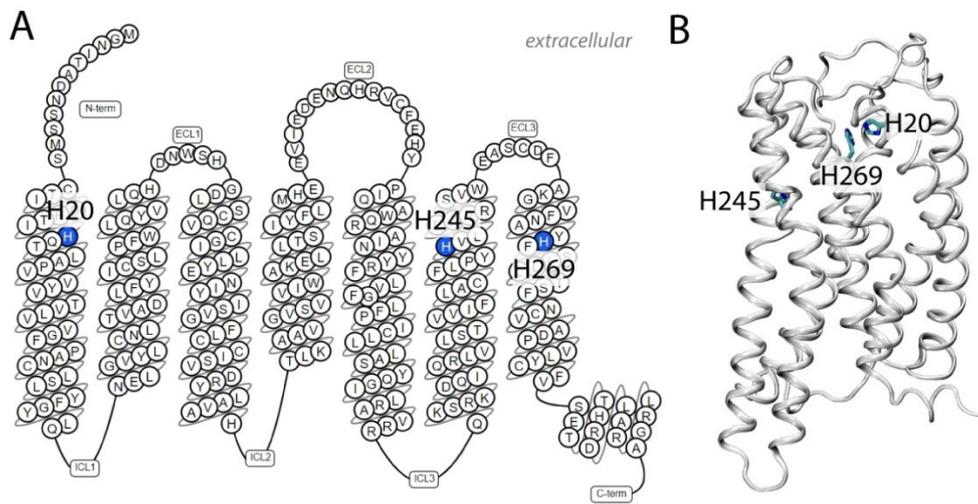


Figure 44: Extracellular and transmembrane histidines of GPR68 that were mutated in this study. (A) Snake diagram of GPR68 highlighting the mutated extracellular and transmembrane histidines. (B) Structural model of GPR68 highlighting spatial orientation of extracellular and transmembrane histidines.

The extracellular histidine mutant, H20<sup>1.31</sup>F and the transmembrane histidine mutant, H269<sup>7.36</sup>F (Figure 44) eliminated the basal activity which was observed for WT-GPR68. The E<sub>max</sub> for H20<sup>1.31</sup>F mutant was decreased by 20% of WT, whereas E<sub>max</sub> for H269<sup>7.36</sup>F

mutant was reduced by about 40% of WT (Figure 43, 44). The  $\text{pH}_{50}$  values for both the mutants reduced significantly (Figure 43, Table 8), demonstrating the requirement of more acidic pH to activate these GPR68 mutants. These mutants were primarily chosen as control mutants for comparison purposes, to compare the data from the literature with the results I obtained using the HTRF assay.

Similar results like the reduced  $E_{\text{max}}$  and a drop in  $\text{pH}_{50}$  values have also been reported previously for the same mutants when tested for  $\text{G}\alpha_{q/11}$  and  $\text{G}\alpha_s$  based signaling assays. At pH values lower than the consensus pH sensing range of GPR68, some mutant activity was observed to be restored [112, 171]. Interestingly, alanine mutant versions of the residues H20<sup>1.31</sup> and H269<sup>7.36</sup> show a lesser drop in  $\text{pH}_{50}$  values in comparison to the phenylalanine mutations, in the proton sensing function of the receptor, and were suggested to unlikely play a functional role [171]. Furthermore,  $\text{pK}_a$  calculations done via computational studies using GPR68 models suggested low  $\text{pK}_a$  values of  $\sim 6.0$  for the transmembrane histidines [370]. A more recent study reported that the mutants H20<sup>1.31</sup>D, H20<sup>1.31</sup>N, H20<sup>1.31</sup>R, H269<sup>7.36</sup>D, and H269<sup>7.36</sup>N were reported to have a decreased  $\text{pH}_{50}$  value and similar or slightly reduced efficacies in comparison to the WT except for H269<sup>7.36</sup>N which demonstrated a slight increase in  $E_{\text{max}}$  value [371]. Additionally, H269<sup>7.36</sup>K showed an increase in  $\text{pH}_{50}$  value by roughly 0.3 pH units along with a decrease of 40% in  $E_{\text{max}}$  [371]. The H245<sup>6.52</sup>F mutant previously has been reported to significantly reduce receptor activity but does not completely abolish the activity [112, 171]. H245<sup>6.52</sup> was suggested to play a structural role as it faces the lipid membrane [112]. The computational studies with GPR68 structural models from my thesis revealed that H245<sup>6.52</sup> faces the inner core of the membrane (section 4.2.5). Thus, I tested H245<sup>6.52</sup>D to check the possibility of a functional role for this buried histidine (Figure 43, Table 8). The H245<sup>6.52</sup>D was observed to significantly reduce receptor function with no basal activity and turned out to be a more drastic mutant than phenylalanine in terms of activity reduction.

### 4.3.3 Insights into buried charged residues

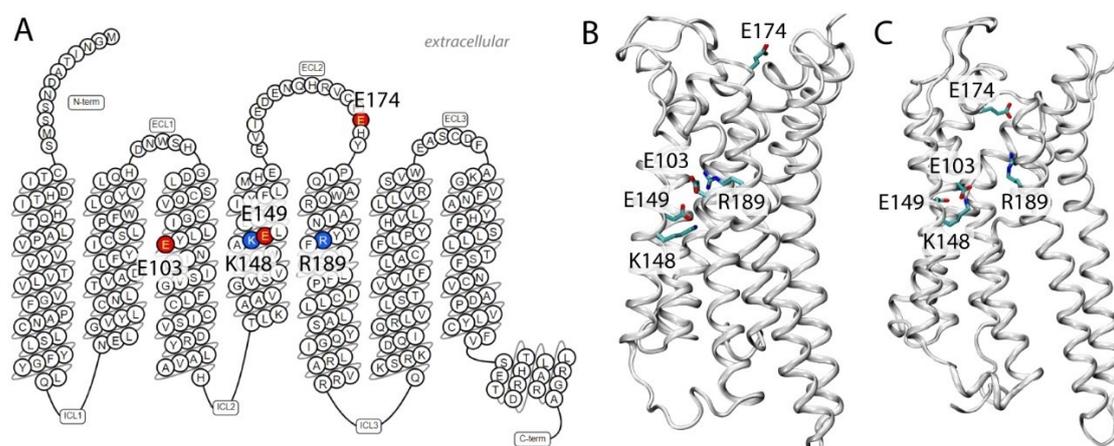


Figure 45: Buried charged residues of GPR68 that were mutated in this study. (A) Snake diagram of GPR68 highlighting the mutated buried charged residues. (B) Structural model of GPR68 highlighting spatial orientation of buried charged residues from a co-ordinate snapshot of standard protonation simulation, which highlights the outward orientation of E174<sup>45,52</sup>. (C) Structural model of GPR68 highlighting spatial orientation of buried charged residues from a co-ordinate snapshot of protonated simulation of E149-Oε1, which highlights the inward orientation of E174<sup>45,52</sup>.

The buried carboxylate E149<sup>4,53</sup>Q was observed to be a dead mutant at all pH values. These results are completely contrary to the results which have been published before, where this mutant was constitutively active when screened with GloSensor cAMP assay with its pH<sub>50</sub> values of approximately 7.4-7.6 [134, 371]. The pK<sub>a</sub> of this buried residue has been suggested to be upshifted to pK<sub>a</sub> 7.0 [134, 370].

The buried acidic triad was proposed to be a crucial element in the proton sensing function of the receptor [134]. In my work based on molecular dynamics (MD) simulations, another buried glutamate, E103<sup>3,34</sup>, was found to be a part of the common hydrogen bond network and was hypothesized to possibly play a role in the proton sensing function of the receptor (section 4.2.4). Thus, to further validate the hypothesis I tested the mutations: E103<sup>3,34</sup>A and E103<sup>3,34</sup>Q. The E103<sup>3,34</sup>A mutant had no basal activity and a reduced pH<sub>50</sub> value by ~1.1 pH like H245<sup>6,52</sup>D and H20<sup>1,31</sup>F mutant (Figure 43, Table 8). Although, this mutant regained function at more acidic pH. On the contrary, E103<sup>3,34</sup>Q mutant shows an increased pH<sub>50</sub> of 7.14±0.04, suggesting a more active mutant, although the E<sub>max</sub> was reduced by 50% (Figure 43, Table 8). In addition to buried negatively charged residues, there exist two buried positively charged residues K148<sup>4,52</sup> and R189<sup>5,42</sup> (Figure 45), which were mutated to their neutral forms: K148<sup>4,52</sup>N and R189<sup>5,42</sup>N. The pH<sub>50</sub> values of these residues is like the WT-GPR68, but their E<sub>max</sub> was decreased (Figure 43, Table 8).

### 4.3.4 Insights into surface charged residues

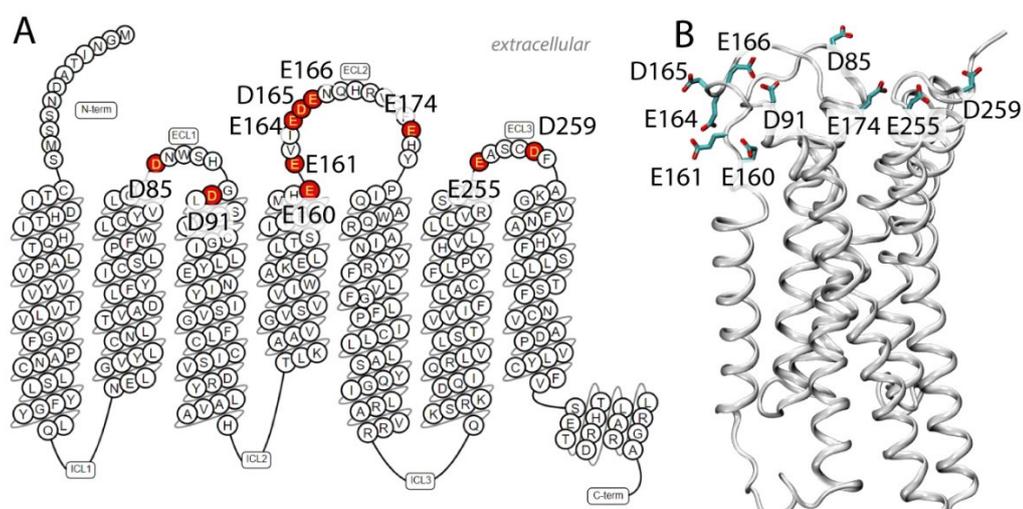


Figure 46: Surface carboxylates of GPR68 that were mutated in this study. (A) Snake diagram of GPR68 highlighting the mutated surface carboxylates. (B) Structural model of GPR68 highlighting spatial orientation of surface carboxylates.

The mutants D85<sup>ECL1</sup>N and D91<sup>3.22</sup>N (extracellular loop 1 region) (Figure 46), and E160<sup>4.64</sup>Q, E161<sup>ECL2</sup>Q, E164<sup>ECL2</sup>Q, D165<sup>ECL2</sup>N and E166<sup>ECL2</sup>Q (extracellular loop 2 region) (Figure 46) showed non-significant differences in  $pH_{50}$  values based on statistical analysis, where the former mutant showed a reduction of  $E_{max}$  by ~50% and the latter mutant by ~20% (Figure 43, Table 8). The alanine mutant of the residue D85<sup>ECL1</sup> has been reported to a dead mutant [370], showing that alanine mutation for this position is more severe than the charge neutralization mutation to glutamine.

Intriguingly, the single mutant E174<sup>45.52</sup>Q, showed a significant reduction in  $pH_{50}$  values by ~0.8 pH units and in  $E_{max}$  by ~25% (Figure 43, Table 8). The double alanine mutation of the E164<sup>ECL2</sup> and E166<sup>ECL2</sup> was reported to behave like WT-GPR68, whereas the single alanine mutant D165<sup>ECL2</sup>A had slightly decreased  $pH_{50}$  values and 50% reduced  $E_{max}$  in comparison to the WT-GPR68 [370]. Following the same trend, the triple alanine mutant of E164<sup>ECL2</sup>A, D165<sup>ECL2</sup>A and E166<sup>ECL2</sup>A showed a significant reduction in  $pH_{50}$  values and 70% reduction in  $E_{max}$  [370]. A more recent investigation of the mutants E174<sup>45.52</sup>Q and E174<sup>45.52</sup>A mutants via GloSensor cAMP assay revealed that the E174Q mutant showed a reduction in  $pH_{50}$  values by ~0.5 pH units and ~60% reduction in  $E_{max}$ , whereas the E174A mutant showed similar  $pH_{50}$  values to WT receptor with a 50% reduction  $E_{max}$  [371]. Currently, it is difficult to explain why the  $pH_{50}$  values of the alanine mutant is similar to WT, but more future studies revolving around this mutant accompanied by the structural

information of GPR68 could help solve the dilemma. Lastly, the extracellular loop 3 double mutant: E255Q, D259N showed a reduction in  $pH_{50}$  values of  $\sim 0.8$  pH units like the E174<sup>45,52</sup>Q mutant (Table 8).

#### 4.3.5 Insights into other polar residues

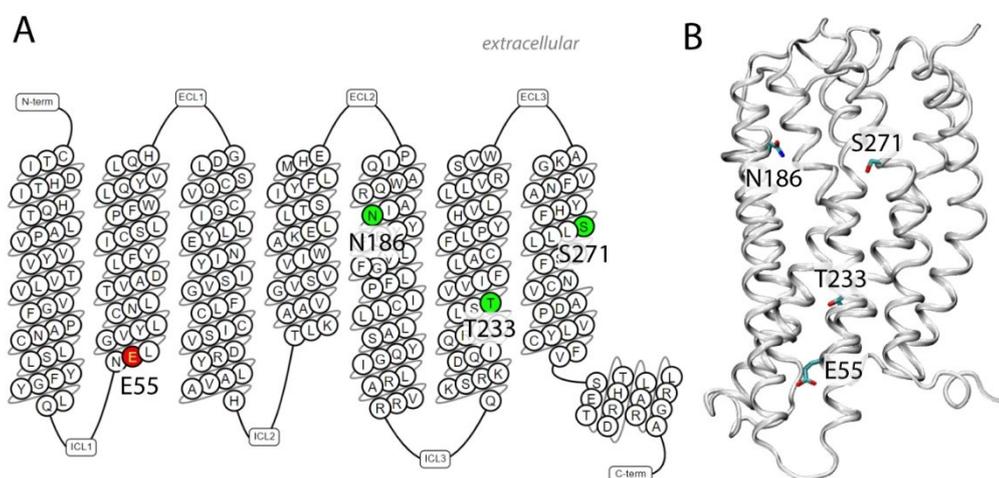


Figure 47: Other residues which were mutated in GPR68 in the study. (A) Snake diagram of GPR68 highlighting the mutated E55, N186, T233 and S271 residues. (B) Structural model of GPR68 highlighting spatial orientation of E55, N186, T233 and S271 residue.

The residue N186<sup>5,39</sup> is in the core of the hydrogen bond network of GPR68 (Figure 39). The mutant N186<sup>5,39</sup>A had similar  $pH_{50}$  values to the WT-GPR68 with a  $\sim 25\%$  decrease in  $E_{max}$  (Figure 43, Table 8). The residue position 7.38 is often involved in ligand binding of class A GPCRs [68]. The endogenous ligand for GPR68 known so far are protons, thus it is difficult to interpret whether a typical class A GPCR like orthosteric pocket exists for this proton sensing GPCR. Thus, the mutant S271<sup>7,38</sup>V was tested to check for a potential role in the pH sensing function. The S271<sup>7,38</sup>V mutant showed a reduced  $pH_{50}$  values by  $\sim 0.5$  pH units but retaining similar  $E_{max}$  to WT-GPR68 at more acidic pH (Figure 43, Table 8). Serine being a polar residue can form water mediated hydrogen bonds and might turn out to be important in the receptor function considering the relevance of this position in other class A GPCRs.

After a more thorough investigation of the GPR68 sequence I found that the position 6.40 in GPR68 contains a hydrophilic amino acid, threonine. It has been reported for class A GPCR, if the position 6.40 contains a hydrophilic amino acid it makes the receptor family more biased towards an active-like conformation increasing the receptor's constitutive activity e.g., melanocortin-4 receptor [372], CXC motif chemokine receptor 1 (CXCR1)

receptor<sup>[102]</sup> and in rhodopsin<sup>[373]</sup>. Usually, this position is hydrophobic and is a part of the commonly found hydrophobic pocket in class A GPCRs comprising of the following residue positions: 2.46, 3.43, 3.40, 3.46, 6.40, 6.41 and 6.44<sup>[374]</sup>. Thus, I mutated T233<sup>6.40</sup> back to a small hydrophobic amino acid, valine with an expectation of reducing receptor activation. The T233<sup>6.40</sup>V mutant showed no basal activity like WT-GPR68 does, and a significant reduction in pH<sub>50</sub> values of ~0.8 pH units was observed (Figure 43, Table 8).

A unique residue E55<sup>2.38</sup> located at the intracellular tip of the TM2 helix (Figure 47). At the position 2.38, 75% of the class A GPCRs possess a hydrophobic residue and the rest have a polar residue that can participate in hydrogen bonding of which only 3% have a negatively charged residue. Intriguingly, the proton sensing GPCRs (GPR68, GPR4 and GPR65), lysophosphatidic acid receptors (LPA4 and LPA6), GPR35, GPR39 and platelet activated factor (PAF) receptor have a glutamate at this position and the chemokine receptor (ACKR3/CXCR7) has an aspartate. The sequence conservation was checked using GPCRdb sequence alignment<sup>[6]</sup>. The mutant E55<sup>2.38</sup>A showed reduced pH<sub>50</sub> values by ~0.8 pH units (Table 8).

## 5. Discussion

In December 2022, the Swiss contract research organization LeadXPro announced that they had successfully determined the structure of an active state GPR65 (<https://leadxpro.com/leadxpro-has-just-elucidated-the-worlds-first-structure-of-a-proton-sensing-gpcr-target/>, as accessed on 05.12.2022), a member of the proton sensing family of receptors. Furthermore, the active state structures of other members of this family (GPR4, GPR68, and GPR65) were very recently determined as part of a publicly available doctoral thesis from University of California, San Francisco, USA <sup>[371, 375]</sup>. However, at the time of submission of this thesis the coordinates of these receptors are not publicly available yet.

Below is a summary of my efforts to study the proton sensing GPR68:

<b>Strategies Tested</b>	<b>Description</b>
Protein engineering	100 constructs Fusions, truncations, fusions and truncation combination, AlphaFold based construct design strategy
Thermostabilizing mutation constructs	~25 constructs
Purifications	~300-350 (500 mL cell pellet)
Species	Human and zebrafish
Expression systems	Insect and mammalian cells
Membrane mimetics	Detergents (Mainly - DDM/CHS, LMNG/CHS, GDN)
MD simulations	Total sampling time of 14 $\mu$ s with GPR68 structural models Structural model validation of GPR68 Specific protonation state simulations with best structural model Application of hydrogen bond network analysis using graph theory
Functional assay setup	Enabling homogenous time-resolved fluorescence assay for testing GPR68 signaling over a range of pH to validate predictions and hypotheses based on MD simulations
Functional assay mutations constructs	~16 constructs

Table 9: Overview of the experiments or simulations performed during my doctoral thesis.

## 5.1 Expression, Purification and Structure Determination Trials of GPR68

The expression and purification of human GPR68 was a challenging task as the wild type GPR68 with tags (C1) had low expression (Figure 19B). The implemented protein engineering strategies such as inserting various fusion proteins in the ICL3 region (Figure 18), utilizing different insertion points of BRIL in ICL3 (Figure 22B), truncations of the termini (Figure 19), or combining different above-mentioned strategies did not help in stabilizing the apoprotein much. Slight variations in the absorbance units from the chromatograms was observed, but mostly the average absorbance was in the range of ~10-50 mAU for all the different types of GPR68 constructs. On average, the yields for all the apoprotein purifications across different constructs were in the range of 20-60  $\mu\text{g}$  /500 mL cell pellet as measured after first affinity purification step. The yields measured after SEC were not reliable as the yields were extremely low for the monomeric protein as most of the sample was just aggregated protein. Overall, the above-described protein engineering strategies improved the protein yields, but the major improvements in total yield and monomer protein yield was observed by addition of small molecule ligands. The stabilization effects of the ligands as judged by presence of a monomeric protein peak, other peak characteristics, and yield based on the general trend from various constructs (sections 4.1.1 – 4.1.6) are mentioned in the descending order Ligand 3 > Ligand 2 >  $\text{Co}^{+2}$  = Ligand 1. The yields on average for GPR68 constructs when purified with Ligand 3 is usually in the range of ~200-300  $\mu\text{g}$  /500 mL cell pellet after size exclusion.

In the beginning of my thesis, I did not have access to the Dionex, which is a high-performance liquid chromatography (HPLC) equipment that enables high throughput sample testing. As a result, most of the purifications were carried out at a scale of 500 mL or more. The Dionex system was only available after more than half of my thesis duration had passed. Nevertheless, the other reasons to pursue purifications at 500 mL scale were poor yields of GPR68, lack of reproducibility when scaling up purifications from 50 mL to 500 mL e.g., mutation constructs (Figure 27C,D) or from 500 mL to 2L (Figure S7) and to continue comparison of newer set of purification results with the older purifications results which were performed at 500 mL scale for more coherent analysis. Direct proportional scaling up of purification protocol from 500 mL to 2L scale leads to change in overall

monomer yield and peak characteristics, which is shown by a representative GPR68 construct, C33 (Figure S7). Thus, to produce more protein 500 mL cell pellets were purified separately and were only pooled together after the eluates from the first affinity purification step were concentrated and then subsequently injected onto an appropriate column for SEC.

The strategies described in the sections 4.1.2, 4.1.3 and 4.1.4 were initially tested at pH 7.8 in the absence of a ligand. Following the screening of many constructs with ligands, an optimal least modified construct, C32, was identified. It was observed that it is not possible to stabilize C32 by pH alone (tested at pH 6.0 and pH 7.8) (Figure S8) and therefore other ligands were required for enhanced stability (Figure 26F). The initial thought was that by altering pH I can introduce a conformation bias in the receptor (towards active conformation to pursue active state structure route), as protons are considered as endogenous ligands for GPR68. Such a change is not observed as the receptor is in the membrane mimetic or non-native environment when extracted and purified using detergent. Exposure of membrane proteins to detergents usually leads to removal of endogenous lipids. Initially, sphingosylphosphorylcholine was proposed as an endogenous lipid ligand for GPR68, but the paper was later retracted <sup>[138]</sup>. Along with the claims of determining an active state structure of GPR68 recently, it was reported that there perhaps exists an extracellular pocket rich in polar and charged residues in GPR68 and other proton sensing GPCRs which could possibly accommodate endogenous molecules like lipids <sup>[375]</sup>. As previously discussed in the section 1.2.2 of my thesis, sensitivity to protons may be a co-incidence detection feature of proton sensing GPCRs and other class A GPCRs like adenosine A<sub>2a</sub> can also be activated by changes in pH <sup>[145]</sup>.

The initial results for C10 (Figure 21A) showed a distinct monomeric protein peak which was not reproducible. The metal ion leaking from the TALON resin due to the resin being old was verified by purification the construct again with fresh batch of TALON resin, FLAG resin and purification in presence of EDTA in combination with TALON resin (Figure 24A). Although the interesting part is that even after adding 100 μM of cobalt chloride and purification via FLAG M2 resin I was not able to reproduce my initial chromatogram (Figure 21A). The divalent metal ion modulation of GPR68 was not a mystery at the time of purification, as it was already reported in the literature. Multiple divalent metal ions (Cd<sup>2+</sup>, Co<sup>2+</sup>, Cu<sup>2+</sup>, Fe<sup>2+</sup>, Mn<sup>2+</sup>, Ni<sup>2+</sup>, and Zn<sup>2+</sup>) have been reported to modulate GPR68 activity in a concentration and pH dependent manner <sup>[171]</sup>. The same study

had reported that cobalt ions can activate the receptor weakly at pH 7.4 and at high concentrations ( $\geq 30 \mu\text{M}$ )<sup>[171]</sup>. The experiments I performed with C10 were at pH 7.8. It is difficult to directly compare the functional assay results to the behavior of the protein in the non-native environment. Therefore, I purified GPR68 constructs with other metal ions like  $\text{Cd}^{2+}$ ,  $\text{Mn}^{2+}$ ,  $\text{Ni}^{2+}$ , and  $\text{Zn}^{2+}$  (data not shown) and found that cobalt ions are the only ions that improve receptor stability the most followed by nickel ions whose effects were minimal at pH 7.8 and the other metal ions had no effect. Interestingly, the structure of active state GPR68 has been claimed to be solved in presence of  $10 \mu\text{M}$  cobalt ions<sup>[371, 375]</sup>. Thus, it is likely that the results that I had produced were likely meaningful.

A second disulfide bond has been predicted for GPR68 (C13<sup>1,24</sup>-C258<sup>ECL3</sup>) (section 1.3.1) and truncation of the first 14 amino acids would prevent disulfide bond formation. The *N*-terminal truncations were interesting especially for constructs where the first 14 amino acids were truncated e.g., C10, C27, C43 and C61 as these could be stabilized via usage of ligands (Figure 24A, 26E,F), which implies that either GPR68 does not have an additional disulfide bond, or this disulfide bond does not influence the structural integrity of the receptor greatly.

The deletion of 60 amino acids from the *C*-terminal end of the receptor as in C35 led to a strange behavior in comparison to when the *C*-terminal end of the receptor is only truncated by 18 amino acids as in C33. I observed that Ligand 2 stabilized C33, but not C35. The initial hypothesis I had was based on the presence of evidence that the *C*-terminal end or Helix 8 of the receptor may be structured as they play a role in hetero-oligomerization with Lysophosphatidic acid receptor 1 (LPA<sub>1</sub>)<sup>[142]</sup> and mechanosensing<sup>[196, 376]</sup> respectively. Additionally, AlphaFold model of GPR68 had a longer helix 8 than typical for class A GPCRs<sup>[339]</sup> which extended until about the residue 317 in GPR68. Later, I developed a second hypothesis when I observed that the Ligand 3 was able to stabilize C33 which had 18 amino acids deletion from the *C*-terminal end (Figure 26F), but not C46 which is basically C33 but with the two potential palmitoylation sites mutated (304, 310) (Figure 27D). It has been often described that many class A GPCRs have these *C*-terminal palmitoylations<sup>[10]</sup> and this post-translational modification can regulate receptor activity of GPCRs<sup>[377]</sup>. C35 which has 60 amino acids deleted from the *C*-terminal end does not contain the “C310” putative palmitoylation. Thus, based on literature data relevance and mutation constructs, it is likely that lack of palmitoylation causes loss of either surface

expression or misfolding of the GPR68 construct thus rendering it insensitive to ligands if a palmitoylation site is missing. The following interpretations led to the usage of C-terminal truncations which were not shorter than residue 319 in GPR68 constructs.

GPR68 has been annotated to have two N-glycosylation sites (N3 and N8) and a putative third N-glycosylation site at N86<sup>23,49</sup> [118]. The consensus motif for N-linked glycosylation is the asparagine-X-serine/threonine (NXS/T) motif, where X represents any amino acid except proline [378]. To understand the effects of mutating the putative N-glycosylation site (N86<sup>23,49</sup>) to glutamine, I tested two combination constructs (C18 and C19) which had a PGS in ICL3, with BRIL (N) and  $\Delta 8$  at the N-terminus in common. However, they differed in the presence of the mutant, N86<sup>23,49</sup>Q. No difference was observed between the C18 and C19 in the gels or chromatograms (Figure S9), thus implying that GPR68 likely lacks a third glycosylation site.

The public release of artificial intelligence based structural modeling tool, AlphaFold was very useful to the work I performed in my thesis. I used this tool to design rigid ICL3 BRIL fusion based GPR68 constructs (human and zebrafish) for cryo-EM. The importance of a rigid BRIL fusion was previously discussed in the section 4.1.8 of my thesis. Various strategies like using a state-of-the-art structural modeling tool, changing BRIL insertion points in ICL3, and changing the expression system from insect cells to mammalian cells to provide as close a native environment still failed to produce good quality human GPR68 for structure determination studies only yielding in fuzzy micelle 2D classes without a distinct fiducial marker as shown previously in Figure 33. The most puzzling behavior was no retention shift in presence of Anti-BRIL FAB for the BRIL fusion constructs (inserted anywhere i.e., N-terminus, ICL1, ICL2 or ICL3) in the chromatograms (section 4.1.8, Figure 29, 30). There is a possibility that the BRIL containing human GPR68 constructs were precipitating upon binding of the FAB as witnessed by the peak size comparison between monomer protein and protein complexed by FAB (Figure 32A). Although the Anti-BRIL FAB and Anti-BRIL-FAB nanobody were still observed on the gel in the complex peak, also identified by MALDI-TOF (Figure S10).

Additionally, a strange effect of salt concentration was witnessed for human GPR68 constructs which is highlighted here by using a representative construct, C61. The retention time or volume and peak characteristics like peak shape, size, tailing, shoulders etc. were

affected by the salt concentration mostly in LMNG/CHS or GDN, but for DDM/CHS only a minimal effect was observed (Figure S11). The preferred stability for DDM/CHS may be since the protein was initially purified using DDM/CHS and then buffer exchanged to LMNG/CHS or GDN by buffer exchange. Thus, as a preventive measure during cryo-EM a range of salt concentrations (0.15 M NaCl – 0.5 M NaCl) was tested considering this observation.

The zebrafish GPR68 construct (C99) was designed a *N*-terminal truncation of 8 amino acids to and the mutation N9Q to remove the potential N-glycosylation site. This leaves a residue before Cys10 to support a potential second disulfide bond formation if one exists similar to one described for the human GPR68. Additionally, a *C*-terminal deletion of 36 amino acids was designed until the first cysteine from the *C*-terminus was observed and an additional residue was left after that (serine in this case) to support palmitoylation to occur if one exists at this cysteine. Apart from these modifications, a thermostabilized green fluorescent protein (GFP) was added at the *N*-terminus and BRIL was inserted in the ICL3 based on AlphaFold structural modeling to allow formation of continuous helix from TM5 and TM6 to the BRIL fusion. Using this construct, a retention volume shift to the correct size upon addition of FAB and NB was observed. Finally, after using this construct for cryo-EM after complexation with FAB and NB, clear 2D classes were obtained with a fiducial marker distinctly visible (Figure 34E), which was never observed for any of the cryo-EM trials performed with human GPR68 (Figure 32). These improved results were obtained quite recently just before writing the thesis and the further steps of data processing were not pursued as the dataset had high noise or high number of empty detergent micelles in the dataset. This problem can likely be counteracted by testing other detergents or other membrane mimetic strategies like usage of amphipol, peptidisc or nanodisc. The MALDI-TOF results for all the components in the gel and the nanoDSF results for C99 in presence of Ligand 3 are shown in Figure S12. The C99 in presence of Ligand 3 had melting temperature of ~60-62°C.

The zebrafish GPR68 has been reported to be activated by acidic pH and was suggested to have an activation mechanism similar to human GPR68 [379]. Later a novel antagonist, ogramorphin was identified in a screen using zebrafish GPR68 [380] and the effects of this ligand has been witnessed against human GPR68 as well [204, 381]. The stabilization of zebrafish GPR68 with the proprietary Ligand 3 (Figure 34B,C), which was screened using

human GPR68, suggests that the ligand binding pocket for both these antagonists have likely a high similarity. Nevertheless, the transmembrane region of zebrafish and human GPR68 has good conservation. The zebrafish GPR68 has ~56 % sequence identity and ~72 % sequence similarity with the human GPR68 based on local sequence alignment and most of the sequence gaps are present in the C-terminus as the C-terminus of zebrafish is shorter than human GPR68 (Figure S13). Moreover, all the corresponding extracellular and transmembrane histidines of human GPR68 are conserved in zebrafish GPR68, except H84<sup>2.67</sup>, H89<sup>23.52</sup> and H169<sup>ECL2</sup> and all the four buried acidic groups are fully conserved (Figure S13).

The reasons for the non-visibility of fiducial markers in the human GPR68 constructs and the failure of complexation between the Anti-BRIL FAB and fused ICL3 BRIL based on appropriate quality control in the human GPR68 constructs were challenging to determine. Nevertheless, it is still a conundrum whether to claim if the human GPR68 constructs were truly stable or not. Although, the opposing evidence that I have in favor of my purification strategies is that the stability imparted via proprietary ligands was very evident, without which the receptor is pretty much unstable. These ligands did not absorb light at 280 nm and thus are not creating false positives in the chromatograms. Moreover, there was stability imparted by cobalt ions as well, using which the active state structure of GPR68 has been reported to be determined <sup>[375]</sup>.

A recently published work reported successful determination of the active state structure of human GPR68 and claimed that “Despite extensive attempts at obtaining a structure of inactive GPR68, we were unable to resolve this conformation of the receptor” <sup>[375]</sup>. This only supports the work I have done as a part of my doctoral thesis highlighting the complexity of working towards the inactive state structure of GPR68. Nevertheless, the advancements I reported with the zebrafish construct (C99) provides a solid foundation to pursue structure determination of the inactive state structure of zebrafish GPR68. The structure determination of zebrafish GPR68 can yield findings that can be applied to human GPR68 because of the similarity between the receptor in these two distinct organisms.

## 5.2 Molecular Dynamics Simulations

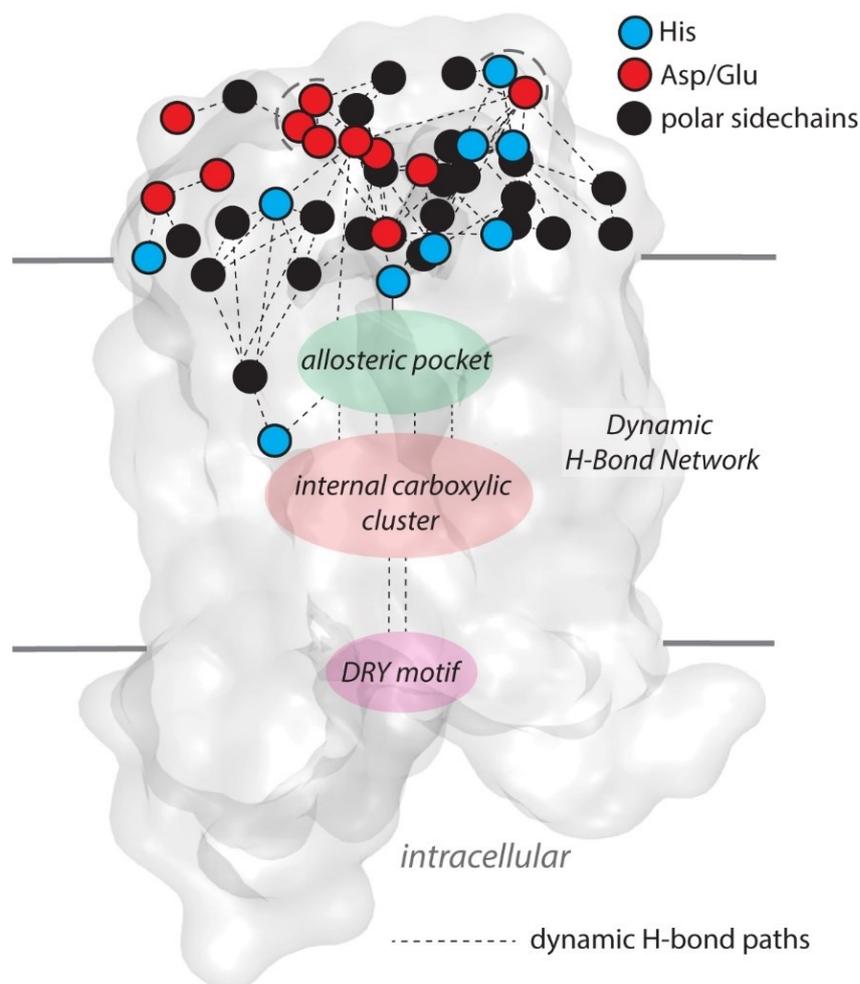


Figure 48: Summary of proposed proton sensing mechanism of GPR68. The figure has been reused from <sup>[339]</sup>.

Class A GPCRs typically have an orthosteric pocket, where their cognate endogenous ligand binds. Though a question for starters would be – But, what if a class A GPCR is endogenously activated by protons? This doctoral thesis proposes a theory regarding the proton-sensing capabilities of GPR68 and identifies the key areas involved in this process. The absence of a publicly available 3D atomic structure of GPR68 was a challenge and additionally the closest homologs of GPR68 only had up to ~20% sequence identities e.g., protease activated receptor 2. To understand the proton sensing mechanism of GPR68, I used different structural models provided to me, which were generated using GOMoDo and Modeller. Additionally, I utilized the structural model of GPR68 created by AlphaFold. I validated these models by comparing their standard protonation simulations in terms of overall rmsd (root mean square deviation) of the structures, number of buried waters,

assessment of secondary structure, and interhelical distances throughout the entire simulation period. Additionally, I compared the orientation of the class A GPCR motifs in these models to existing literature data for all initial structural models. An ideal model was then chosen based on the above-mentioned criteria. To understand protonation coupled dynamics of GPR68, select residues were protonated and their effect was studied using graph theory-based hydrogen bond network analysis.

GPR68 encompasses a highly dynamic water mediated hydrogen bond network which undergoes changes based on certain specific protonation states of buried acidic residues. I found that the extracellular histidine-carboxylate cluster and the buried carboxylates are a part of the larger common water mediated hydrogen bond network, although this network in standard protonation state of GPR68 only spans up to the internal carboxylates (section 4.2.3) and can extend up to the DR<sup>3.50</sup>Y motif in the cytoplasmic side depending on specific protonation states of select residues (section 4.2.4). The R3.50 of the DRY motif has been described to often have a direct interaction with G proteins thus aiding in activation of the receptors<sup>[101]</sup>. Earlier works on GPR68 have suggested that extracellular histidines<sup>[112]</sup> and buried carboxylates (D67<sup>2.50</sup>, E149<sup>4.53</sup> and D282<sup>7.49</sup>)<sup>[134]</sup> play a role in the proton sensing function of the receptor. However, the literature lacks further discussion on the mechanisms underlying this phenomenon.

I hypothesized that the bulk exposed histidine-carboxylate cluster and the buried carboxylate cluster work in tandem to bring about a conformational change in the receptor upon activation via protons. These connections are mostly long-distance water mediated interactions. I also observed that the previously proposed putative positive allosteric modulator pocket lies between these two crucial regions and thus likely controls the activation state of the receptor (Figure 48). I also observed that the buried carboxylate E103<sup>3.34</sup> is in spatial proximity to E149<sup>4.53</sup> and using MD simulation I found that it might play a role in the proton sensing function based on the hydrogen bond network analysis (section 4.2.4). The potential role of E103<sup>3.34</sup> has also been described after comparing the recently determined active structure of GPR68 with an AlphaFold model representing the inactive state conformation of GPR68<sup>[375]</sup>.

Interestingly, extracellular carboxylates (except E255<sup>ECL3</sup>), buried carboxylates (except D67<sup>2.50</sup>), and extracellular / TM histidines (H89<sup>23.52</sup>, H159<sup>4.63</sup>, and H245<sup>6.52</sup>), and E160<sup>4.64</sup>,

R189<sup>5.42</sup>, and Y244<sup>6.51</sup> of the allosteric pocket are a part of the cancer-associated mutations in GPR68. These compiled mutations shown in Figure S14 were identified directly from the tissues of the cancer patients and were compiled in Catalogue Of Somatic Mutations In Cancer (COSMIC) database [382, 383]. Thus, further studies are required to classify these mutants as either gain of function or loss of function mutations to further understand the link between these residues and cancer. These residues were selectively shown in the Figure 38 as they play a central role in the dynamic H-Bond network of GPR68 by forming a part of the core H-bond network of GPR68. The bulk exposed carboxylate and histidine cluster rich regions like GPR68 are also described for otopetrin proteins, where two extracellular loops which have multiple negatively charged residues accompanied by spatially proximal histidines have been described to be essential in proton sensing function of these proteins [384]. Many such other membrane proteins have similar regions and have been reviewed previously describing the presence and importance of proton binding motifs or proton antennas [385]. Usually, the bulk exposed negatively charged residues also assist in the formation of long tightly connected water wires, thus increasing the possibility of proton capture from the environment [386].

Lastly, an important observation was also made about the behavior of the sodium ion which is a negative allosteric modulator in class A GPCRs. Earlier computational studies in receptors like opioid and muscarinic acetylcholine receptors showed that the sodium ion usually enters from the extracellular side but may exit from either the extracellular side or the cytosolic side [387]. Later it was suggested that the cytosolic side exit maybe more favorable in muscarinic receptor [169] and in case of the protease activated receptor the sodium ion could enter from the intracellular side, as the allosteric pocket was suggested to be blocked from the extracellular side [388]. In my best compatible structural model of GPR68, the sodium ion also enters from the intracellular side, which suggests that potentially the entry to the allosteric is likely blocked from the extracellular side.

The hypotheses derived during the investigation of the proton sensing function of GPR68 using computational tools and strategies required further validation via experimental work. I subsequently accomplished this work, and the key results are discussed in the following section.

## 5.3 Functional Assay

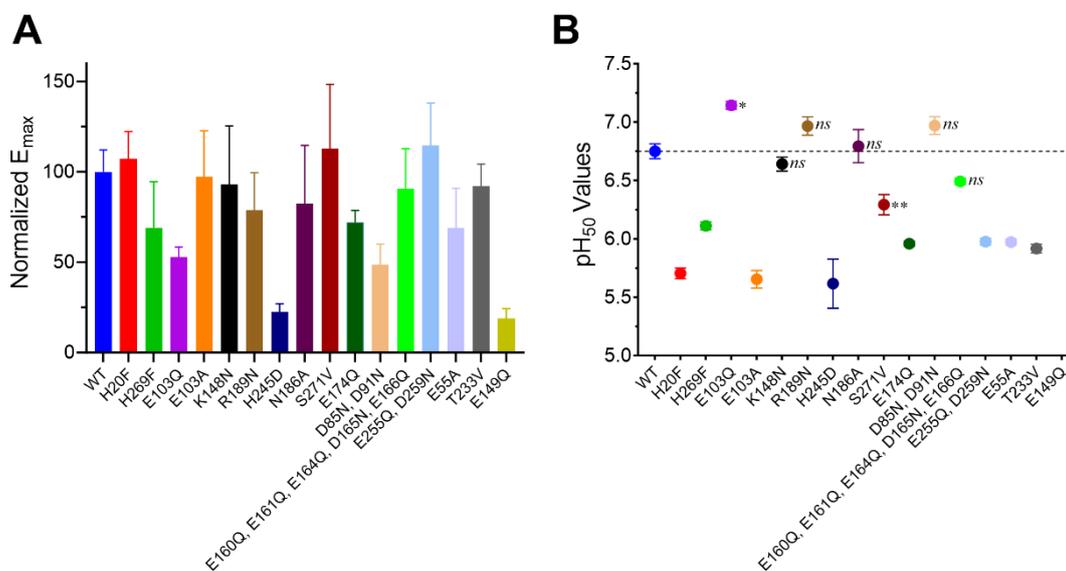


Figure 49: Summary of the functional analysis of GPR68 mutants. (A) Normalized efficacy maximum ( $E_{max}$ ) for GPR68 mutants. The data is represented as mean from 3 individual experiments with error bars representing the standard error. (B) Mean  $pH_{50}$  values for GPR68 mutants shown with error bars of standard error. The dotted line represents the  $pH_{50}$  value of WT-GPR68. The significance values have been marked as according to the p-value scale: ns ( $P > 0.05$ ), \* ( $P \leq 0.05$ ), \*\* ( $P \leq 0.01$ ), \*\*\* ( $P \leq 0.001$ ), and \*\*\*\* ( $P \leq 0.0001$ ). In cases of mutants where no symbol is marked implies \*\*\*\* P-value.

The first study investigating WT-GPR68 pH sensing range and half maximal-activation values reported the pH sensing range of GPR68 to be  $\sim$ pH 7.8 (fully inactive) to  $\sim$ pH 6.8 (fully active) and a half-maximal activation at  $\sim$ 7.44<sup>[112]</sup>. The values were highly similar when tested in HEK293 cells, CCL39 hamster fibroblasts, MG63 human osteosarcoma cells, and primary human preosteoblasts<sup>[112]</sup>. This study involves measurement of IP<sub>1</sub> accumulation, but via a different method which involves extraction from cells followed by isolation via batch column chromatography. The half-maximal activation reported in the study would suggest 50% of GPR68 population would be active at approximately pH 7.4, as the half-maximal activation value overlaps with the physiological pH in blood. This would be peculiar because GPR68 overexpression has been documented in various types of cancer<sup>[170, 389]</sup>, and extracellular acidification is a commonly observed feature in cancer<sup>[390]</sup>.

While, comparatively recent studies, suggest a half-maximal activation value of  $6.74 \pm 0.01$  for GPR68 and the pH sensing range to be approximately in the range of pH 7.4 (fully inactive) to pH 6.1 (fully active), when following cyclic adenosine monophosphate (cAMP)

accumulation via GloSensor cAMP assay in HEK293T cells [190, 371]. These results are in line with results generated by me using the HTRF assay based on IP<sub>1</sub> accumulation. Interestingly, another study utilizing the GloSensor cAMP assay format and based on the reported results the half-maximal activation of GPR68 was observed to be at approximately pH 6.1-6.3 [370]. The pH sensing range which could be approximately estimated from this research work is pH 6.7 (fully inactive) to pH 5.2 (fully active). Although, this study claims to have tracked cAMP production instead of cAMP accumulation as an explanation to deviation of results from previous studies [370]. This difference claimed by the previous report remains unclear.

Another study based on bio-luminescence resonance energy transfer (BRET) that tracks conformational changes within GPR68 was performed using mini-G $\alpha_{si}$  chimera primarily resembling G $\alpha_i$ -coupling indicated a pH<sub>50</sub> value of 6.35±0.02 in ‘BRET assay buffer’ and 6.70±0.11 in ‘Na<sup>+</sup>-PBS buffer’ for the wild-type receptor expressed in HEK293T/17 cells [134]. The pH sensing range which could be approximately estimated from this research work is pH 7.6 (fully inactive) to pH 6.2 (fully active) in ‘BRET assay buffer’ or pH 7.2 (fully inactive) to pH 5.5 (fully active) in ‘Na<sup>+</sup>-PBS buffer’.

The half-maximal activation for WT-GPR68 measured by me (section 4.3.1) deviates from the previously reported values of WT-GPR68. Deviations in the results of other mutants on comparison with previous literature data were also observed. These deviations could be mainly associated to the differences within the type of signaling pathway investigated for GPR68 along with overall assay setup, like the usage of different cell lines, buffer systems and assay method. Despite variations, as a consensus it seems the pH sensing range for WT-GPR68 falls in the pH range of pH 7.5 – pH 6.0 with a pH<sub>50</sub> value of ~6.7.

The data from WT-GPR68 and mutants H20<sup>1.31</sup>F, H269<sup>6.52</sup>F, and E149<sup>4.53</sup>Q which were used as literature control mutants [112, 134, 171] are very similar to the literature data. This signified that the assay I established was relevant showing the relevance of the modified HTRF assay for measuring GPR68 activity. As a general trend, high standard errors were observed in the efficacy maximum values for WT-GPR68 and its mutants, such variabilities may be due to usage of transiently transfected cells [391]. However, the pH<sub>50</sub> values only had relatively high standard error for the mutant, H245<sup>6.52</sup> which was the mutant causing the most reduction in proton sensing function of the receptor (Figure 49). As an exception, I

observed that the E149<sup>4.53</sup>Q mutant was a dead mutant, which was the data point with the largest discrepancy with the literature, where it was previously described as a constitutively active mutant when investigated based on GPR68-G $\alpha_s$  coupling [134]. Such significant differences may be due to the choice of screening assay used to investigate GPR68. I have used GPR68-G $\alpha_{q/11}$  coupling whereas the literature data is reported for GPR68-G $\alpha_s$  coupling and this suggests that perhaps this mutant could induce a G protein coupling bias. Although further experimentation would be required to answer this specific question.

I had proposed new set of residues (the surface carboxylates and the buried carboxylate, E103<sup>3.34</sup>) based on my MD simulation work, that could potentially play a role in the pH sensing mechanism of GPR68. This hypothesis was validated by E103<sup>3.34</sup>Q mutant which showed an upshift in pH<sub>50</sub> value by ~0.4 units signifying a decreased dependence on acidic pH for receptor activation. This suggests that E103<sup>3.34</sup> could be protonated during the receptor activation process as this mutant acts as a gain-of-function mutant. Thus, this mutant would support the idea that buried carboxylate E103<sup>3.34</sup>Q can be involved in the pH sensing function of the receptor, as neutralizing the buried carboxylate would mimic the protonated glutamate and potentially have upshifted pK<sub>a</sub> value like E149<sup>4.53</sup>Q as reported in previous research works [134].

Based on the best compatible structural model I had assessed, buried glutamates E149<sup>4.53</sup> and E103<sup>3.34</sup> are in proximity of K148<sup>4.52</sup> and R189<sup>5.42</sup> respectively. Considering the evidence that E149<sup>4.53</sup> and E103<sup>3.34</sup> could both have potentially upshifted pK<sub>a</sub> values, the major question that remains is whether these buried positively charged residues have a downshifted pK<sub>a</sub>? The negligible effects of the buried positively charged residue mutants R189<sup>5.42</sup>N, and K148<sup>4.52</sup>N suggests that these could potentially have downshifted pK<sub>a</sub> values resulting in non-significant differences in pH<sub>50</sub> values when compared to WT-GPR68 (Figure 49, Table 8).

The role of extracellular carboxylates functioning as potential proton antenna as carboxylate-carboxylate or carboxylate-histidine pair for GPR68 was discussed in the section 5.2 of my thesis. The proton concentration and proton capture are crucial initial steps in proton transfer which are often rate limiting [392–394]. The proton capture is facilitated by surface carboxylates [394, 395]. Thus, neutralization of the surface charge is likely to affect proton sensing function of GPR68. Considering the highly robust water

mediated hydrogen bond network of GPR68 (section 4.2.3), it is unlikely that single mutations of carboxylates to their neutral forms are going to exhibit considerable effect. Thus, the surface carboxylates were mostly mutated in groups: D85<sup>ECL1</sup>N and D91<sup>3.22</sup>N (extracellular loop 1 region), E160<sup>4.64</sup>Q, E161<sup>ECL2</sup>Q, E164<sup>ECL2</sup>Q, D165<sup>ECL2</sup>N and E166<sup>ECL2</sup>Q (extracellular loop 2 region), E174<sup>45.52</sup>Q (extracellular loop 2 region), and E255<sup>ECL3</sup>Q and D259<sup>ECL3</sup>Q (extracellular loop 3 region).

From the surface carboxylate mutants, the mutant group at ECL1 (D85<sup>ECL1</sup>N, D91<sup>3.22</sup>N) and the mutant group at ECL2 (E160Q, E161Q, E164Q, D165N and E166Q) did not show significant shift in pH<sub>50</sub> values, but the E<sub>max</sub> for the mutant D85N, D91N was reduced by 50% of WT-GPR68. Moreover, the single ECL2 mutant, E174<sup>45.52</sup>Q, showed a significant drop in both pH<sub>50</sub> value and efficacy maximum. Interestingly, only the surface carboxylate group in ECL3 (E255Q, D259N) showed significant drop in pH<sub>50</sub> value and a rather similar E<sub>max</sub> to WT. This indicates that of all the surface carboxylates, E174<sup>45.52</sup> is the most important for receptor function, followed by the group E255<sup>ECL3</sup>Q, D259<sup>ECL3</sup>N and the group D85N, D91N, whereas ECL2 surface carboxylate grouped mutant barely had any effects on receptor function showing that the residues in the extracellular loop 3 are collectively more important for receptor function than the multi-neutralizing extracellular loop 2 mutant (E160Q, E161Q, E164Q, D165Q and E166Q). Based on the active state structure description of GPR68 and chimeric construct-based investigation of extracellular loop regions of GPR68 revealed that ECL2-ECL3 region is likely the crucial proton sensing element in GPR68 [375]. E174<sup>45.52</sup>Q, a single mutant in the extracellular loop 2 was more effective in reducing pH sensing function of GPR68 over the multi-carboxylate neutralization in the extracellular loop 2. It is important to note that the position 45.52 is highly conserved within class A G protein-coupled receptors (GPCRs) and often points towards the inner core of the 7TM fold of the class A GPCRs [51] and this residue was also considered crucial based on mutation study and comparison of the active state structure of GPR68 with the AlphaFold model of GPR68 which was claimed to represent inactive state of GPR68 [375]. This residue is also a part of the highly dynamic hydrogen bond network of GPR68 which is very well connected to the buried carboxylates E103<sup>3.34</sup> and E149<sup>4.53</sup> (Figure 39).

Finally, it was also crucial to investigate the intracellular side of GPR68, where I found that a unique residue deviating drastically from sequence conservation patterns from other class

A GPCRs is conserved within proton sensing GPCRs (GPR4, GPR65 and GPR68) (section 4.3.5). The mutant E55<sup>2.38</sup>A also drastically affects GPR68 half maximal activation by a drop in ~0.8 pH units (Table 8) like the following mutants whose importance have been discussed heavily earlier in this section E174<sup>45.52</sup>Q, and the ECL3 surface charge neutral mutant (E255<sup>ECL3</sup>Q and D259<sup>ECL3</sup>N).

As next steps, I plan to perform fluorescence microscopy studies using a V5 tag and an Anti-V5 antibody conjugated with a fluorophore to check for cell-surface expression quantitatively. Furthermore, I also plan to perform the HTRF assay modified for G $\alpha_s$  coupling for WT and all mutants, as this data from different signaling pathways can act as consensus for overall results and may provide insights mainly for addressing the unexpected behavior of E149<sup>4.53</sup>Q mutant.

## References

- [1] Chu, R.; Takei, J.; Knowlton, J. R.; Andrykovitch, M.; Pei, W.; Kajava, A. V.; Steinbach, P. J.; Ji, X.; Bai, Y. Redesign of a Four-Helix Bundle Protein by Phage Display Coupled with Proteolysis and Structural Characterization by NMR and X-Ray Crystallography. *J. Mol. Biol.*, **2002**, *323* (2), 253–262. [https://doi.org/10.1016/s0022-2836\(02\)00884-7](https://doi.org/10.1016/s0022-2836(02)00884-7).
- [2] Bjarnadóttir, T. K.; Gloriam, D. E.; Hellstrand, S. H.; Kristiansson, H.; Fredriksson, R.; Schiöth, H. B. Comprehensive Repertoire and Phylogenetic Analysis of the G Protein-Coupled Receptors in Human and Mouse. *Genomics*, **2006**, *88* (3), 263–273. <https://doi.org/10.1016/j.ygeno.2006.04.001>.
- [3] Zhang, Y.; DeVries, M. E.; Skolnick, J. Structure Modeling of All Identified G Protein-Coupled Receptors in the Human Genome. *PLoS Comput. Biol.*, **2006**, *2* (2), e13. <https://doi.org/10.1371/journal.pcbi.0020013>.
- [4] Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Trong, I. L.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; et al. Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. *Science*, **2000**, *289* (5480), 739–745. <https://doi.org/10.1126/science.289.5480.739>.
- [5] Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; et al. High-Resolution Crystal Structure of an Engineered Human B2-Adrenergic G Protein-Coupled Receptor. *Science*, **2007**, *318* (5854), 1258–1265. <https://doi.org/10.1126/science.1150577>.
- [6] Pándy-Szekeres, G.; Munk, C.; Tsonkov, T. M.; Mordalski, S.; Harpsøe, K.; Hauser, A. S.; Bojarski, A. J.; Gloriam, D. E. GPCRdb in 2018: Adding GPCR Structure Models and Ligands. *Nucleic Acids Res.*, **2017**, *46* (Database issue), gkx1109. <https://doi.org/10.1093/nar/gkx1109>.
- [7] Ji, T. H.; Grossmann, M.; Ji, I. G Protein-Coupled Receptors I. DIVERSITY OF RECEPTOR-LIGAND INTERACTIONS. *J. Biol. Chem.*, **1998**, *273* (28), 17299–17302. <https://doi.org/10.1074/jbc.273.28.17299>.
- [8] Sato, T.; Kawasaki, T.; Mine, S.; Matsumura, H. Functional Role of the C-Terminal Amphipathic Helix 8 of Olfactory Receptors and Other G Protein-Coupled Receptors. *Int. J. Mol. Sci.*, **2016**, *17* (11), 1930. <https://doi.org/10.3390/ijms17111930>.
- [9] Davis, D. P.; Segaloff, D. L. N-Linked Carbohydrates on G Protein-Coupled Receptors: Mapping Sites of Attachment and Determining Functional Roles. *Methods Enzym.*, **2002**, *343*, 200–212. [https://doi.org/10.1016/s0076-6879\(02\)43136-9](https://doi.org/10.1016/s0076-6879(02)43136-9).

- [10] Zhang, B.; Li, S.; Shui, W. Post-Translational Modifications of G Protein–Coupled Receptors Revealed by Proteomics and Structural Biology. *Front. Chem.*, **2022**, *10*, 843502. <https://doi.org/10.3389/fchem.2022.843502>.
- [11] Karnik, S. S.; Gogonea, C.; Patil, S.; Saad, Y.; Takezako, T. Activation of G-Protein-Coupled Receptors: A Common Molecular Mechanism. *Trends Endocrinol. Metab.*, **2003**, *14* (9), 431–437. <https://doi.org/10.1016/j.tem.2003.09.007>.
- [12] Ohyama, K.; Yamano, Y.; Sano, T.; Nakagomi, Y.; Hamakubo, T.; Morishima, I.; Inagami, T. Disulfide Bridges in Extracellular Domains of Angiotensin II Receptor Type IA. *Regul. Pept.*, **1995**, *57* (2), 141–147. [https://doi.org/10.1016/0167-0115\(95\)00030-f](https://doi.org/10.1016/0167-0115(95)00030-f).
- [13] Wu, B.; Chien, E. Y. T.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; et al. Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists. *Science*, **2010**, *330* (6007), 1066–1071. <https://doi.org/10.1126/science.1194396>.
- [14] Cook, J. V. F.; Eidne, K. A. An Intramolecular Disulfide Bond between Conserved Extracellular Cysteines in the Gonadotropin-Releasing Hormone Receptor Is Essential for Binding and Activation 1. *Endocrinology*, **1997**, *138* (7), 2800–2806. <https://doi.org/10.1210/endo.138.7.5233>.
- [15] Jaakola, V.-P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y. T.; Lane, J. R.; IJzerman, A. P.; Stevens, R. C. The 2.6 Angstrom Crystal Structure of a Human A2A Adenosine Receptor Bound to an Antagonist. *Science*, **2008**, *322* (5905), 1211–1217. <https://doi.org/10.1126/science.1164772>.
- [16] Tarnow, P.; Schöneberg, T.; Krude, H.; Grüters, A.; Biebermann, H. Mutationally Induced Disulfide Bond Formation within the Third Extracellular Loop Causes Melanocortin 4 Receptor Inactivation in Patients with Obesity. *J. Biol. Chem.*, **2003**, *278* (49), 48666–48673. <https://doi.org/10.1074/jbc.m309941200>.
- [17] Duvernay, M. T.; Filipeanu, C. M.; Wu, G. The Regulatory Mechanisms of Export Trafficking of G Protein-Coupled Receptors. *Cell. Signal.*, **2005**, *17* (12), 1457–1465. <https://doi.org/10.1016/j.cellsig.2005.05.020>.
- [18] Drake, M. T.; Shenoy, S. K.; Lefkowitz, R. J. Trafficking of G Protein–Coupled Receptors. *Circ. Res.*, **2006**, *99* (6), 570–582. <https://doi.org/10.1161/01.res.0000242563.47507.ce>.
- [19] Dong, C.; Filipeanu, C. M.; Duvernay, M. T.; Wu, G. Regulation of G Protein-Coupled Receptor Export Trafficking. *Biochim. Biophys. Acta (BBA) - Biomembr.*, **2007**, *1768* (4), 853–870. <https://doi.org/10.1016/j.bbamem.2006.09.008>.
- [20] Schneider, J.; Korshunova, K.; Musiani, F.; Alfonso-Prieto, M.; Giorgetti, A.; Carloni, P. Predicting Ligand Binding Poses for Low-Resolution Membrane Protein Models: Perspectives from Multiscale Simulations. *Biochem. Biophys. Res. Commun.*, **2018**, *498* (2), 366–374. <https://doi.org/10.1016/j.bbrc.2018.01.160>.

- [21] Foster, S. R.; Hauser, A. S.; Vedel, L.; Strachan, R. T.; Huang, X.-P.; Gavin, A. C.; Shah, S. D.; Nayak, A. P.; Haugaard-Kedström, L. M.; Penn, R. B.; et al. Discovery of Human Signaling Systems: Pairing Peptides to G Protein-Coupled Receptors. *Cell*, **2019**, *179* (4), 895-908.e21. <https://doi.org/10.1016/j.cell.2019.10.010>.
- [22] Kobilka, B. K. G Protein Coupled Receptor Structure and Activation. *Biochim. Biophys. Acta (BBA) - Biomembr.*, **2007**, *1768* (4), 794–807. <https://doi.org/10.1016/j.bbamem.2006.10.021>.
- [23] Rosenbaum, D. M.; Rasmussen, S. G. F.; Kobilka, B. K. The Structure and Function of G-Protein-Coupled Receptors. *Nature*, **2009**, *459* (7245), 356–363. <https://doi.org/10.1038/nature08144>.
- [24] Strange, P. G. Agonist Binding, Agonist Affinity and Agonist Efficacy at G Protein-coupled Receptors. *Br. J. Pharmacol.*, **2008**, *153* (7), 1353–1363. <https://doi.org/10.1038/sj.bjp.0707672>.
- [25] Bourne, H. R.; Sanders, D. A.; McCormick, F. The GTPase Superfamily: Conserved Structure and Molecular Mechanism. *Nature*, **1991**, *349* (6305), 117–127. <https://doi.org/10.1038/349117a0>.
- [26] Oldham, W. M.; Hamm, H. E. Heterotrimeric G Protein Activation by G-Protein-Coupled Receptors. *Nat Rev Mol Cell Bio*, **2008**, *9* (1), 60–71. <https://doi.org/10.1038/nrm2299>.
- [27] Weis, W. I.; Kobilka, B. K. The Molecular Basis of G Protein–Coupled Receptor Activation. *Annu Rev Biochem*, **2014**, *87* (1), 897–919. <https://doi.org/10.1146/annurev-biochem-060614-033910>.
- [28] Tobin, A. B. G-protein-coupled Receptor Phosphorylation: Where, When and by Whom. *Br. J. Pharmacol.*, **2008**, *153* (S1), S167–S176. <https://doi.org/10.1038/sj.bjp.0707662>.
- [29] Paradis, J. S.; Ly, S.; Blondel-Tepaz, É.; Galan, J. A.; Beaufrais, A.; Scott, M. G. H.; Enslin, H.; Marullo, S.; Roux, P. P.; Bouvier, M. Receptor Sequestration in Response to  $\beta$ -Arrestin-2 Phosphorylation by ERK1/2 Governs Steady-State Levels of GPCR Cell-Surface Expression. *Proc. Natl. Acad. Sci.*, **2015**, *112* (37), E5160–E5168. <https://doi.org/10.1073/pnas.1508836112>.
- [30] Gurevich, V. V.; Gurevich, E. V. GPCR Signaling Regulation: The Role of GRKs and Arrestins. *Front. Pharmacol.*, **2019**, *10*, 125. <https://doi.org/10.3389/fphar.2019.00125>.
- [31] Attramadal, H.; Arriza, J. L.; Aoki, C.; Dawson, T. M.; Codina, J.; Kwatra, M. M.; Snyder, S. H.; Caron, M. G.; Lefkowitz, R. J. Beta-Arrestin2, a Novel Member of the Arrestin/Beta-Arrestin Gene Family. *J. Biol. Chem.*, **1992**, *267* (25), 17882–17890. [https://doi.org/10.1016/s0021-9258\(19\)37125-x](https://doi.org/10.1016/s0021-9258(19)37125-x).

- [32] Marchese, A.; Paing, M. M.; Temple, B. R. S.; Trejo, J. G Protein–Coupled Receptor Sorting to Endosomes and Lysosomes. *Annu. Rev. Pharmacol. Toxicol.*, **2008**, *48* (1), 601–629. <https://doi.org/10.1146/annurev.pharmtox.48.113006.094646>.
- [33] Zastrow, M. von; Kobilka, B. K. Antagonist-Dependent and -Independent Steps in the Mechanism of Adrenergic Receptor Internalization. *J. Biol. Chem.*, **1994**, *269* (28), 18448–18452. [https://doi.org/10.1016/s0021-9258\(17\)32329-3](https://doi.org/10.1016/s0021-9258(17)32329-3).
- [34] Claing, A.; Laporte, S. A.; Caron, M. G.; Lefkowitz, R. J. Endocytosis of G Protein–Coupled Receptors: Roles of G Protein–Coupled Receptor Kinases and  $\beta$ -Arrestin Proteins. *Prog. Neurobiol.*, **2002**, *66* (2), 61–79. [https://doi.org/10.1016/s0301-0082\(01\)00023-5](https://doi.org/10.1016/s0301-0082(01)00023-5).
- [35] Thomas, C. M.; Smart, E. J. Caveolae Structure and Function. *J. Cell. Mol. Med.*, **2008**, *12* (3), 796–809. <https://doi.org/10.1111/j.1582-4934.2008.00295.x>.
- [36] Sandoval, I. V.; Bakke, O. Targeting of Membrane Proteins to Endosomes and Lysosomes. *Trends Cell Biol.*, **1994**, *4* (8), 292–297. [https://doi.org/10.1016/0962-8924\(94\)90220-8](https://doi.org/10.1016/0962-8924(94)90220-8).
- [37] Bouley, R.; Sun, T.-X.; Chenard, M.; McLaughlin, M.; McKee, M.; Lin, H. Y.; Brown, D.; Ausiello, D. A. Functional Role of the NPxxY Motif in Internalization of the Type 2 Vasopressin Receptor in LLC-PK1 Cells. *Am J Physiol-cell Ph*, **2003**, *285* (4), C750–C762. <https://doi.org/10.1152/ajpcell.00477.2002>.
- [38] He, R.; Browning, D. D.; Ye, R. D. Differential Roles of the NPXXY Motif in Formyl Peptide Receptor Signaling. *J. Immunol.*, **2001**, *166* (6), 4099–4105. <https://doi.org/10.4049/jimmunol.166.6.4099>.
- [39] Patwardhan, A.; Cheng, N.; Trejo, J. Post-Translational Modifications of G Protein–Coupled Receptors Control Cellular Signaling Dynamics in Space and Time. *Pharmacol. Rev.*, **2021**, *73* (1), 120–151. <https://doi.org/10.1124/pharmrev.120.000082>.
- [40] Cao, T. T.; Deacon, H. W.; Reczek, D.; Bretscher, A.; Zastrow, M. von. A Kinase-Regulated PDZ-Domain Interaction Controls Endocytic Sorting of the B2-Adrenergic Receptor. *Nature*, **1999**, *401* (6750), 286–290. <https://doi.org/10.1038/45816>.
- [41] Nordström, K. J. V.; Almén, M. S.; Edstam, M. M.; Fredriksson, R.; Schiöth, H. B. Independent HHsearch, Needleman–Wunsch-Based, and Motif Analyses Reveal the Overall Hierarchy for Most of the G Protein–Coupled Receptor Families. *Mol. Biol. Evol.*, **2011**, *28* (9), 2471–2480. <https://doi.org/10.1093/molbev/msr061>.
- [42] Schiöth, H. B.; Fredriksson, R. The GRAFS Classification System of G-Protein Coupled Receptors in Comparative Perspective. *Gen. Comp. Endocrinol.*, **2005**, *142* (1–2), 94–101. <https://doi.org/10.1016/j.ygcen.2004.12.018>.
- [43] Katritch, V.; Cherezov, V.; Stevens, R. C. Diversity and Modularity of G Protein–Coupled Receptor Structures. *Trends Pharmacol. Sci.*, **2012**, *33* (1), 17–27. <https://doi.org/10.1016/j.tips.2011.09.003>.

- [44] Horn, F.; Bettler, E.; Oliveira, L.; Campagne, F.; Cohen, F. E.; Vriend, G. GPCRDB Information System for G Protein-Coupled Receptors. *Nucleic Acids Res.*, **2003**, *31* (1), 294–297. <https://doi.org/10.1093/nar/gkg103>.
- [45] Tang, X.; Wang, Y.; Li, D.; Luo, J.; Liu, M. Orphan G Protein-Coupled Receptors (GPCRs): Biological Functions and Potential Drug Targets. *Acta Pharmacol Sin*, **2012**, *33* (3), 363–371. <https://doi.org/10.1038/aps.2011.210>.
- [46] Ballesteros, J. A.; Weinstein, H. Integrated Methods for the Construction of Three-Dimensional Models and Computational Probing of Structure-Function Relations in G Protein-Coupled Receptors. *Methods Neurosci*, **1995**, *25*, 366–428. [https://doi.org/10.1016/s1043-9471\(05\)80049-7](https://doi.org/10.1016/s1043-9471(05)80049-7).
- [47] Wootten, D.; Simms, J.; Miller, L. J.; Christopoulos, A.; Sexton, P. M. Polar Transmembrane Interactions Drive Formation of Ligand-Specific and Signal Pathway-Biased Family B G Protein-Coupled Receptor Conformations. *Proc. Natl. Acad. Sci.*, **2013**, *110* (13), 5211–5216. <https://doi.org/10.1073/pnas.1221585110>.
- [48] Pin, J.-P.; Galvez, T.; Prézeau, L. Evolution, Structure, and Activation Mechanism of Family 3/C G-Protein-Coupled Receptors. *Pharmacol. Ther.*, **2003**, *98* (3), 325–354. [https://doi.org/10.1016/s0163-7258\(03\)00038-x](https://doi.org/10.1016/s0163-7258(03)00038-x).
- [49] Wang, C.; Wu, H.; Evron, T.; Vardy, E.; Han, G. W.; Huang, X.-P.; Hufeisen, S. J.; Mangano, T. J.; Urban, D. J.; Katritch, V.; et al. Structural Basis for Smoothed Receptor Modulation and Chemoresistance to Anticancer Drugs. *Nat. Commun.*, **2014**, *5* (1), 4355. <https://doi.org/10.1038/ncomms5355>.
- [50] Smit, M. J.; Vischer, H. F.; Bakker, R. A.; Jongejan, A.; Timmerman, H.; Pardo, L.; Leurs, R. Pharmacogenomic and Structural Analysis of Constitutive G Protein-Coupled Receptor Activity. *Annu. Rev. Pharmacol. Toxicol.*, **2007**, *47* (1), 53–87. <https://doi.org/10.1146/annurev.pharmtox.47.120505.105126>.
- [51] Isberg, V.; Graaf, C. de; Bortolato, A.; Cherezov, V.; Katritch, V.; Marshall, F. H.; Mordalski, S.; Pin, J.-P.; Stevens, R. C.; Vriend, G.; et al. Generic GPCR Residue Numbers – Aligning Topology Maps While Minding the Gaps. *Trends Pharmacol Sci*, **2015**, *36* (1), 22–31. <https://doi.org/10.1016/j.tips.2014.11.001>.
- [52] Syrovatkina, V.; Alegre, K. O.; Dey, R.; Huang, X.-Y. Regulation, Signaling, and Physiological Functions of G-Proteins. *J. Mol. Biol.*, **2016**, *428* (19), 3850–3868. <https://doi.org/10.1016/j.jmb.2016.08.002>.
- [53] Strathmann, M.; Wilkie, T. M.; Simon, M. I. Diversity of the G-Protein Family: Sequences from Five Additional Alpha Subunits in the Mouse. *Proc. Natl. Acad. Sci.*, **1989**, *86* (19), 7407–7409. <https://doi.org/10.1073/pnas.86.19.7407>.
- [54] Raymond, J. H.; Aktary, Z.; Larue, L.; Delmas, V. Targeting GPCRs and Their Signaling as a Therapeutic Option in Melanoma. *Cancers*, **2022**, *14* (3), 706. <https://doi.org/10.3390/cancers14030706>.

- [55] Wettschureck, N.; Offermanns, S. Mammalian G Proteins and Their Cell Type Specific Functions. *Physiol. Rev.*, **2005**, 85 (4), 1159–1204. <https://doi.org/10.1152/physrev.00003.2005>.
- [56] Rhee, S. G.; Bae, Y. S. Regulation of Phosphoinositide-Specific Phospholipase C Isozymes. *J. Biol. Chem.*, **1997**, 272 (24), 15045–15048. <https://doi.org/10.1074/jbc.272.24.15045>.
- [57] Dhanasekaran, N.; Dermott, J. M. Signaling by the G12 Class of G Proteins. *Cell. Signal.*, **1996**, 8 (4), 235–245. [https://doi.org/10.1016/0898-6568\(96\)00048-4](https://doi.org/10.1016/0898-6568(96)00048-4).
- [58] Kozasa, T.; Jiang, X.; Hart, M. J.; Sternweis, P. M.; Singer, W. D.; Gilman, A. G.; Bollag, G.; Sternweis, P. C. P115 RhoGEF, a GTPase Activating Protein for G $\alpha$ 12 and G $\alpha$ 13. *Science*, **1998**, 280 (5372), 2109–2111. <https://doi.org/10.1126/science.280.5372.2109>.
- [59] Cerione, R. A.; Codina, J.; Benovic, J. L.; Lefkowitz, R. J.; Birnbaumer, L.; Caron, M. G. Mammalian  $\beta$ 2-Adrenergic Receptor: Reconstitution of Functional Interactions between Pure Receptor and Pure Stimulatory Nucleotide Binding Protein of the Adenylate Cyclase System. *Biochemistry*, **1984**, 23 (20), 4519–4525. <https://doi.org/10.1021/bi00315a003>.
- [60] Costa, T.; Herz, A. Antagonists with Negative Intrinsic Activity at Delta Opioid Receptors Coupled to GTP-Binding Proteins. *Proc. Natl. Acad. Sci.*, **1989**, 86 (19), 7321–7325. <https://doi.org/10.1073/pnas.86.19.7321>.
- [61] Szkudlinski, M. W. New Frontier in Glycoprotein Hormones and Their Receptors Structure–Function. *Front. Endocrinol.*, **2015**, 6, 155. <https://doi.org/10.3389/fendo.2015.00155>.
- [62] Alix, K.; Khatri, S.; Mosier, P. D.; Casterlow, S.; Yan, D.; Nyce, H. L.; White, M. M.; Schulte, M. K.; Dukat, M. Superagonist, Full Agonist, Partial Agonist, and Antagonist Actions of Arylguanidines at 5-Hydroxytryptamine-3 (5-HT<sub>3</sub>) Subunit A Receptors. *ACS Chem. Neurosci.*, **2016**, 7 (11), 1565–1574. <https://doi.org/10.1021/acschemneuro.6b00196>.
- [63] Wacker, D.; Stevens, R. C.; Roth, B. L. How Ligands Illuminate GPCR Molecular Pharmacology. *Cell*, **2017**, 170 (3), 414–427. <https://doi.org/10.1016/j.cell.2017.07.009>.
- [64] Canto, I.; Soh, U. J. K.; Trejo, J. Allosteric Modulation of Protease-Activated Receptor Signaling. *Mini-Rev. Med. Chem.*, **2012**, 12 (9), 804–811. <https://doi.org/10.2174/138955712800959116>.
- [65] Avlani, V. A.; Gregory, K. J.; Morton, C. J.; Parker, M. W.; Sexton, P. M.; Christopoulos, A. Critical Role for the Second Extracellular Loop in the Binding of Both Orthosteric and Allosteric G Protein-Coupled Receptor Ligands. *J. Biol. Chem.*, **2007**, 282 (35), 25677–25686. <https://doi.org/10.1074/jbc.m702311200>.

- [66] Woolley, Michael. J.; Watkins, H. A.; Taddese, B.; Karakullukcu, Z. G.; Barwell, J.; Smith, K. J.; Hay, D. L.; Poyner, D. R.; Reynolds, C. A.; Conner, A. C. The Role of ECL2 in CGRP Receptor Activation: A Combined Modelling and Experimental Approach. *J. R. Soc. Interface*, **2013**, *10* (88), 20130589. <https://doi.org/10.1098/rsif.2013.0589>.
- [67] Maeda, S.; Shiimura, Y.; Asada, H.; Hirata, K.; Luo, F.; Nango, E.; Tanaka, N.; Toyomoto, M.; Inoue, A.; Aoki, J.; et al. Endogenous Agonist–Bound S1PR3 Structure Reveals Determinants of G Protein–Subtype Bias. *Sci. Adv.*, **2021**, *7* (24), eabf5325. <https://doi.org/10.1126/sciadv.abf5325>.
- [68] Chan, H. C. S.; Li, Y.; Dahoun, T.; Vogel, H.; Yuan, S. New Binding Sites, New Opportunities for GPCR Drug Discovery. *Trends Biochem. Sci.*, **2019**, *44* (4), 312–330. <https://doi.org/10.1016/j.tibs.2018.11.011>.
- [69] Oswald, C.; Rappas, M.; Kean, J.; Doré, A. S.; Errey, J. C.; Bennett, K.; Deflorian, F.; Christopher, J. A.; Jazayeri, A.; Mason, J. S.; et al. Intracellular Allosteric Antagonism of the CCR9 Receptor. *Nature*, **2016**, *540* (7633), 462–465. <https://doi.org/10.1038/nature20606>.
- [70] Valant, C.; Sexton, P. M.; Christopoulos, A. Orthosteric/Allosteric Bitopic Ligands: Going Hybrid at GPCRs. *Mol. Interv.*, **2009**, *9* (3), 125–135. <https://doi.org/10.1124/mi.9.3.6>.
- [71] Kenakin, T.; Christopoulos, A. Signalling Bias in New Drug Discovery: Detection, Quantification and Therapeutic Impact. *Nat. Rev. Drug Discov.*, **2013**, *12* (3), 205–216. <https://doi.org/10.1038/nrd3954>.
- [72] Wootten, D.; Christopoulos, A.; Marti-Solano, M.; Babu, M. M.; Sexton, P. M. Mechanisms of Signalling and Biased Agonism in G Protein-Coupled Receptors. *Nat Rev Mol Cell Bio*, **2018**, *19* (10), 638–653. <https://doi.org/10.1038/s41580-018-0049-3>.
- [73] Smith, J. S.; Lefkowitz, R. J.; Rajagopal, S. Biased Signalling: From Simple Switches to Allosteric Microprocessors. *Nat. Rev. Drug Discov.*, **2018**, *17* (4), 243–260. <https://doi.org/10.1038/nrd.2017.229>.
- [74] Alexander, S. P. H.; Benson, H. E.; Faccenda, E.; Pawson, A. J.; Sharman, J. L.; Spedding, M.; Peters, J. A.; Harmar, A. J.; Collaborators, C. G Protein-coupled Receptors. *Br. J. Pharmacol.*, **2013**, *170* (8), 1459–1581. <https://doi.org/10.1111/bph.12445>.
- [75] Yang, D.; Zhou, Q.; Labroska, V.; Qin, S.; Darbalaei, S.; Wu, Y.; Yuliantie, E.; Xie, L.; Tao, H.; Cheng, J.; et al. G Protein-Coupled Receptors: Structure- and Function-Based Drug Discovery. *Signal Transduct. Target. Ther.*, **2021**, *6* (1), 7. <https://doi.org/10.1038/s41392-020-00435-w>.
- [76] Deupi, X.; Standfuss, J. Structural Insights into Agonist-Induced Activation of G-Protein-Coupled Receptors. *Curr. Opin. Struct. Biol.*, **2011**, *21* (4), 541–551. <https://doi.org/10.1016/j.sbi.2011.06.002>.

- [77] Olivella, M.; Caltabiano, G.; Cordero, A. The Role of Cysteine 6.47 in Class A GPCRs. *Bmc Struct Biol*, **2013**, *13* (1), 3. <https://doi.org/10.1186/1472-6807-13-3>.
- [78] Wacker, D.; Wang, C.; Katritch, V.; Han, G. W.; Huang, X.-P.; Vardy, E.; McCorvy, J. D.; Jiang, Y.; Chu, M.; Siu, F. Y.; et al. Structural Features for Functional Selectivity at Serotonin Receptors. *Science*, **2013**, *340* (6132), 615–619. <https://doi.org/10.1126/science.1232808>.
- [79] Rasmussen, S. G. F.; Choi, H.-J.; Fung, J. J.; Pardon, E.; Casarosa, P.; Chae, P. S.; DeVree, B. T.; Rosenbaum, D. M.; Thian, F. S.; Kobilka, T. S.; et al. Structure of a Nanobody-Stabilized Active State of the B2 Adrenoceptor. *Nature*, **2011**, *469* (7329), 175–180. <https://doi.org/10.1038/nature09648>.
- [80] Katritch, V.; Fenalti, G.; Abola, E. E.; Roth, B. L.; Cherezov, V.; Stevens, R. C. Allosteric Sodium in Class A GPCR Signaling. *Trends Biochem Sci*, **2014**, *39* (5), 233–244. <https://doi.org/10.1016/j.tibs.2014.03.002>.
- [81] Standfuss, J.; Edwards, P. C.; D’Antona, A.; Fransen, M.; Xie, G.; Oprian, D. D.; Schertler, G. F. X. The Structural Basis of Agonist-Induced Activation in Constitutively Active Rhodopsin. *Nature*, **2011**, *471* (7340), 656–660. <https://doi.org/10.1038/nature09795>.
- [82] Mitchell, R.; McCulloch, D.; Lutz, E.; Johnson, M.; MacKenzie, C.; Fennell, M.; Fink, G.; Zhou, W.; Sealfon, S. C. Rhodopsin-Family Receptors Associate with Small G Proteins to Activate Phospholipase D. *Nature*, **1998**, *392* (6674), 411–414. <https://doi.org/10.1038/32937>.
- [83] Fritze, O.; Filipek, S.; Kuksa, V.; Palczewski, K.; Hofmann, K. P.; Ernst, O. P. Role of the Conserved NPxxY(x)5,6F Motif in the Rhodopsin Ground State and during Activation. *Proc National Acad Sci*, **2003**, *100* (5), 2290–2295. <https://doi.org/10.1073/pnas.0435715100>.
- [84] Rovati, G. E.; Capra, V.; Neubig, R. R. The Highly Conserved DRY Motif of Class A G Protein-Coupled Receptors: Beyond the Ground State. *Mol Pharmacol*, **2007**, *71* (4), 959–964. <https://doi.org/10.1124/mol.106.029470>.
- [85] Thomson, N. J.; Vickery, O. N.; Ives, C. M.; Zachariae, U. Ion-Water Coupling Controls Class A GPCR Signal Transduction Pathways. *bioRxiv*, **2021**, 2020.08.28.271510. <https://doi.org/10.1101/2020.08.28.271510>.
- [86] Olivella, M.; Deupi, X.; Govaerts, C.; Pardo, L. Influence of the Environment in the Conformation of  $\alpha$ -Helices Studied by Protein Database Search and Molecular Dynamics Simulations. *Biophys. J.*, **2002**, *82* (6), 3207–3213. [https://doi.org/10.1016/s0006-3495\(02\)75663-4](https://doi.org/10.1016/s0006-3495(02)75663-4).
- [87] Ruprecht, J. J.; Mielke, T.; Vogel, R.; Villa, C.; Schertler, G. F. Electron Crystallography Reveals the Structure of Metarhodopsin I. *EMBO J.*, **2004**, *23* (18), 3609–3620. <https://doi.org/10.1038/sj.emboj.7600374>.

- [88] Shi, L.; Liapakis, G.; Xu, R.; Guarnieri, F.; Ballesteros, J. A.; Javitch, J. A. B2 Adrenergic Receptor Activation MODULATION OF THE PROLINE KINK IN TRANSMEMBRANE 6 BY A ROTAMER TOGGLE SWITCH. *J. Biol. Chem.*, **2002**, 277 (43), 40989–40996. <https://doi.org/10.1074/jbc.m206801200>.
- [89] Holst, B.; Nygaard, R.; Valentin-Hansen, L.; Bach, A.; Engelstoft, M. S.; Petersen, P. S.; Frimurer, T. M.; Schwartz, T. W. A Conserved Aromatic Lock for the Tryptophan Rotameric Switch in TM-VI of Seven-Transmembrane Receptors2. *J. Biol. Chem.*, **2010**, 285 (6), 3973–3985. <https://doi.org/10.1074/jbc.m109.064725>.
- [90] Pei, Y.; Mercier, R. W.; Anday, J. K.; Thakur, G. A.; Zvonok, A. M.; Hurst, D.; Reggio, P. H.; Janero, D. R.; Makriyannis, A. Ligand-Binding Architecture of Human CB2 Cannabinoid Receptor: Evidence for Receptor Subtype-Specific Binding Motif and Modeling GPCR Activation. *Chem. Biol.*, **2008**, 15 (11), 1207–1219. <https://doi.org/10.1016/j.chembiol.2008.10.011>.
- [91] Trzaskowski, B.; Latek, D.; Yuan, S.; Ghoshdastider, U.; Debinski, A.; Filipek, S. Action of Molecular Switches in GPCRs - Theoretical and Experimental Studies. *Curr Med Chem*, **2012**, 19 (8), 1090–1109. <https://doi.org/10.2174/092986712799320556>.
- [92] Hilger, D. The Role of Structural Dynamics in GPCR-mediated Signaling. *FEBS J.*, **2021**, 288 (8), 2461–2489. <https://doi.org/10.1111/febs.15841>.
- [93] Massink, A.; Gutiérrez-de-Terán, H.; Lenselink, E. B.; Zacarías, N. V. O.; Xia, L.; Heitman, L. H.; Katritch, V.; Stevens, R. C.; IJzerman, A. P. Sodium Ion Binding Pocket Mutations and Adenosine A2A Receptor Function. *Mol. Pharmacol.*, **2015**, 87 (2), 305–313. <https://doi.org/10.1124/mol.114.095737>.
- [94] Ragnarsson, L.; Andersson, Å.; Thomas, W. G.; Lewis, R. J. Mutations in the NPxxY Motif Stabilize Pharmacologically Distinct Conformational States of the A1B- and B2-Adrenoceptors. *Sci. Signal.*, **2019**, 12 (572). <https://doi.org/10.1126/scisignal.aas9485>.
- [95] Miao, Y.; Caliman, A. D.; McCammon, J. A. Allosteric Effects of Sodium Ion Binding on Activation of the M3 Muscarinic G-Protein-Coupled Receptor. *Biophys. J.*, **2015**, 108 (7), 1796–1806. <https://doi.org/10.1016/j.bpj.2015.03.003>.
- [96] Zhou, Q.; Yang, D.; Wu, M.; Guo, Y.; Guo, W.; Zhong, L.; Cai, X.; Dai, A.; Jang, W.; Shakhnovich, E. I.; et al. Common Activation Mechanism of Class A GPCRs. *Elife*, **2019**, 8, e50279. <https://doi.org/10.7554/elife.50279>.
- [97] Prioleau, C.; Visiers, I.; Ebersole, B. J.; Weinstein, H.; Sealfon, S. C. Conserved Helix 7 Tyrosine Acts as a Multistate Conformational Switch in the 5HT2C Receptor IDENTIFICATION OF A NOVEL “LOCKED-ON” PHENOTYPE AND DOUBLE REVERTANT MUTATIONS. *J. Biol. Chem.*, **2002**, 277 (39), 36577–36584. <https://doi.org/10.1074/jbc.m206223200>.
- [98] Bhattacharya, S.; Hall, S. E.; Vaidehi, N. Agonist-Induced Conformational Changes in Bovine Rhodopsin: Insight into Activation of G-Protein-Coupled Receptors. *J. Mol. Biol.*, **2008**, 382 (2), 539–555. <https://doi.org/10.1016/j.jmb.2008.06.084>.

- [99] Cohen, G. B.; Yang, T.; Robinson, P. R.; Oprian, D. D. Constitutive Activation of Opsin: Influence of Charge at Position 134 and Size at Position 296. *Biochemistry*, **1993**, 32 (23), 6111–6115. <https://doi.org/10.1021/bi00074a024>.
- [100] Springael, J.-Y.; Poorter, C. de; Deupi, X.; Durme, J. V.; Pardo, L.; Parmentier, M. The Activation Mechanism of Chemokine Receptor CCR5 Involves Common Structural Changes but a Different Network of Interhelical Interactions Relative to Rhodopsin. *Cell. Signal.*, **2007**, 19 (7), 1446–1456. <https://doi.org/10.1016/j.cellsig.2007.01.022>.
- [101] Mnpotra, J. S.; Qiao, Z.; Cai, J.; Lynch, D. L.; Grossfield, A.; Leioatts, N.; Hurst, D. P.; Pitman, M. C.; Song, Z.-H.; Reggio, P. H. Structural Basis of G Protein-Coupled Receptor-Gi Protein Interaction FORMATION OF THE CANNABINOID CB2 RECEPTOR-Gi PROTEIN COMPLEX. *J. Biol. Chem.*, **2014**, 289 (29), 20259–20272. <https://doi.org/10.1074/jbc.m113.539916>.
- [102] Han, X.; Tachado, S. D.; Koziel, H.; Boisvert, W. A. Leu1283.43 (L128) and Val2476.40 (V247) of CXCR1 Are Critical Amino Acid Residues for G Protein Coupling and Receptor Activation. *Plos One*, **2012**, 7 (8), e42765. <https://doi.org/10.1371/journal.pone.0042765>.
- [103] Tehan, B. G.; Bortolato, A.; Blaney, F. E.; Weir, M. P.; Mason, J. S. Unifying Family A GPCR Theories of Activation. *Pharmacol Therapeut*, **2014**, 143 (1), 51–60. <https://doi.org/10.1016/j.pharmthera.2014.02.004>.
- [104] Hauser, A. S.; Kooistra, A. J.; Munk, C.; Heydenreich, F. M.; Vepintsev, D. B.; Bouvier, M.; Babu, M. M.; Gloriam, D. E. GPCR Activation Mechanisms across Classes and Macro/Microscales. *Nat Struct Mol Biol*, **2021**, 28 (11), 879–888. <https://doi.org/10.1038/s41594-021-00674-7>.
- [105] Arnam, E. B. V.; Lester, H. A.; Dougherty, D. A. Dissecting the Functions of Conserved Prolines within Transmembrane Helices of the D2 Dopamine Receptor. *ACS Chem. Biol.*, **2011**, 6 (10), 1063–1068. <https://doi.org/10.1021/cb200153g>.
- [106] Venkatakrisnan, A. J.; Deupi, X.; Lebon, G.; Tate, C. G.; Schertler, G. F.; Babu, M. M. Molecular Signatures of G-Protein-Coupled Receptors. *Nature*, **2013**, 494 (7436), 185–194. <https://doi.org/10.1038/nature11896>.
- [107] Quade, B. N.; Parker, M. D.; Occhipinti, R. The Therapeutic Importance of Acid-Base Balance. *Biochem. Pharmacol.*, **2021**, 183, 114278. <https://doi.org/10.1016/j.bcp.2020.114278>.
- [108] Swietach, P.; Vaughan-Jones, R. D.; Harris, A. L.; Hulikova, A. The Chemistry, Physiology and Pathology of PH in Cancer. *Philos. Trans. R. Soc. B: Biol. Sci.*, **2014**, 369 (1638), 20130099. <https://doi.org/10.1098/rstb.2013.0099>.
- [109] Roux, B. Ion Channels and Ion Selectivity. *Essays Biochem.*, **2017**, 61 (2), 201–209. <https://doi.org/10.1042/ebc20160074>.

- [110] Dunbar, L. A.; Caplan, M. J. The Cell Biology of Ion Pumps: Sorting and Regulation. *Eur. J. Cell Biol.*, **2000**, *79* (8), 557–563. <https://doi.org/10.1078/0171-9335-00079>.
- [111] Lindskog, S. Structure and Mechanism of Carbonic Anhydrase. *Pharmacol. Ther.*, **1997**, *74* (1), 1–20. [https://doi.org/10.1016/s0163-7258\(96\)00198-2](https://doi.org/10.1016/s0163-7258(96)00198-2).
- [112] Ludwig, M.-G.; Vanek, M.; Guerini, D.; Gasser, J. A.; Jones, C. E.; Junker, U.; Hofstetter, H.; Wolf, R. M.; Seuwen, K. Proton-Sensing G-Protein-Coupled Receptors. *Nature*, **2003**, *425* (6953), 93–98. <https://doi.org/10.1038/nature01905>.
- [113] Mahnensmith, R. L.; Aronson, P. S. The Plasma Membrane Sodium-Hydrogen Exchanger and Its Role in Physiological and Pathophysiological Processes. *Circulation Research*, **1985**, *56*, 773–788. <https://doi.org/10.1161/01.res.56.6.773>.
- [114] Halestrap, A. P. The Monocarboxylate Transporter Family—Structure and Functional Characterization. *IUBMB Life*, **2012**, *64* (1), 1–9. <https://doi.org/10.1002/iub.573>.
- [115] Soleimani, M.; Burnham, C. E. Physiologic and Molecular Aspects of TheNa<sup>+</sup>:HCO<sub>3</sub><sup>-</sup> Cotransporter in Health and Disease Processes. *Kidney Int.*, **2000**, *57* (2), 371–384. <https://doi.org/10.1046/j.1523-1755.2000.00857.x>.
- [116] Sisignano, M.; Fischer, M. J. M.; Geisslinger, G. Proton-Sensing GPCRs in Health and Disease. *Cells*, **2021**, *10* (8), 2050. <https://doi.org/10.3390/cells10082050>.
- [117] Silva, P. H. I.; Wagner, C. A. Physiological Relevance of Proton-Activated GPCRs. *Pflugers Archiv*, **2022**, *474* (5), 487–504. <https://doi.org/10.1007/s00424-022-02671-1>.
- [118] Xu, Y.; Casey, G. Identification of Human OGR1, a Novel G Protein-Coupled Receptor That Maps to Chromosome 14. *Genomics*, **1996**, *35* (2), 397–402. <https://doi.org/10.1006/geno.1996.0377>.
- [119] Mahadevan, M. S.; Baird, S.; BAILLY, J. E.; Shutler, G. G.; Sabourin, L. A.; Tsilfidis, C.; Neville, C. E.; Narang, M.; Korneluk, R. G. Isolation of a Novel G Protein-Coupled Receptor (GPR4) Localized to Chromosome 19q13.3. *Genomics*, **1995**, *30* (1), 84–88. <https://doi.org/10.1006/geno.1995.0013>.
- [120] Kyaw, H.; Zeng, Z.; Su, K.; Fan, P.; Shell, B. K.; Carter, K. C.; Li, Y. Cloning, Characterization, and Mapping of Human Homolog of Mouse T-Cell Death-Associated Gene. *DNA Cell Biol.*, **1998**, *17* (6), 493–500. <https://doi.org/10.1089/dna.1998.17.493>.
- [121] Weng, Z.; Fluckiger, A.-C.; Nisitani, S.; Wahl, M. I.; Le, L. Q.; Hunter, C. A.; Fernal, A. A.; Beau, M. M. L.; Witte, O. N. A DNA Damage and Stress Inducible G Protein-Coupled Receptor Blocks Cells in G2/M. *Proc. Natl. Acad. Sci.*, **1998**, *95* (21), 12334–12339. <https://doi.org/10.1073/pnas.95.21.12334>.
- [122] Mashiko, M.; Kurosawa, A.; Tani, Y.; Tsuji, T.; Takeda, S. GPR31 and GPR151 Are Activated under Acidic Conditions. *J Biochem*, **2019**, *166* (4), 317–322. <https://doi.org/10.1093/jb/mvz042>.

- [123] Karlsson, M.; Zhang, C.; Méar, L.; Zhong, W.; Digre, A.; Katona, B.; Sjöstedt, E.; Butler, L.; Odeberg, J.; Dusart, P.; et al. A Single-Cell Type Transcriptomics Map of Human Tissues. *Sci. Adv.*, **2021**, 7 (31), eabh2169. <https://doi.org/10.1126/sciadv.abh2169>.
- [124] Uhlén, M.; Fagerberg, L.; Hallström, B. M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; Sjöstedt, E.; Asplund, A.; et al. Tissue-Based Map of the Human Proteome. *Science*, **2015**, 347 (6220), 1260419. <https://doi.org/10.1126/science.1260419>.
- [125] Xu, J.; Mathur, J.; Vessières, E.; Hammack, S.; Nonomura, K.; Favre, J.; Grimaud, L.; Petrus, M.; Francisco, A.; Li, J.; et al. GPR68 Senses Flow and Is Essential for Vascular Physiology. *Cell*, **2018**, 173 (3), 762-775.e16. <https://doi.org/10.1016/j.cell.2018.03.076>.
- [126] Ye, S.; Zhu, Y.; Zhong, D.; Song, X.; Li, J.; Xiao, F.; Huang, Z.; Zhang, W.; Wu, M.; Zhang, K.; et al. G Protein-Coupled Receptor GPR68 Inhibits Lymphocyte Infiltration and Contributes to Gender-Dependent Melanoma Growth. *Front. Oncol.*, **2023**, 13, 1202750. <https://doi.org/10.3389/fonc.2023.1202750>.
- [127] Marie, M. A.; Sanderlin, E. J.; Satturwar, S.; Hong, H.; Lertpiriyapong, K.; Donthi, D.; Yang, L. V. GPR65 (TDAG8) Inhibits Intestinal Inflammation and Colitis-Associated Colorectal Cancer Development in Experimental Mouse Models. *Biochim. Biophys. Acta (BBA) - Mol. Basis Dis.*, **2022**, 1868 (1), 166288. <https://doi.org/10.1016/j.bbadis.2021.166288>.
- [128] Tovar, G. M.; Bergsbaken, T. The Role of GPR132 in Regulating T Cell Function. *J. Immunol.*, **2023**, 210 (1\_Supplement), 239.13-239.13. <https://doi.org/10.4049/jimmunol.210.supp.239.13>.
- [129] Sato, K.; Tobo, A.; Mogi, C.; Tobo, M.; Yamane, N.; Tosaka, M.; Tomura, H.; Im, D.-S.; Okajima, F. The Protective Role of Proton-Sensing TDAG8 in the Brain Injury in a Mouse Ischemia Reperfusion Model. *Sci. Rep.*, **2020**, 10 (1), 17193. <https://doi.org/10.1038/s41598-020-74372-7>.
- [130] Matoba, K.; Yamashita, S.; Isaksen, T. J.; Yamashita, T. Proton-Sensing Receptor GPR132 Facilitates Migration of Astrocytes. *Neurosci. Res.*, **2021**, 170, 106–113. <https://doi.org/10.1016/j.neures.2020.10.001>.
- [131] Krewson, E. A.; Sanderlin, E. J.; Marie, M. A.; Akhtar, S. N.; Velcicky, J.; Loetscher, P.; Yang, L. V. The Proton-Sensing GPR4 Receptor Regulates Paracellular Gap Formation and Permeability of Vascular Endothelial Cells. *iScience*, **2020**, 23 (2), 100848. <https://doi.org/10.1016/j.isci.2020.100848>.
- [132] Giudici, L.; Velic, A.; Daryadel, A.; Bettoni, C.; Mohebbi, N.; Suply, T.; Seuwen, K.; Ludwig, M.-G.; Wagner, C. A. The Proton-Activated Receptor GPR4 Modulates Glucose Homeostasis by Increasing Insulin Sensitivity. *Cell. Physiol. Biochem.*, **2013**, 32 (5), 1403–1416. <https://doi.org/10.1159/000356578>.

- [133] Sanderlin, E.; Justus, C.; Krewson, E.; Yang, L. Emerging Roles for the PH-Sensing G Protein-Coupled Receptors in Response to Acidotic Stress. *Cell Heal. Cytoskelet.*, **2015**, 99. <https://doi.org/10.2147/chc.s60508>.
- [134] Rowe, J. B.; Kapolka, N. J.; Taghon, G. J.; Morgan, W. M.; Isom, D. G. The Evolution and Mechanism of GPCR Proton Sensing. *J Biological Chem*, **2020**, 296, 100167. <https://doi.org/10.1074/jbc.ra120.016352>.
- [135] Wang, J.-Q.; Kon, J.; Mogi, C.; Tobo, M.; Damirin, A.; Sato, K.; Komachi, M.; Malchinkhuu, E.; Murata, N.; Kimura, T.; et al. TDAG8 Is a Proton-Sensing and Psychosine-Sensitive G-Protein-Coupled Receptor. *J. Biol. Chem.*, **2004**, 279 (44), 45626–45633. <https://doi.org/10.1074/jbc.m406966200>.
- [136] Guo, Y.; Zhang, W.; Giroux, C.; Cai, Y.; Ekambaram, P.; Dilly, A.; Hsu, A.; Zhou, S.; Maddipati, K. R.; Liu, J.; et al. Identification of the Orphan G Protein-Coupled Receptor GPR31 as a Receptor for 12-(S)-Hydroxyeicosatetraenoic Acid. *J. Biol. Chem.*, **2011**, 286 (39), 33832–33840. <https://doi.org/10.1074/jbc.m110.216564>.
- [137] Kabarowski, J. H. S.; Zhu, K.; Le, L. Q.; Witte, O. N.; Xu, Y. Lysophosphatidylcholine as a Ligand for the Immunoregulatory Receptor G2A. *Science*, **2001**, 293 (5530), 702–705. <https://doi.org/10.1126/science.1061781>.
- [138] Xu, Y.; Zhu, K.; Hong, G.; Wu, W.; Baudhuin, L. M.; Xiao, Y.; Damron, D. S. Sphingosylphosphorylcholine Is a Ligand for Ovarian Cancer G-Protein-Coupled Receptor 1. *Nat Cell Biol*, **2000**, 2 (5), 261–267. <https://doi.org/10.1038/35010529>.
- [139] Zhu, K.; Baudhuin, L. M.; Hong, G.; Williams, F. S.; Cristina, K. L.; Kabarowski, J. H. S.; Witte, O. N.; Xu, Y. Sphingosylphosphorylcholine and Lysophosphatidylcholine Are Ligands for the G Protein-Coupled Receptor GPR4. *J. Biol. Chem.*, **2001**, 276 (44), 41325–41335. <https://doi.org/10.1074/jbc.m008057200>.
- [140] Ignatov, A.; Hermans-Borgmeyer, I.; Schaller, H. C. Cloning and Characterization of a Novel G-Protein-Coupled Receptor with Homology to Galanin Receptors. *Neuropharmacology*, **2004**, 46 (8), 1114–1120. <https://doi.org/10.1016/j.neuropharm.2004.02.004>.
- [141] Holmes, F. E.; Kerr, N.; Chen, Y.-J.; Vanderplank, P.; McArdle, C. A.; Wynick, D. Targeted Disruption of the Orphan Receptor Gpr151 Does Not Alter Pain-Related Behaviour despite a Strong Induction in Dorsal Root Ganglion Expression in a Model of Neuropathic Pain. *Mol. Cell. Neurosci.*, **2017**, 78, 35–40. <https://doi.org/10.1016/j.mcn.2016.11.010>.
- [142] Zaslavsky, A.; Singh, L. S.; Tan, H.; Ding, H.; Liang, Z.; Xu, Y. Homo- and Hetero-Dimerization of LPA/S1P Receptors, OGR1 and GPR4. *Biochimica Et Biophysica Acta Bba - Mol Cell Biology Lipids*, **2006**, 1761 (10), 1200–1212. <https://doi.org/10.1016/j.bbalip.2006.08.011>.

- [143] Felce, J. H.; MacRae, A.; Davis, S. J. Constraints on GPCR Heterodimerization Revealed by the Type-4 Induced-Association BRET Assay. *Biophys. J.*, **2019**, *116* (1), 31–41. <https://doi.org/10.1016/j.bpj.2018.09.034>.
- [144] Glitsch, M. D. Recent Advances in Acid Sensing by G Protein Coupled Receptors. *Pflügers Arch. - Eur. J. Physiol.*, **2024**, 1–11. <https://doi.org/10.1007/s00424-024-02919-Y>.
- [145] Kapolka, N. J.; Rowe, J. B.; Taghon, G. J.; Morgan, W. M.; O’Shea, C. R.; Isom, D. G. Proton-Gated Coincidence Detection Is a Common Feature of GPCR Signaling. *Proc. Natl. Acad. Sci.*, **2021**, *118* (28), e2100171118. <https://doi.org/10.1073/pnas.2100171118>.
- [146] Kapolka, N. J.; Taghon, G. J.; Rowe, J. B.; Morgan, W. M.; Enten, J. F.; Lambert, N. A.; Isom, D. G. DCyFIR: A High-Throughput CRISPR Platform for Multiplexed G Protein-Coupled Receptor Profiling and Ligand Discovery. *Proc National Acad Sci*, **2020**, *117* (23), 13117–13126. <https://doi.org/10.1073/pnas.2000430117>.
- [147] Ghanouni, P.; Schambye, H.; Seifert, R.; Lee, T. W.; Rasmussen, S. G. F.; Gether, U.; Kobilka, B. K. The Effect of PH on B2 Adrenoceptor Function EVIDENCE FOR PROTONATION-DEPENDENT ACTIVATION. *J. Biol. Chem.*, **2000**, *275* (5), 3121–3127. <https://doi.org/10.1074/jbc.275.5.3121>.
- [148] Quinn, S. J.; Bai, M.; Brown, E. M. PH Sensing by the Calcium-Sensing Receptor\*. *J. Biol. Chem.*, **2004**, *279* (36), 37241–37249. <https://doi.org/10.1074/jbc.m404520200>.
- [149] Banani, H. E.; Bernard, M.; Baetz, D.; Cabanes, E.; Cozzone, P.; Lucien, A.; Feuvray, D. Changes in Intracellular Sodium and PH during Ischaemia–Reperfusion Are Attenuated by Trimetazidine. *Cardiovasc. Res.*, **2000**, *47* (4), 688–696. [https://doi.org/10.1016/s0008-6363\(00\)00136-x](https://doi.org/10.1016/s0008-6363(00)00136-x).
- [150] Milliken, A. S.; Ciesla, J. H.; Nadtochiy, S. M.; Brookes, P. S. Distinct Effects of Intracellular vs. Extracellular Acidic PH on the Cardiac Metabolome during Ischemia and Reperfusion. *J. Mol. Cell. Cardiol.*, **2023**, *174*, 101–114. <https://doi.org/10.1016/j.yjmcc.2022.11.008>.
- [151] Tyrtysnaia, A. A.; Lysenko, L. V.; Madamba, F.; Manzhulo, I. V.; Khotimchenko, M. Y.; Kleschevnikov, A. M. Acute Neuroinflammation Provokes Intracellular Acidification in Mouse Hippocampus. *J. Neuroinflammation*, **2016**, *13* (1), 283. <https://doi.org/10.1186/s12974-016-0747-8>.
- [152] Isom, D. G.; Sridharan, V.; Baker, R.; Clement, S. T.; Smalley, D. M.; Dohlman, H. G. Protons as Second Messenger Regulators of G Protein Signaling. *Mol. Cell*, **2013**, *51* (4), 531–538. <https://doi.org/10.1016/j.molcel.2013.07.012>.
- [153] Yamashiro, D. J.; Fluss, S. R.; Maxfield, F. R. Acidification of Endocytic Vesicles by an ATP-Dependent Proton Pump. *J. cell Biol.*, **1983**, *97* (3), 929–934. <https://doi.org/10.1083/jcb.97.3.929>.

- [154] Gidon, A.; Al-Bataineh, M. M.; Jean-Alphonse, F. G.; Stevenson, H. P.; Watanabe, T.; Louet, C.; Khatri, A.; Calero, G.; Pastor-Soler, N. M.; Gardella, T. J.; et al. Endosomal GPCR Signaling Turned off by Negative Feedback Actions of PKA and V-ATPase. *Nat. Chem. Biol.*, **2014**, *10* (9), 707–709. <https://doi.org/10.1038/nchembio.1589>.
- [155] Feinstein, T. N.; Yui, N.; Webber, M. J.; Wehbi, V. L.; Stevenson, H. P.; King, J. D.; Hallows, K. R.; Brown, D.; Bouley, R.; Vilaradaga, J.-P. Noncanonical Control of Vasopressin Receptor Type 2 Signaling by Retromer and Arrestin. *J. Biol. Chem.*, **2013**, *288* (39), 27849–27860. <https://doi.org/10.1074/jbc.m112.445098>.
- [156] Wehbi, V. L.; Stevenson, H. P.; Feinstein, T. N.; Calero, G.; Romero, G.; Vilaradaga, J.-P. Noncanonical GPCR Signaling Arising from a PTH Receptor–Arrestin–G $\beta\gamma$  Complex. *Proc. Natl. Acad. Sci.*, **2013**, *110* (4), 1530–1535. <https://doi.org/10.1073/pnas.1205756110>.
- [157] Irannejad, R.; Tomshine, J. C.; Tomshine, J. R.; Chevalier, M.; Mahoney, J. P.; Steyaert, J.; Rasmussen, S. G. F.; Sunahara, R. K.; El-Samad, H.; Huang, B.; et al. Conformational Biosensors Reveal GPCR Signalling from Endosomes. *Nature*, **2013**, *495* (7442), 534–538. <https://doi.org/10.1038/nature12000>.
- [158] Thomsen, A. R. B.; Plouffe, B.; Cahill, T. J.; Shukla, A. K.; Tarrasch, J. T.; Dosey, A. M.; Kahsai, A. W.; Strachan, R. T.; Pani, B.; Mahoney, J. P.; et al. GPCR-G Protein- $\beta$ -Arrestin Super-Complex Mediates Sustained G Protein Signaling. *Cell*, **2016**, *166* (4), 907–919. <https://doi.org/10.1016/j.cell.2016.07.004>.
- [159] Thomsen, A. R. B.; Jensen, D. D.; Hicks, G. A.; Bunnett, N. W. Therapeutic Targeting of Endosomal G-Protein-Coupled Receptors. *Trends Pharmacol. Sci.*, **2018**, *39* (10), 879–891. <https://doi.org/10.1016/j.tips.2018.08.003>.
- [160] Fasciani, I.; Carli, M.; Petragano, F.; Colaianni, F.; Aloisi, G.; Maggio, R.; Scarselli, M.; Rossi, M. GPCRs in Intracellular Compartments: New Targets for Drug Discovery. *Biomolecules*, **2022**, *12* (10), 1343. <https://doi.org/10.3390/biom12101343>.
- [161] Meyer, J.; Vecchio, G. D.; Seitz, V.; Massaly, N.; Stein, C. Modulation of M-opioid Receptor Activation by Acidic PH Is Dependent on Ligand Structure and an Ionizable Amino Acid Residue. *Br. J. Pharmacol.*, **2019**, *176* (23), 4510–4520. <https://doi.org/10.1111/bph.14810>.
- [162] Arunan, E.; Desiraju, G. R.; Klein, R. A.; Sadlej, J.; Scheiner, S.; Alkorta, I.; Clary, D. C.; Crabtree, R. H.; Dannenberg, J. J.; Hobza, P.; et al. Definition of the Hydrogen Bond (IUPAC Recommendations 2011). *Pure Appl. Chem.*, **2011**, *83* (8), 1637–1641. <https://doi.org/10.1351/pac-rec-10-01-02>.
- [163] Bertalan, É.; Lešnik, S.; Bren, U.; Bondar, A.-N. Protein-Water Hydrogen-Bond Networks of G Protein-Coupled Receptors: Graph-Based Analyses of Static Structures and Molecular Dynamics. *J. Struct. Biol.*, **2020**, *212* (3), 107634. <https://doi.org/10.1016/j.jsb.2020.107634>.

- [164] Venkatakrishnan, A. J.; Ma, A. K.; Fonseca, R.; Latorraca, N. R.; Kelly, B.; Betz, R. M.; Asawa, C.; Kobilka, B. K.; Dror, R. O. Diverse GPCRs Exhibit Conserved Water Networks for Stabilization and Activation. *P Natl Acad Sci Usa*, **2019**, *116* (8), 3288–3293. <https://doi.org/10.1073/pnas.1809251116>.
- [165] Zhang, X. C.; Sun, K.; Zhang, L.; Li, X.; Cao, C. GPCR Activation: Protonation and Membrane Potential. *Protein Cell*, **2013**, *4* (10), 747–760. <https://doi.org/10.1007/s13238-013-3073-2>.
- [166] Ishikita, H.; Saito, K. Proton Transfer Reactions and Hydrogen-Bond Networks in Protein Environments. *J Roy Soc Interface*, **2014**, *11* (91), 20130518. <https://doi.org/10.1098/rsif.2013.0518>.
- [167] Zhang, X. C.; Cao, C.; Zhou, Y.; Zhao, Y. Proton Transfer-Mediated GPCR Activation. *Protein Cell*, **2015**, *6* (1), 12–17. <https://doi.org/10.1007/s13238-014-0106-4>.
- [168] Wilson, M. H.; Highfield, H. A.; Limbird, L. E. The Role of a Conserved Inter-Transmembrane Domain Interface in Regulating A2a-Adrenergic Receptor Conformational Stability and Cell-Surface Turnover. *Mol. Pharmacol.*, **2001**, *59* (4), 929–938. <https://doi.org/10.1124/mol.59.4.929>.
- [169] Vickery, O. N.; Carvalheda, C. A.; Zaidi, S. A.; Pislakov, A. V.; Katritch, V.; Zachariae, U. Intracellular Transfer of Na<sup>+</sup> in an Active-State G-Protein-Coupled Receptor. *Structure*, **2018**, *26* (1), 171-180.e2. <https://doi.org/10.1016/j.str.2017.11.013>.
- [170] Wiley, S. Z.; Sriram, K.; Salmerón, C.; Insel, P. A. GPR68: An Emerging Drug Target in Cancer. *Int J Mol Sci*, **2019**, *20* (3), 559. <https://doi.org/10.3390/ijms20030559>.
- [171] 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-us*, **2020**, *59* (38), 3594–3614. <https://doi.org/10.1021/acs.biochem.0c00576>.
- [172] Maeyashiki, C.; Melhem, H.; Hering, L.; Baebler, K.; Cosin-Roger, J.; Schefer, F.; Weder, B.; Hausmann, M.; Scharl, M.; Rogler, G.; et al. Activation of PH-Sensing Receptor OGR1 (GPR68) Induces ER Stress Via the IRE1 $\alpha$ /JNK Pathway in an Intestinal Epithelial Cell Model. *Sci. Rep.*, **2020**, *10* (1), 1438. <https://doi.org/10.1038/s41598-020-57657-9>.
- [173] Matsuzaki, S.; Ishizuka, T.; Yamada, H.; Kamide, Y.; Hisada, T.; Ichimonji, I.; Aoki, H.; Yatomi, M.; Komachi, M.; Tsurumaki, H.; et al. Extracellular Acidification Induces Connective Tissue Growth Factor Production through Proton-Sensing Receptor OGR1 in Human Airway Smooth Muscle Cells. *Biochem Bioph Res Co*, **2011**, *413* (4), 499–503. <https://doi.org/10.1016/j.bbrc.2011.08.087>.
- [174] Chandra, V.; Karamitri, A.; Richards, P.; Cormier, F.; Ramond, C.; Jockers, R.; Armanet, M.; Albagli-Curiel, O.; Scharfmann, R. Extracellular Acidification Stimulates GPR68 Mediated IL-8 Production in Human Pancreatic  $\beta$  Cells. *Sci. Rep.*, **2016**, *6* (1), 25765. <https://doi.org/10.1038/srep25765>.

- [175] Huang, X.-P.; Karpiak, J.; Kroeze, W. K.; Zhu, H.; Chen, X.; Moy, S. S.; Sadoris, K. A.; Nikolova, V. D.; Farrell, M. S.; Wang, S.; et al. Allosteric Ligands for the Pharmacologically Dark Receptors GPR68 and GPR65. *Nature*, **2015**, 527 (7579), 477–483. <https://doi.org/10.1038/nature15699>.
- [176] Wang, J.; Sun, Y.; Tomura, H.; Okajima, F. Ovarian Cancer G-Protein-Coupled Receptor 1 Induces the Expression of the Pain Mediator Prostaglandin E2 in Response to an Acidic Extracellular Environment in Human Osteoblast-like Cells. *Int. J. Biochem. Cell Biol.*, **2012**, 44 (11), 1937–1941. <https://doi.org/10.1016/j.biocel.2012.07.015>.
- [177] Singh, L. S.; Berk, M.; Oates, R.; Zhao, Z.; Tan, H.; Jiang, Y.; Zhou, A.; Kirmani, K.; Steinmetz, R.; Lindner, D.; et al. Ovarian Cancer G Protein–Coupled Receptor 1, a New Metastasis Suppressor Gene in Prostate Cancer. *Jnci J National Cancer Inst*, **2007**, 99 (17), 1313–1327. <https://doi.org/10.1093/jnci/djm107>.
- [178] Li, J.; Guo, B.; Wang, J.; Cheng, X.; Xu, Y.; Sang, J. Ovarian Cancer G Protein Coupled Receptor 1 Suppresses Cell Migration of MCF7 Breast Cancer Cells via a Gα12/13 -Rho-Rac1 Pathway. *J. Mol. Signal.*, **2013**, 8 (0), Art. 6. <https://doi.org/10.1186/1750-2187-8-6>.
- [179] Afrasiabi, E.; Blom, T.; Ekokoski, E.; Tuominen, R. K.; Törnquist, K. Sphingosylphosphorylcholine Enhances Calcium Entry in Thyroid FRO Cells by a Mechanism Dependent on Protein Kinase C. *Cell. Signal.*, **2006**, 18 (10), 1671–1678. <https://doi.org/10.1016/j.cellsig.2006.01.005>.
- [180] Li, H.; Wang, D.; Singh, L. S.; Berk, M.; Tan, H.; Zhao, Z.; Steinmetz, R.; Kirmani, K.; Wei, G.; Xu, Y. Abnormalities in Osteoclastogenesis and Decreased Tumorigenesis in Mice Deficient for Ovarian Cancer G Protein-Coupled Receptor 1. *PLoS ONE*, **2009**, 4 (5), e5705. <https://doi.org/10.1371/journal.pone.0005705>.
- [181] Murata, N.; Mogi, C.; Tobo, M.; Nakakura, T.; Sato, K.; Tomura, H.; Okajima, F. Inhibition of Superoxide Anion Production by Extracellular Acidification in Neutrophils. *Cell. Immunol.*, **2009**, 259 (1), 21–26. <https://doi.org/10.1016/j.cellimm.2009.05.008>.
- [182] McAleer, J. P.; Fan, J.; Roar, B.; Primerano, D. A.; Denvir, J. Cytokine Regulation in Human CD4 T Cells by the Aryl Hydrocarbon Receptor and Gq-Coupled Receptors. *Sci. Rep.*, **2018**, 8 (1), 10954. <https://doi.org/10.1038/s41598-018-29262-4>.
- [183] Tomura, H.; Wang, J.-Q.; Komachi, M.; Damirin, A.; Mogi, C.; Tobo, M.; Kon, J.; Misawa, N.; Sato, K.; Okajima, F. Prostaglandin I2 Production and cAMP Accumulation in Response to Acidic Extracellular PH through OGR1 in Human Aortic Smooth Muscle Cells. *J. Biol. Chem.*, **2005**, 280 (41), 34458–34464. <https://doi.org/10.1074/jbc.m505287200>.
- [184] Saxena, H.; Deshpande, D.; Tiegs, B.; Yan, H.; Battafarano, R.; Burrows, W.; Damera, G.; Panettieri, R.; Jr, T. D.; An, S.; et al. The GPCR OGR1 (GPR68) Mediates Diverse Signalling and Contraction of Airway Smooth Muscle in Response to Small Reductions in Extracellular PH. *Br. J. Pharmacol.*, **2012**, 166 (3), 981–990. <https://doi.org/10.1111/j.1476-5381.2011.01807.x>.

- [185] Wiley, S. Z.; Sriram, K.; Liang, W.; Chang, S. E.; French, R.; McCann, T.; Sicklick, J.; Nishihara, H.; Lowy, A. M.; Insel, P. A. GPR68, a Proton-sensing GPCR, Mediates Interaction of Cancer-associated Fibroblasts and Cancer Cells. *FASEB J.*, **2018**, *32* (3), 1170–1183. <https://doi.org/10.1096/fj.201700834r>.
- [186] Horman, S. R.; To, J.; Lamb, J.; Zoll, J. H.; Leonetti, N.; Tu, B.; Moran, R.; Newlin, R.; Walker, J. R.; Orth, A. P. Functional Profiling of Microtumors to Identify Cancer Associated Fibroblast-Derived Drug Targets. *Oncotarget*, **2017**, *8* (59), 99913–99930. <https://doi.org/10.18632/oncotarget.21915>.
- [187] Perren, L.; Busch, M.; Schuler, C.; Ruiz, P. A.; Foti, F.; Weibel, N.; Vallière, C. de; Morsy, Y.; Seuwen, K.; Hausmann, M.; et al. OGR1 (GPR68) and TDAG8 (GPR65) Have Antagonistic Effects in Models of Colonic Inflammation. *Int. J. Mol. Sci.*, **2023**, *24* (19), 14855. <https://doi.org/10.3390/ijms241914855>.
- [188] Nagel, D. J.; Rackow, A. R.; Ku, W.-Y.; Bell, T. J.; Sime, P. J.; Kottmann, R. M. Cell-Type-Specific Effects of the Ovarian Cancer G-Protein Coupled Receptor (OGR1) on Inflammation and Fibrosis; Potential Implications for Idiopathic Pulmonary Fibrosis. *Cells*, **2022**, *11* (16), 2540. <https://doi.org/10.3390/cells11162540>.
- [189] Parry, D. A.; Smith, C. E. L.; El-Sayed, W.; Poulter, J. A.; Shore, R. C.; Logan, C. V.; Mogi, C.; Sato, K.; Okajima, F.; Harada, A.; et al. Mutations in the PH-Sensing G-Protein-Coupled Receptor GPR68 Cause Amelogenesis Imperfecta. *Am J Hum Genetics*, **2016**, *99* (4), 984–990. <https://doi.org/10.1016/j.ajhg.2016.08.020>.
- [190] 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*, **2020**, *59* (38), 3594–3614. <https://doi.org/10.1021/acs.biochem.0c00576>.
- [191] Wei, W.-C.; Bianchi, F.; Wang, Y.-K.; Tang, M.-J.; Ye, H.; Glitsch, M. D. Coincidence Detection of Membrane Stretch and Extracellular PH by the Proton-Sensing Receptor OGR1 (GPR68). *Curr Biol*, **2018**, *28* (23), 3815–3823.e4. <https://doi.org/10.1016/j.cub.2018.10.046>.
- [192] Iliff, A. J.; Xu, X. Z. S. A Mechanosensitive GPCR That Detects the Bloody Force. *Cell*, **2018**, *173* (3), 542–544. <https://doi.org/10.1016/j.cell.2018.04.001>.
- [193] Marullo, S.; Doly, S.; Saha, K.; Enslin, H.; Scott, M. G. H.; Coureuil, M. Mechanical GPCR Activation by Traction Forces Exerted on Receptor N-Glycans. *ACS Pharmacol. Transl. Sci.*, **2020**, *3* (2), 171–178. <https://doi.org/10.1021/acsptsci.9b00106>.
- [194] Wilde, C.; Mitgau, J.; Suchý, T.; Schoeneberg, T.; Liebscher, I. Translating the Force - Mechano-Sensing GPCRs. *Am J Physiol-cell Ph*, **2022**. <https://doi.org/10.1152/ajpcell.00465.2021>.
- [195] Virion, Z.; Doly, S.; Saha, K.; Lambert, M.; Guillonneau, F.; Bied, C.; Duke, R. M.; Rudd, P. M.; Robbe-Masselot, C.; Nassif, X.; et al. Sialic Acid Mediated Mechanical

Activation of B2 Adrenergic Receptors by Bacterial Pili. *Nat. Commun.*, **2019**, *10* (1), 4752. <https://doi.org/10.1038/s41467-019-12685-6>.

[196] Erdogmus, S.; Storch, U.; Danner, L.; Becker, J.; Winter, M.; Ziegler, N.; Wirth, A.; Offermanns, S.; Hoffmann, C.; Gudermann, T.; et al. Helix 8 Is the Essential Structural Motif of Mechanosensitive GPCRs. *Nat. Commun.*, **2019**, *10* (1), 5784. <https://doi.org/10.1038/s41467-019-13722-0>.

[197] Abe-Ohya, R.; Ishikawa, T.; Shiozawa, H.; Suda, K.; Nara, F. Identification of Metals from Osteoblastic ST-2 Cell Supernatants as Novel OGR1 Agonists. *J Recept Sig Transd*, **2015**, *35* (5), 485–492. <https://doi.org/10.3109/10799893.2015.1015736>.

[198] Mochimaru, Y.; Negishi, J.; Murakami, S.; Musha, S.; Sato, K.; Okajima, F.; Tomura, H. Metals Differentially Activate Ovarian Cancer G Protein-Coupled Receptor 1 in Various Species. *Zool Sci*, **2018**, *35* (2), 109–114. <https://doi.org/10.2108/zs170145>.

[199] Pera, T.; Deshpande, D. A.; Ippolito, M.; Wang, B.; Gavrila, A.; Michael, J. V.; Nayak, A. P.; Tompkins, E.; Farrell, E.; Kroeze, W. K.; et al. Biased Signaling of the Proton-sensing Receptor OGR1 by Benzodiazepines. *Faseb J*, **2018**, *32* (2), 862–874. <https://doi.org/10.1096/fj.201700555r>.

[200] Yu, X.; Huang, X.-P.; Kenakin, T. P.; Slocum, S. T.; Chen, X.; Martini, M. L.; Liu, J.; Jin, J. Design, Synthesis, and Characterization of Ogerin-Based Positive Allosteric Modulators for G Protein-Coupled Receptor 68 (GPR68). *J Med Chem*, **2019**, *62* (16), 7557–7574. <https://doi.org/10.1021/acs.jmedchem.9b00869>.

[201] Fukuda, H.; Ito, S.; Watari, K.; Mogi, C.; Arisawa, M.; Okajima, F.; Kurose, H.; Shuto, S. Identification of a Potent and Selective GPR4 Antagonist as a Drug Lead for the Treatment of Myocardial Infarction. *ACS Med. Chem. Lett.*, **2016**, *7* (5), 493–497. <https://doi.org/10.1021/acsmchemlett.6b00014>.

[202] Yamanaka, T.; Harimoto, N.; Yokobori, T.; Muranushi, R.; Hoshino, K.; Hagiwara, K.; Gantumur, D.; Handa, T.; Ishii, N.; Tsukagoshi, M.; et al. Conophylline Inhibits Hepatocellular Carcinoma by Inhibiting Activated Cancer-Associated Fibroblasts Through Suppression of G Protein-Coupled Receptor 68. *Mol Cancer Ther*, **2021**, *20* (6), 1019–1028. <https://doi.org/10.1158/1535-7163.mct-20-0150>.

[203] Yoshida, Y.; Fukuoka, K.; Sakugawa, M.; Kurogi, M.; Hamamura, K.; Hamasaki, K.; Tsurusaki, F.; Sotono, K.; Nishi, T.; Fukuda, T.; et al. Inhibition of G Protein-Coupled Receptor 68 Using Homoharringtonine Attenuates Chronic Kidney Disease-Associated Cardiac Impairment. *Transl. Res.*, **2024**. <https://doi.org/10.1016/j.trsl.2024.02.004>.

[204] Williams, C. H.; Neitzel, L. R.; Cornell, J.; Rea, S.; Mills, I.; Silver, M. S.; Ahmad, J. D.; Birukov, K. G.; Birukova, A.; Brem, H.; et al. GPR68-ATF4 Signaling Is a Novel Prosurvival Pathway in Glioblastoma Activated by Acidic Extracellular Microenvironment. *Exp. Hematol. Oncol.*, **2024**, *13* (1), 13. <https://doi.org/10.1186/s40164-023-00468-1>.

- [205] Addis, P.; Bali, U.; Baron, F.; Campbell, A.; Harborne, S.; Jagger, L.; Milne, G.; Pearce, M.; Rosethorne, E. M.; Satchell, R.; et al. Key Aspects of Modern GPCR Drug Discovery. *SLAS Discov.*, **2023**, *29* (1), 1–22. <https://doi.org/10.1016/j.slasd.2023.08.007>.
- [206] Zhang, R.; Xie, X. Tools for GPCR Drug Discovery. *Acta Pharmacol. Sin.*, **2012**, *33* (3), 372–384. <https://doi.org/10.1038/aps.2011.173>.
- [207] Becker, O. M.; Marantz, Y.; Shacham, S.; Inbal, B.; Heifetz, A.; Kalid, O.; Bar-Haim, S.; Warshaviak, D.; Fichman, M.; Noiman, S. G Protein-Coupled Receptors: In Silico Drug Discovery in 3D. *Proc. Natl. Acad. Sci.*, **2004**, *101* (31), 11304–11309. <https://doi.org/10.1073/pnas.0401862101>.
- [208] Congreve, M.; Graaf, C. de; Swain, N. A.; Tate, C. G. Impact of GPCR Structures on Drug Discovery. *Cell*, **2020**, *181* (1), 81–91. <https://doi.org/10.1016/j.cell.2020.03.003>.
- [209] Congreve, M.; Oswald, C.; Marshall, F. H. Applying Structure-Based Drug Design Approaches to Allosteric Modulators of GPCRs. *Trends Pharmacol. Sci.*, **2017**, *38* (9), 837–847. <https://doi.org/10.1016/j.tips.2017.05.010>.
- [210] Tarcsay, A.; Paragi, G.; Vass, M.; Jójárt, B.; Bogár, F.; Keserű, G. M. The Impact of Molecular Dynamics Sampling on the Performance of Virtual Screening against GPCRs. *J. Chem. Inf. Model.*, **2013**, *53* (11), 2990–2999. <https://doi.org/10.1021/ci400087b>.
- [211] Pandey, A.; Shin, K.; Patterson, R. E.; Liu, X.-Q.; Rainey, J. K. Current Strategies for Protein Production and Purification Enabling Membrane Protein Structural Biology 1. *Biochem. Cell Biol.*, **2016**, *94* (6), 507–527. <https://doi.org/10.1139/bcb-2015-0143>.
- [212] Ribeiro, R. P.; Giorgetti, A. PyGOMoDo: GPCRs Modeling and Docking with Python. *Bioinformatics*, **2023**, *39* (5), btad294. <https://doi.org/10.1093/bioinformatics/btad294>.
- [213] Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Židek, A.; Potapenko, A.; et al. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature*, **2021**, *596* (7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
- [214] Chen, W.; Song, C.; Leng, L.; Zhang, S.; Chen, S. The Application of Artificial Intelligence Accelerates G Protein-Coupled Receptor Ligand Discovery. *Engineering*, **2023**. <https://doi.org/10.1016/j.eng.2023.09.011>.
- [215] Sykes, D. A.; Stoddart, L. A.; Kilpatrick, L. E.; Hill, S. J. Binding Kinetics of Ligands Acting at GPCRs. *Mol. Cell. Endocrinol.*, **2019**, *485*, 9–19. <https://doi.org/10.1016/j.mce.2019.01.018>.
- [216] Kunig, V.; Potowski, M.; Gohla, A.; Brunschweiler, A. DNA-Encoded Libraries – an Efficient Small Molecule Discovery Technology for the Biomedical Sciences. *Biol. Chem.*, **2018**, *399* (7), 691–710. <https://doi.org/10.1515/hsz-2018-0119>.

- [217] Locatelli-Hoops, S.; Yeliseev, A. A.; Gawrisch, K.; Gorshkova, I. Surface Plasmon Resonance Applied to G Protein-Coupled Receptors. *Biomed. Spectrosc. Imaging*, **2013**, *2* (3), 155–181. <https://doi.org/10.3233/bsi-130045>.
- [218] Ward, R. J.; Milligan, G. Structural and Biophysical Characterisation of G Protein-Coupled Receptor Ligand Binding Using Resonance Energy Transfer and Fluorescent Labelling Techniques. *Biochim. Biophys. Acta (BBA) - Biomembr.*, **2014**, *1838* (1), 3–14. <https://doi.org/10.1016/j.bbamem.2013.04.007>.
- [219] Uddin, R.; Simms, J.; Poyner, D. Functional Characterisation of G Protein-Coupled Receptors. *Methods*, **2018**, *147*, 213–220. <https://doi.org/10.1016/j.ymeth.2018.02.018>.
- [220] Jaakola, V.-P.; Prilusky, J.; Sussman, J. L.; Goldman, A. G Protein-Coupled Receptors Show Unusual Patterns of Intrinsic Unfolding. *Protein Eng. Des. Sel.*, **2005**, *18* (2), 103–110. <https://doi.org/10.1093/protein/gzi004>.
- [221] Heng, J.; Hu, Y.; Pérez-Hernández, G.; Inoue, A.; Zhao, J.; Ma, X.; Sun, X.; Kawakami, K.; Ikuta, T.; Ding, J.; et al. Function and Dynamics of the Intrinsically Disordered Carboxyl Terminus of B2 Adrenergic Receptor. *Nat. Commun.*, **2023**, *14* (1), 2005. <https://doi.org/10.1038/s41467-023-37233-1>.
- [222] Jakubík, J.; El-Fakahany, E. E. Allosteric Modulation of GPCRs of Class A by Cholesterol. *Int. J. Mol. Sci.*, **2021**, *22* (4), 1953. <https://doi.org/10.3390/ijms22041953>.
- [223] Grisshammer, R.; Duckworth, R.; Henderson, R. Expression of a Rat Neurotensin Receptor in Escherichia Coli. *Biochem. J.*, **1993**, *295* (2), 571–576. <https://doi.org/10.1042/bj2950571>.
- [224] Weiß, H. M.; Grisshammer, R. Purification and Characterization of the Human Adenosine A2a Receptor Functionally Expressed in Escherichia Coli. *Eur. J. Biochem.*, **2002**, *269* (1), 82–92. <https://doi.org/10.1046/j.0014-2956.2002.02618.x>.
- [225] Stanasila, L.; Massotte, D.; Kieffer, B. L.; Pattus, F. Expression of  $\delta$ ,  $\kappa$  and  $\mu$  Human Opioid Receptors in Escherichia Coli and Reconstitution of the High-affinity State for Agonist with Heterotrimeric G Proteins. *Eur. J. Biochem.*, **1999**, *260* (2), 430–438. <https://doi.org/10.1046/j.1432-1327.1999.00187.x>.
- [226] Wiseman, D. N.; Otchere, A.; Patel, J. H.; Uddin, R.; Pollock, N. L.; Routledge, S. J.; Rothnie, A. J.; Slack, C.; Poyner, D. R.; Bill, R. M.; et al. Expression and Purification of Recombinant G Protein-Coupled Receptors: A Review. *Protein Express Purif*, **2020**, *167*, 105524. <https://doi.org/10.1016/j.pep.2019.105524>.
- [227] Saarenpää, T.; Jaakola, V.-P.; Goldman, A. Chapter Nine Baculovirus-Mediated Expression of GPCRs in Insect Cells. *Methods Enzym.*, **2015**, *556*, 185–218. <https://doi.org/10.1016/bs.mie.2014.12.033>.
- [228] McKenzie, E. A.; Abbott, W. M. Expression of Recombinant Proteins in Insect and Mammalian Cells. *Methods*, **2018**, *147*, 40–49. <https://doi.org/10.1016/j.ymeth.2018.05.013>.

- [229] Milić, D.; Veprintsev, D. B. Large-Scale Production and Protein Engineering of G Protein-Coupled Receptors for Structural Studies. *Front Pharmacol*, **2015**, *6*, 66. <https://doi.org/10.3389/fphar.2015.00066>.
- [230] Thomas, J. A.; Tate, C. G. Quality Control in Eukaryotic Membrane Protein Overproduction. *J. Mol. Biol.*, **2014**, *426* (24), 4139–4154. <https://doi.org/10.1016/j.jmb.2014.10.012>.
- [231] Malm, M.; Saghaleyni, R.; Lundqvist, M.; Giudici, M.; Chotteau, V.; Field, R.; Varley, P. G.; Hatton, D.; Grassi, L.; Svensson, T.; et al. Evolution from Adherent to Suspension: Systems Biology of HEK293 Cell Line Development. *Sci. Rep.*, **2020**, *10* (1), 18996. <https://doi.org/10.1038/s41598-020-76137-8>.
- [232] Chakraborty, R.; Xu, B.; Bhullar, R. P.; Chelikani, P. Chapter Twelve Expression of G Protein-Coupled Receptors in Mammalian Cells. *Methods Enzymol.*, **2015**, *556*, 267–281. <https://doi.org/10.1016/bs.mie.2014.12.013>.
- [233] Guan, X. M.; Kobilka, T. S.; Kobilka, B. K. Enhancement of Membrane Insertion and Function in a Type IIIb Membrane Protein Following Introduction of a Cleavable Signal Peptide. *J. Biol. Chem.*, **1992**, *267* (31), 21995–21998.
- [234] Possee, R. D.; Howard, S. C. Analysis of the Polyhedrin Gene Promoter of the Autographa Californica Nuclear Polyhedrosis Virus. *Nucleic Acids Res.*, **1987**, *15* (24), 10233–10248. <https://doi.org/10.1093/nar/15.24.10233>.
- [235] Johari, Y. B.; Scarrott, J. M.; Pohle, T. H.; Liu, P.; Mayer, A.; Brown, A. J.; James, D. C. Engineering of the CMV Promoter for Controlled Expression of Recombinant Genes in HEK293 Cells. *Biotechnol. J.*, **2022**, *17* (8), e2200062. <https://doi.org/10.1002/biot.202200062>.
- [236] Chun, E.; Thompson, A. A.; Liu, W.; Roth, C. B.; Griffith, M. T.; Katritch, V.; Kunken, J.; Xu, F.; Cherezov, V.; Hanson, M. A.; et al. Fusion Partner Toolchest for the Stabilization and Crystallization of G Protein-Coupled Receptors. *Structure*, **2012**, *20* (6), 967–976. <https://doi.org/10.1016/j.str.2012.04.010>.
- [237] Zhang, K.; Wu, H.; Hoppe, N.; Manglik, A.; Cheng, Y. Fusion Protein Strategies for Cryo-EM Study of G Protein-Coupled Receptors. *Nat Commun*, **2022**, *13* (1), 4366. <https://doi.org/10.1038/s41467-022-32125-2>.
- [238] Xu, J.; Chen, G.; Wang, H.; Cao, S.; Heng, J.; Deupi, X.; Du, Y.; Kobilka, B. K. Calcineurin-Fusion Facilitates Cryo-EM Structure Determination of a Family A GPCR. *bioRxiv*, **2022**, 2022.03.27.485993. <https://doi.org/10.1101/2022.03.27.485993>.
- [239] Liang, Y.-L.; Zhao, P.; Draper-Joyce, C.; Baltos, J.-A.; Glukhova, A.; Truong, T. T.; May, L. T.; Christopoulos, A.; Wootten, D.; Sexton, P. M.; et al. Dominant Negative G Proteins Enhance Formation and Purification of Agonist-GPCR-G Protein Complexes for Structure Determination. *Acs Pharmacol Transl Sci*, **2018**, *1* (1), 12–20. <https://doi.org/10.1021/acsptsci.8b00017>.

- [240] Lv, X.; Liu, J.; Shi, Q.; Tan, Q.; Wu, D.; Skinner, J. J.; Walker, A. L.; Zhao, L.; Gu, X.; Chen, N.; et al. In Vitro Expression and Analysis of the 826 Human G Protein-Coupled Receptors. *Protein Cell*, **2016**, 7 (5), 325–337. <https://doi.org/10.1007/s13238-016-0263-8>.
- [241] Bertheleme, N.; Chae, P. S.; Singh, S.; Mossakowska, D.; Hann, M. M.; Smith, K. J.; Hubbard, J. A.; Dowell, S. J.; Byrne, B. Unlocking the Secrets of the Gatekeeper: Methods for Stabilizing and Crystallizing GPCRs. *Biochimica Et Biophysica Acta Bba - Biomembr*, **2013**, 1828 (11), 2583–2591. <https://doi.org/10.1016/j.bbamem.2013.07.013>.
- [242] Schneberg, T.; Liu, J.; Wess, J. Plasma Membrane Localization and Functional Rescue of Truncated Forms of a G Protein-Coupled Receptor. *J. Biol. Chem.*, **1995**, 270 (30), 18000–18006. <https://doi.org/10.1074/jbc.270.30.18000>.
- [243] André, N.; Cherouati, N.; Prual, C.; Steffan, T.; Zeder-Lutz, G.; Magnin, T.; Pattus, F.; Michel, H.; Wagner, R.; Reinhart, C. Enhancing Functional Production of G Protein-coupled Receptors in *Pichia Pastoris* to Levels Required for Structural Studies via a Single Expression Screen. *Protein Sci*, **2006**, 15 (5), 1115–1126. <https://doi.org/10.1110/ps.062098206>.
- [244] Tate, C. G.; Schertler, G. F. Engineering G Protein-Coupled Receptors to Facilitate Their Structure Determination. *Curr. Opin. Struct. Biol.*, **2009**, 19 (4), 386–395. <https://doi.org/10.1016/j.sbi.2009.07.004>.
- [245] Heydenreich, F. M.; Vuckovic, Z.; Matkovic, M.; Veprintsev, D. B. Stabilization of G Protein-Coupled Receptors by Point Mutations. *Front. Pharmacol.*, **2015**, 6, 82. <https://doi.org/10.3389/fphar.2015.00082>.
- [246] Popov, P.; Peng, Y.; Shen, L.; Stevens, R. C.; Cherezov, V.; Liu, Z.-J.; Katritch, V. Computational Design of Thermostabilizing Point Mutations for G Protein-Coupled Receptors. *Elife*, **2018**, 7, e34729. <https://doi.org/10.7554/elife.34729>.
- [247] Grisshammer, R. Chapter 36 Purification of Recombinant G-Protein-Coupled Receptors. *Methods Enzym.*, **2009**, 463, 631–645. [https://doi.org/10.1016/s0076-6879\(09\)63036-6](https://doi.org/10.1016/s0076-6879(09)63036-6).
- [248] Yeliseev, A.; Zoubak, L.; Schmidt, T. G. M. Application of Strep-Tactin XT for Affinity Purification of Twin-Strep-Tagged CB2, a G Protein-Coupled Cannabinoid Receptor. *Protein Expr. Purif.*, **2017**, 131, 109–118. <https://doi.org/10.1016/j.pep.2016.11.006>.
- [249] Waugh, D. S. An Overview of Enzymatic Reagents for the Removal of Affinity Tags. *Protein Expr. Purif.*, **2011**, 80 (2), 283–293. <https://doi.org/10.1016/j.pep.2011.08.005>.
- [250] Dewey, J. S.; Struck, D. K.; Young, R. Thiol Protection in Membrane Protein Purifications: A Study with Phage Holins. *Anal. Biochem.*, **2009**, 390 (2), 221–223. <https://doi.org/10.1016/j.ab.2009.04.031>.

- [251] Stetsenko, A.; Guskov, A. An Overview of the Top Ten Detergents Used for Membrane Protein Crystallization. *Crystals*, **2017**, *7* (7), 197. <https://doi.org/10.3390/cryst7070197>.
- [252] Ratkeviciute, G.; Cooper, B. F.; Knowles, T. J. Methods for the Solubilisation of Membrane Proteins: The Micelle-Aneous World of Membrane Protein Solubilisation. *Biochem. Soc. Trans.*, **2021**, *49* (4), 1763–1777. <https://doi.org/10.1042/bst20210181>.
- [253] Contreras, F.-X.; Ernst, A. M.; Wieland, F.; Brügger, B. Specificity of Intramembrane Protein–Lipid Interactions. *Cold Spring Harb. Perspect. Biol.*, **2011**, *3* (6), a004705. <https://doi.org/10.1101/cshperspect.a004705>.
- [254] Mueller, K. Structural Dimorphism of Bile Salt/Lecithin Mixed Micelles. A Possible Regulatory Mechanism for Cholesterol Solubility in Bile? X-Ray Structural Analysis. *Biochemistry*, **1981**, *20* (2), 404–414. <https://doi.org/10.1021/bi00505a028>.
- [255] Diller, A.; Loudet, C.; Aussenac, F.; Raffard, G.; Fournier, S.; Laguerre, M.; Grélard, A.; Opella, S. J.; Marassi, F. M.; Dufourc, E. J. Bicelles: A Natural ‘Molecular Goniometer’ for Structural, Dynamical and Topological Studies of Molecules in Membranes. *Biochimie*, **2009**, *91* (6), 744–751. <https://doi.org/10.1016/j.biochi.2009.02.003>.
- [256] Tribet, C.; Audebert, R.; Popot, J.-L. Amphipols: Polymers That Keep Membrane Proteins Soluble in Aqueous Solutions. *Proc. Natl. Acad. Sci.*, **1996**, *93* (26), 15047–15050. <https://doi.org/10.1073/pnas.93.26.15047>.
- [257] Dörr, J. M.; Scheidelaar, S.; Koorengel, M. C.; Dominguez, J. J.; Schäfer, M.; Walree, C. A. van; Killian, J. A. The Styrene–Maleic Acid Copolymer: A Versatile Tool in Membrane Research. *Eur. Biophys. J.*, **2016**, *45* (1), 3–21. <https://doi.org/10.1007/s00249-015-1093-y>.
- [258] Frauenfeld, J.; Löving, R.; Armache, J.-P.; Sonnen, A. F.-P.; Guettou, F.; Moberg, P.; Zhu, L.; Jegerschöld, C.; Flayhan, A.; Briggs, J. A. G.; et al. A Saposin-Lipoprotein Nanoparticle System for Membrane Proteins. *Nat. Methods*, **2016**, *13* (4), 345–351. <https://doi.org/10.1038/nmeth.3801>.
- [259] Denisov, I. G.; Sligar, S. G. Nanodiscs for Structural and Functional Studies of Membrane Proteins. *Nat. Struct. Mol. Biol.*, **2016**, *23* (6), 481–486. <https://doi.org/10.1038/nsmb.3195>.
- [260] Akbarzadeh, A.; Rezaei-Sadabady, R.; Davaran, S.; Joo, S. W.; Zarghami, N.; Hanifehpour, Y.; Samiei, M.; Kouhi, M.; Nejati-Koshki, K. Liposome: Classification, Preparation, and Applications. *Nanoscale Res. Lett.*, **2013**, *8* (1), 102. <https://doi.org/10.1186/1556-276x-8-102>.
- [261] Carlson, M. L.; Young, J. W.; Zhao, Z.; Fabre, L.; Jun, D.; Li, J.; Li, J.; Dhupar, H. S.; Wason, I.; Mills, A. T.; et al. The Peptidisc, a Simple Method for Stabilizing Membrane Proteins in Detergent-Free Solution. *eLife*, **2018**, *7*, e34085. <https://doi.org/10.7554/elife.34085>.

- [262] Dubochet, J.; Adrian, M.; Chang, J.-J.; Homo, J.-C.; Lepault, J.; McDowell, A. W.; Schultz, P. Cryo-Electron Microscopy of Vitrified Specimens. *Q. Rev. Biophys.*, **1988**, *21* (2), 129–228. <https://doi.org/10.1017/s0033583500004297>.
- [263] Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C. A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature*, **1958**, *181* (4610), 662–666. <https://doi.org/10.1038/181662a0>.
- [264] García-Nafría, J.; Tate, C. G. Structure Determination of GPCRs: Cryo-EM Compared with X-Ray Crystallography. *Biochem Soc T*, **2021**, *49* (5), 2345–2355. <https://doi.org/10.1042/bst20210431>.
- [265] Hardy, D.; Bill, R. M.; Jawhari, A.; Rothnie, A. J. Overcoming Bottlenecks in the Membrane Protein Structural Biology Pipeline. *Biochem. Soc. Trans.*, **2016**, *44* (3), 838–844. <https://doi.org/10.1042/bst20160049>.
- [266] Midgett, C. R.; Madden, D. R. Breaking the Bottleneck: Eukaryotic Membrane Protein Expression for High-Resolution Structural Studies. *J. Struct. Biol.*, **2007**, *160* (3), 265–274. <https://doi.org/10.1016/j.jsb.2007.07.001>.
- [267] Landau, E. M.; Rosenbusch, J. P. Lipidic Cubic Phases: A Novel Concept for the Crystallization of Membrane Proteins. *Proc. Natl. Acad. Sci.*, **1996**, *93* (25), 14532–14535. <https://doi.org/10.1073/pnas.93.25.14532>.
- [268] Caffrey, M. A Comprehensive Review of the Lipid Cubic Phase or in Meso Method for Crystallizing Membrane and Soluble Proteins and Complexes. *Acta Crystallogr. Sect. F*, **2015**, *71* (1), 3–18. <https://doi.org/10.1107/s2053230x14026843>.
- [269] Magnani, F.; Serrano-Vega, M. J.; Shibata, Y.; Abdul-Hussein, S.; Lebon, G.; Miller-Gallacher, J.; Singhal, A.; Strege, A.; Thomas, J. A.; Tate, C. G. A Mutagenesis and Screening Strategy to Generate Optimally Thermostabilized Membrane Proteins for Structural Studies. *Nat. Protoc.*, **2016**, *11* (8), 1554–1571. <https://doi.org/10.1038/nprot.2016.088>.
- [270] Salom, D.; Padayatti, P. S.; Palczewski, K. Chapter 24 Crystallization of G Protein-Coupled Receptors. *Methods Cell Biol.*, **2013**, *117*, 451–468. <https://doi.org/10.1016/b978-0-12-408143-7.00024-4>.
- [271] Liang, Y.-L.; Khoshouei, M.; Radjainia, M.; Zhang, Y.; Glukhova, A.; Tarrasch, J.; Thal, D. M.; Furness, S. G. B.; Christopoulos, G.; Coudrat, T.; et al. Phase-Plate Cryo-EM Structure of a Class B GPCR–G-Protein Complex. *Nature*, **2017**, *546* (7656), 118–123. <https://doi.org/10.1038/nature22327>.
- [272] Liu, Y.; Huynh, D. T.; Yeates, T. O. A 3.8 Å Resolution Cryo-EM Structure of a Small Protein Bound to an Imaging Scaffold. *Nat. Commun.*, **2019**, *10* (1), 1864. <https://doi.org/10.1038/s41467-019-09836-0>.

- [273] Du, Y.; Duc, N. M.; Rasmussen, S. G. F.; Hilger, D.; Kubiak, X.; Wang, L.; Bohon, J.; Kim, H. R.; Wegrecki, M.; Asuru, A.; et al. Assembly of a GPCR-G Protein Complex. *Cell*, **2019**, *177* (5), 1232-1242.e11. <https://doi.org/10.1016/j.cell.2019.04.022>.
- [274] Huang, W.; Masureel, M.; Qu, Q.; Janetzko, J.; Inoue, A.; Kato, H. E.; Robertson, M. J.; Nguyen, K. C.; Glenn, J. S.; Skiniotis, G.; et al. Structure of the Neurotensin Receptor 1 in Complex with  $\beta$ -Arrestin 1. *Nature*, **2020**, *579* (7798), 303–308. <https://doi.org/10.1038/s41586-020-1953-1>.
- [275] Mukherjee, S.; Erramilli, S. K.; Ammirati, M.; Alvarez, F. J. D.; Fennell, K. F.; Purdy, M. D.; Skrobek, B. M.; Radziwon, K.; Coukos, J.; Kang, Y.; et al. Synthetic Antibodies against BRIL as Universal Fiducial Marks for Single-particle CryoEM Structure Determination of Membrane Proteins. *Nat. Commun.*, **2020**, *11* (1), 1598. <https://doi.org/10.1038/s41467-020-15363-0>.
- [276] Miyagi, H.; Asada, H.; Suzuki, M.; Takahashi, Y.; Yasunaga, M.; Suno, C.; Iwata, S.; Saito, J. The Discovery of a New Antibody for BRIL-Fused GPCR Structure Determination. *Sci Rep-uk*, **2020**, *10* (1), 11669. <https://doi.org/10.1038/s41598-020-68355-x>.
- [277] Hough, M. A.; Prischi, F.; Worrall, J. A. R. Perspective: Structure Determination of Protein-Ligand Complexes at Room Temperature Using X-Ray Diffraction Approaches. *Front. Mol. Biosci.*, **2023**, *10*, 1113762. <https://doi.org/10.3389/fmolb.2023.1113762>.
- [278] Wang, H.-W. A Commentary of “Cryo-EM Achieves Atomic Resolution” in 10 Remarkable Discoveries from 2020 in Nature. *Fundam. Res.*, **2022**, *2* (2), 349–350. <https://doi.org/10.1016/j.fmre.2022.01.014>.
- [279] Ashmore, J.; Carragher, B.; Rosenthal, P. B.; Weis, W. A Resolution Record for CryoEM. *Fac. Rev.*, **2021**, *10*, 64. <https://doi.org/10.12703/r-01-000002>.
- [280] He, J.; Li, T.; Huang, S.-Y. Improvement of Cryo-EM Maps by Simultaneous Local and Non-Local Deep Learning. *Nat. Commun.*, **2023**, *14* (1), 3217. <https://doi.org/10.1038/s41467-023-39031-1>.
- [281] Dahl, S. G.; Edvardsen, O.; Sylte, I. Molecular Dynamics of Dopamine at the D2 Receptor. *Proc. Natl. Acad. Sci.*, **1991**, *88* (18), 8111–8115. <https://doi.org/10.1073/pnas.88.18.8111>.
- [282] Torrens-Fontanals, M.; Stepniewski, T. M.; Aranda-García, D.; Morales-Pastor, A.; Medel-Lacruz, B.; Selent, J. How Do Molecular Dynamics Data Complement Static Structural Data of GPCRs. *Int. J. Mol. Sci.*, **2020**, *21* (16), 5933. <https://doi.org/10.3390/ijms21165933>.
- [283] Baker, J. A.; Hirst, J. D. Molecular Dynamics Simulations Using Graphics Processing Units. *Mol. Inform.*, **2011**, *30* (6-7), 498–504. <https://doi.org/10.1002/minf.201100042>.

- [284] Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-based Graphical User Interface for CHARMM. *J Comput Chem*, **2008**, *29* (11), 1859–1865. <https://doi.org/10.1002/jcc.20945>.
- [285] Mayol, E.; García-Recio, A.; Tiemann, J. K. S.; Hildebrand, P. W.; Guixà-González, R.; Olivella, M.; Codomí, A. HomolWat: A Web Server Tool to Incorporate ‘Homologous’ Water Molecules into GPCR Structures. *Nucleic Acids Res.*, **2020**, *48* (W1), gkaa440-. <https://doi.org/10.1093/nar/gkaa440>.
- [286] Sandal, M.; Duy, T. P.; Cona, M.; Zung, H.; Carloni, P.; Musiani, F.; Giorgetti, A. GOMoDo: A GPCRs Online Modeling and Docking Webserver. *Plos One*, **2013**, *8* (9), e74092. <https://doi.org/10.1371/journal.pone.0074092>.
- [287] Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.*, **2011**, *32* (10), 2319–2327. <https://doi.org/10.1002/jcc.21787>.
- [288] Rodríguez-Espigares, I.; Torrens-Fontanals, M.; Tiemann, J. K. S.; Aranda-García, D.; Ramírez-Anguita, J. M.; Stepniewski, T. M.; Worp, N.; Varela-Rial, A.; Morales-Pastor, A.; Medel-Lacruz, B.; et al. GPCRmd Uncovers the Dynamics of the 3D-GPCRome. *Nat. Methods*, **2020**, *17* (8), 777–787. <https://doi.org/10.1038/s41592-020-0884-y>.
- [289] Latorraca, N. R.; Venkatakrisnan, A. J.; Dror, R. O. GPCR Dynamics: Structures in Motion. *Chem Rev*, **2017**, *117* (1), 139–155. <https://doi.org/10.1021/acs.chemrev.6b00177>.
- [290] Oddi, S.; Stepniewski, T. M.; Totaro, A.; Selent, J.; Scipioni, L.; Dufrusine, B.; Fezza, F.; Dainese, E.; Maccarrone, M. Palmitoylation of Cysteine 415 of CB1 Receptor Affects Ligand-Stimulated Internalization and Selective Interaction with Membrane Cholesterol and Caveolin 1. *Biochim. Biophys. Acta (BBA) - Mol. Cell Biol. Lipids*, **2017**, *1862* (5), 523–532. <https://doi.org/10.1016/j.bbalip.2017.02.004>.
- [291] Guixà-González, R.; Albasanz, J. L.; Rodríguez-Espigares, I.; Pastor, M.; Sanz, F.; Martí-Solano, M.; Manna, M.; Martínez-Seara, H.; Hildebrand, P. W.; Martín, M.; et al. Membrane Cholesterol Access into a G-Protein-Coupled Receptor. *Nat. Commun.*, **2017**, *8* (1), 14505. <https://doi.org/10.1038/ncomms14505>.
- [292] Ramírez-Anguita, J. M.; Rodríguez-Espigares, I.; Guixà-González, R.; Bruno, A.; Torrens-Fontanals, M.; Varela-Rial, A.; Selent, J. Membrane Cholesterol Effect on the 5-HT<sub>2A</sub> Receptor: Insights into the Lipid-induced Modulation of an Antipsychotic Drug Target. *Biotechnol. Appl. Biochem.*, **2018**, *65* (1), 29–37. <https://doi.org/10.1002/bab.1608>.
- [293] Guixà-González, R.; Javanainen, M.; Gómez-Soler, M.; Cordobilla, B.; Domingo, J. C.; Sanz, F.; Pastor, M.; Ciruela, F.; Martínez-Seara, H.; Selent, J. Membrane Omega-3 Fatty Acids Modulate the Oligomerisation Kinetics of Adenosine A<sub>2A</sub> and Dopamine D<sub>2</sub> Receptors. *Sci. Rep.*, **2016**, *6* (1), 19839. <https://doi.org/10.1038/srep19839>.
- [294] Angel, T. E.; Chance, M. R.; Palczewski, K. Conserved Waters Mediate Structural and Functional Activation of Family A (Rhodopsin-like) G Protein-Coupled Receptors.

*Proc National Acad Sci*, **2009**, *106* (21), 8555–8560.  
<https://doi.org/10.1073/pnas.0903545106>.

[295] Yuan, S.; Filipek, S.; Palczewski, K.; Vogel, H. Activation of G-Protein-Coupled Receptors Correlates with the Formation of a Continuous Internal Water Pathway. *Nat Commun*, **2014**, *5* (1), 4733. <https://doi.org/10.1038/ncomms5733>.

[296] Bruno, A.; Costantino, G. Molecular Dynamics Simulations of G Protein-Coupled Receptors. *Mol. Inform.*, **2012**, *31* (3-4), 222–230.  
<https://doi.org/10.1002/minf.201100138>.

[297] Wang, T.; Zhou, G.; He, M.; Xu, Y.; Rusyniak, W. G.; Xu, Y.; Ji, Y.; Simon, R. P.; Xiong, Z.-G.; Zha, X. GPR68 Is a Neuroprotective Proton Receptor in Brain Ischemia. *Stroke*, **2020**, *51* (12), 3690–3700. <https://doi.org/10.1161/strokeaha.120.031479>.

[298] Haak, A. J.; Ducharme, M. T.; Espinosa, A. M. D.; Tschumperlin, D. J. Targeting GPCR Signaling for Idiopathic Pulmonary Fibrosis Therapies. *Trends Pharmacol. Sci.*, **2020**, *41* (3), 172–182. <https://doi.org/10.1016/j.tips.2019.12.008>.

[299] Zeng, Z.; Mukherjee, A.; Varghese, A. P.; Yang, X.-L.; Chen, S.; Zhang, H. Roles of G Protein-Coupled Receptors in Inflammatory Bowel Disease. *World J. Gastroenterol.*, **2020**, *26* (12), 1242–1261. <https://doi.org/10.3748/wjg.v26.i12.1242>.

[300] Siemers, M.; Lazaratos, M.; Karathanou, K.; Guerra, F.; Brown, L. S.; Bondar, A.-N. Bridge: A Graph-Based Algorithm to Analyze Dynamic H-Bond Networks in Membrane Proteins. *J Chem Theory Comput*, **2019**, *15* (12), 6781–6798.  
<https://doi.org/10.1021/acs.jctc.9b00697>.

[301] Siemers, M.; Bondar, A.-N. Interactive Interface for Graph-Based Analyses of Dynamic H-Bond Networks: Application to Spike Protein S. *J Chem Inf Model*, **2021**, *61* (6), 2998–3014. <https://doi.org/10.1021/acs.jcim.1c00306>.

[302] Degorce, F.; Card, A.; Soh, S.; Trinquet, E.; Knapik, G. P.; Xie, B. HTRF: A Technology Tailored for Drug Discovery –A Review of Theoretical Aspects and Recent Applications. *Curr. Chem. Genom.*, **2009**, *3* (1), 22–32.  
<https://doi.org/10.2174/1875397300903010022>.

[303] Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr. Sect. D: Biol. Crystallogr.*, **2010**, *66* (4), 486–501.  
<https://doi.org/10.1107/s0907444910007493>.

[304] Wilkins, M. R.; Gasteiger, E.; Bairoch, A.; Sanchez, J.-C.; Williams, K. L.; Appel, R. D.; Hochstrasser, D. F. 2-D Proteome Analysis Protocols. *Methods Mol. Biol. (Clifton, NJ)*, **1998**, *112*, 531–552. <https://doi.org/10.1385/1-59259-584-7:531>.

[305] Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; et al. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *J Chem Phys*, **2020**, *153* (4), 044130.  
<https://doi.org/10.1063/5.0014475>.

- [306] Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkóczi, G.; Chen, V. B.; Croll, T. I.; Hintze, B.; Hung, L.-W.; Jain, S.; McCoy, A. J.; et al. Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments in Phenix. *Acta Crystallogr. Sect. D*, **2019**, *75* (Pt 10), 861–877. <https://doi.org/10.1107/s2059798319011471>.
- [307] Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J Mol Graphics*, **1996**, *14* (1), 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
- [308] Miyagi, H.; Suzuki, M.; Yasunaga, M.; Asada, H.; Iwata, S.; Saito, J. Structural Insight into an Anti-BRIL Fab as a G-Protein-Coupled Receptor Crystallization Chaperone. *Acta Crystallogr. Sect. D*, **2023**, *79* (Pt 5), 435–441. <https://doi.org/10.1107/s205979832300311x>.
- [309] Bloch, J. S.; Mukherjee, S.; Kowal, J.; Filippova, E. V.; Niederer, M.; Pardon, E.; Steyaert, J.; Kossiakoff, A. A.; Locher, K. P. Development of a Universal Nanobody-Binding Fab Module for Fiducial-Assisted Cryo-EM Studies of Membrane Proteins. *Proc. Natl. Acad. Sci.*, **2021**, *118* (47), e2115435118. <https://doi.org/10.1073/pnas.2115435118>.
- [310] Pace, C. N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. How to Measure and Predict the Molar Absorption Coefficient of a Protein. *Protein Sci.*, **1995**, *4* (11), 2411–2423. <https://doi.org/10.1002/pro.5560041120>.
- [311] Cavaluzzi, M. J.; Borer, P. N. Revised UV Extinction Coefficients for Nucleoside-5'-monophosphates and Unpaired DNA and RNA. *Nucleic Acids Res.*, **2004**, *32* (1), e13–e13. <https://doi.org/10.1093/nar/gnh015>.
- [312] Laemmli, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, **1970**, *227* (5259), 680–685. <https://doi.org/10.1038/227680a0>.
- [313] Smith, B. J. Basic Protein and Peptide Protocols. *Methods Mol. Biol. (Clifton, NJ)*, **1994**, *32*, 23–34. <https://doi.org/10.1385/0-89603-268-x:23>.
- [314] Sagné, C.; Isambert, M.-F.; Henry, J.-P.; Gasnier, B. SDS-Resistant Aggregation of Membrane Proteins: Application to the Purification of the Vesicular Monoamine Transporter. *Biochem. J.*, **1996**, *316* (3), 825–831. <https://doi.org/10.1042/bj3160825>.
- [315] Rath, A.; Glibowicka, M.; Nadeau, V. G.; Chen, G.; Deber, C. M. Detergent Binding Explains Anomalous SDS-PAGE Migration of Membrane Proteins. *Proc. Natl. Acad. Sci.*, **2009**, *106* (6), 1760–1765. <https://doi.org/10.1073/pnas.0813167106>.
- [316] Kim, S. H.; Yoo, H. J.; Park, E. J.; Na, D. H. Nano Differential Scanning Fluorimetry-Based Thermal Stability Screening and Optimal Buffer Selection for Immunoglobulin G. *Pharmaceuticals*, **2021**, *15* (1), 29. <https://doi.org/10.3390/ph15010029>.
- [317] Gooran, N.; Kopra, K. Fluorescence-Based Protein Stability Monitoring—A Review. *Int. J. Mol. Sci.*, **2024**, *25* (3), 1764. <https://doi.org/10.3390/ijms25031764>.

- [318] Real-Hohn, A.; Groznica, M.; Löffler, N.; Blaas, D.; Kowalski, H. NanoDSF: In Vitro Label-Free Method to Monitor Picornavirus Uncoating and Test Compounds Affecting Particle Stability. *Front. Microbiol.*, **2020**, *11*, 1442. <https://doi.org/10.3389/fmicb.2020.01442>.
- [319] Caffrey, M.; Cherezov, V. Crystallizing Membrane Proteins Using Lipidic Mesophases. *Nat. Protoc.*, **2009**, *4* (5), 706–731. <https://doi.org/10.1038/nprot.2009.31>.
- [320] Ewald, P. P. Max von Laue, 1879-1960. *Biogr. Mem. Fellows R. Soc.*, **1960**, *6* (6), 134–156. <https://doi.org/10.1098/rsbm.1960.0028>.
- [321] Friedrich, W.; Knipping, P.; Laue, M. Interferenzerscheinungen Bei Röntgenstrahlen. *Ann. Phys.*, **1913**, *346* (10), 971–988. <https://doi.org/10.1002/andp.19133461004>.
- [322] Benvenuti, M.; Mangani, S. Crystallization of Soluble Proteins in Vapor Diffusion for X-Ray Crystallography. *Nat. Protoc.*, **2007**, *2* (7), 1633–1651. <https://doi.org/10.1038/nprot.2007.198>.
- [323] Ng, J. D.; Gavira, J. A.; García-Ruíz, J. M. Protein Crystallization by Capillary Counterdiffusion for Applied Crystallographic Structure Determination. *J. Struct. Biol.*, **2003**, *142* (1), 218–231. [https://doi.org/10.1016/s1047-8477\(03\)00052-2](https://doi.org/10.1016/s1047-8477(03)00052-2).
- [324] Thomas, D. H.; Rob, A.; Rice, D. W. A Novel Dialysis Procedure for the Crystallization of Proteins. *Protein Eng., Des. Sel.*, **1989**, *2* (6), 489–491. <https://doi.org/10.1093/protein/2.6.489>.
- [325] Chayen, N. E.; Stewart, P. D. S.; Blow, D. M. Microbatch Crystallization under Oil — a New Technique Allowing Many Small-Volume Crystallization Trials. *J. Cryst. Growth*, **1992**, *122* (1–4), 176–180. [https://doi.org/10.1016/0022-0248\(92\)90241-a](https://doi.org/10.1016/0022-0248(92)90241-a).
- [326] Rayment, I. Small-Scale Batch Crystallization of Proteins Revisited An Underutilized Way to Grow Large Protein Crystals. *Structure*, **2002**, *10* (2), 147–151. [https://doi.org/10.1016/s0969-2126\(02\)00711-6](https://doi.org/10.1016/s0969-2126(02)00711-6).
- [327] Smyth, M. S.; Martin, J. H. J. X Ray Crystallography. *Mol. Pathol.*, **2000**, *53* (1), 8. <https://doi.org/10.1136/mp.53.1.8>.
- [328] Pomés, A.; Chruszcz, M.; Gustchina, A.; Minor, W.; Mueller, G. A.; Pedersen, L. C.; Wlodawer, A.; Chapman, M. D. 100 Years Later: Celebrating the Contributions of x-Ray Crystallography to Allergy and Clinical Immunology. *J. Allergy Clin. Immunol.*, **2015**, *136* (1), 29-37.e10. <https://doi.org/10.1016/j.jaci.2015.05.016>.
- [329] Fecher, G. H.; Kübler, J.; Felser, C. Chirality in the Solid State: Chiral Crystal Structures in Chiral and Achiral Space Groups. *Materials*, **2022**, *15* (17), 5812. <https://doi.org/10.3390/ma15175812>.
- [330] Rossmann, M. G. The Molecular Replacement Method. *Acta Crystallogr. Sect. A*, **1990**, *46* (2), 73–82. <https://doi.org/10.1107/s0108767389009815>.

- [331] Green, D. W.; Ingram, V. M.; Perutz, M. F. The Structure of Haemoglobin - IV. Sign Determination by the Isomorphous Replacement Method. *Proc. R. Soc. Lond. Ser. A Math. Phys. Sci.*, **1954**, 225 (1162), 287–307. <https://doi.org/10.1098/rspa.1954.0203>.
- [332] Guss, J. M.; Merritt, E. A.; Phizackerley, R. P.; Hedman, B.; Murata, M.; Hodgson, K. O.; Freeman, H. C. Phase Determination by Multiple-Wavelength X-Ray Diffraction: Crystal Structure of a Basic “Blue” Copper Protein from Cucumbers. *Science*, **1988**, 241 (4867), 806–811. <https://doi.org/10.1126/science.3406739>.
- [333] Rose, J. P.; Wang, B.-C. SAD Phasing: History, Current Impact and Future Opportunities. *Arch. Biochem. Biophys.*, **2016**, 602, 80–94. <https://doi.org/10.1016/j.abb.2016.03.018>.
- [334] Eyck, L. F. T. Crystallographic Fast Fourier Transforms. *Acta Crystallogr. Sect. A*, **1973**, 29 (2), 183–191. <https://doi.org/10.1107/s0567739473000458>.
- [335] Yin, X.; Xu, H.; Hanson, M.; Liu, W. GPCR Crystallization Using Lipidic Cubic Phase Technique. *Curr. Pharm. Biotechnol.*, **2014**, 15 (10), 971–979. <https://doi.org/10.2174/1389201015666140922110325>.
- [336] Milne, J. L. S.; Borgnia, M. J.; Bartesaghi, A.; Tran, E. E. H.; Earl, L. A.; Schauder, D. M.; Lengyel, J.; Pierson, J.; Patwardhan, A.; Subramaniam, S. Cryo-electron Microscopy – a Primer for the Non-microscopist. *FEBS J.*, **2013**, 280 (1), 28–45. <https://doi.org/10.1111/febs.12078>.
- [337] Glaeser, R. M.; Taylor, K. A. Radiation Damage Relative to Transmission Electron Microscopy of Biological Specimens at Low Temperature: A Review. *J. Microsc.*, **1978**, 112 (1), 127–138. <https://doi.org/10.1111/j.1365-2818.1978.tb01160.x>.
- [338] Punjani, A.; Rubinstein, J. L.; Fleet, D. J.; Brubaker, M. A. CryoSPARC: Algorithms for Rapid Unsupervised Cryo-EM Structure Determination. *Nat. Methods*, **2017**, 14 (3), 290–296. <https://doi.org/10.1038/nmeth.4169>.
- [339] Kapur, B.; Baldessari, F.; Lazaratos, M.; Nar, H.; Schnapp, G.; Giorgetti, A.; Bondar, A.-N. Protons Taken Hostage: Dynamic H-Bond Networks of the PH-Sensing GPR68. *Comput. Struct. Biotechnol. J.*, **2023**, 21, 4370–4384. <https://doi.org/10.1016/j.csbj.2023.08.034>.
- [340] Fiser, A.; Do, R. K. G.; Šali, A. Modeling of Loops in Protein Structures. *Protein Sci*, **2000**, 9 (9), 1753–1773. <https://doi.org/10.1110/ps.9.9.1753>.
- [341] Eilers, M.; Hornak, V.; Smith, S. O.; Konopka, J. B. Comparison of Class A and D G Protein-Coupled Receptors: Common Features in Structure and Activation †. *Biochemistry*, **2005**, 44 (25), 8959–8975. <https://doi.org/10.1021/bi047316u>.
- [342] Lomize, M. A.; Pogozheva, I. D.; Joo, H.; Mosberg, H. I.; Lomize, A. L. OPM Database and PPM Web Server: Resources for Positioning of Proteins in Membranes. *Nucleic Acids Res*, **2012**, 40 (Database issue), D370–D376. <https://doi.org/10.1093/nar/gkr703>.

- [343] Wu, E. L.; Cheng, X.; Jo, S.; Rui, H.; Song, K. C.; Dávila-Contreras, E. M.; Qi, Y.; Lee, J.; Monje-Galvan, V.; Venable, R. M.; et al. CHARMM-GUI Membrane Builder toward Realistic Biological Membrane Simulations. *J Comput Chem*, **2014**, *35* (27), 1997–2004. <https://doi.org/10.1002/jcc.23702>.
- [344] Subczynski, W. K.; Pasenkiewicz-Gierula, M.; Widomska, J.; Mainali, L.; Raguz, M. High Cholesterol/Low Cholesterol: Effects in Biological Membranes: A Review. *Cell Biochem. Biophys.*, **2017**, *75* (3–4), 369–385. <https://doi.org/10.1007/s12013-017-0792-7>.
- [345] Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; Groot, B. L. de; Grubmüller, H.; MacKerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat Methods*, **2017**, *14* (1), 71–73. <https://doi.org/10.1038/nmeth.4067>.
- [346] Price, D. J.; Brooks, C. L. A Modified TIP3P Water Potential for Simulation with Ewald Summation. *J. Chem. Phys.*, **2004**, *121* (20), 10096–10103. <https://doi.org/10.1063/1.1808117>.
- [347] Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J Comput Chem*, **2005**, *26* (16), 1781–1802. <https://doi.org/10.1002/jcc.20289>.
- [348] Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: Greater Scalability for Parallel Molecular Dynamics. *J. Comput. Phys.*, **1999**, *151* (1), 283–312. <https://doi.org/10.1006/jcph.1999.6201>.
- [349] Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J Chem Phys*, **1995**, *103* (11), 4613–4621. <https://doi.org/10.1063/1.470648>.
- [350] Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A*, **1985**, *31* (3), 1695–1697. <https://doi.org/10.1103/physreva.31.1695>.
- [351] Nosé, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys.*, **1984**, *81* (1), 511–519. <https://doi.org/10.1063/1.447334>.
- [352] Dalton, J. A.; Lans, I.; Giraldo, J. Quantifying Conformational Changes in GPCRs: Glimpse of a Common Functional Mechanism. *Bmc Bioinformatics*, **2015**, *16* (1), 124. <https://doi.org/10.1186/s12859-015-0567-3>.
- [353] Nørskov-Lauritsen, L.; Thomsen, A.; Bräuner-Osborne, H. G Protein-Coupled Receptor Signaling Analysis Using Homogenous Time-Resolved Förster Resonance Energy Transfer (HTRF®) Technology. *Int. J. Mol. Sci.*, **2014**, *15* (2), 2554–2572. <https://doi.org/10.3390/ijms15022554>.
- [354] Kanemaru, K.; Nakamura, Y. Activation Mechanisms and Diverse Functions of Mammalian Phospholipase C. *Biomolecules*, **2023**, *13* (6), 915. <https://doi.org/10.3390/biom13060915>.

- [355] Vadnal, R.; Parthasarathy, R. Myo-Inositol Monophosphatase: Diverse Effects of Lithium, Carbamazepine, and Valproate. *Neuropsychopharmacology*, **1995**, *12* (4), 277–285. [https://doi.org/10.1016/0893-133x\(94\)00088-h](https://doi.org/10.1016/0893-133x(94)00088-h).
- [356] Gavriilidou, A. F. M.; Hunziker, H.; Mayer, D.; Vuckovic, Z.; Veprintsev, D. B.; Zenobi, R. Insights into the Basal Activity and Activation Mechanism of the B1 Adrenergic Receptor Using Native Mass Spectrometry. *J. Am. Soc. Mass Spectrom.*, **2019**, *30* (3), 529–537. <https://doi.org/10.1007/s13361-018-2110-z>.
- [357] Weatherall, M. The Meaning and Importance of Drug Potency in Medicine. *Clin. Pharmacol. Ther.*, **1966**, *7* (5), 577–582. <https://doi.org/10.1002/cpt196675577>.
- [358] Vanaken, T.; Foxall-Vanaken, S.; Castleman, S.; Ferguson-Miller, S. Alkyl Glycoside Detergents: Synthesis and Applications to the Study of Membrane Proteins. *Methods Enzym.*, **1986**, *125*, 27–35. [https://doi.org/10.1016/s0076-6879\(86\)25005-3](https://doi.org/10.1016/s0076-6879(86)25005-3).
- [359] Chaptal, V.; Delolme, F.; Kilburg, A.; Magnard, S.; Montigny, C.; Picard, M.; Prier, C.; Monticelli, L.; Bornert, O.; Agez, M.; et al. Quantification of Detergents Complexed with Membrane Proteins. *Sci. Rep.*, **2017**, *7* (1), 41751. <https://doi.org/10.1038/srep41751>.
- [360] Ehsan, M.; Das, M.; Stern, V.; Du, Y.; Mortensen, J. S.; Hariharan, P.; Byrne, B.; Loland, C. J.; Kobilka, B. K.; Guan, L.; et al. Steroid-Based Amphiphiles for Membrane Protein Study: The Importance of Alkyl Spacers for Protein Stability. *ChemBioChem*, **2018**, *19* (13), 1433–1443. <https://doi.org/10.1002/cbic.201800106>.
- [361] Lee, H. J.; Lee, H. S.; Youn, T.; Byrne, B.; Chae, P. S. Impact of Novel Detergents on Membrane Protein Studies. *Chem*, **2022**, *8* (4), 980–1013. <https://doi.org/10.1016/j.chempr.2022.02.007>.
- [362] Pándy-Szekeres, G.; Caroli, J.; Mamyrbekov, A.; Kermani, A. A.; Keserű, G. M.; Kooistra, A. J.; Gloriam, D. E. GPCRdb in 2023: State-Specific Structure Models Using AlphaFold2 and New Ligand Resources. *Nucleic Acids Res.*, **2022**, *51* (D1), D395–D402. <https://doi.org/10.1093/nar/gkac1013>.
- [363] Grahl, A.; Abiko, L. A.; Isogai, S.; Sharpe, T.; Grzesiek, S. A High-Resolution Description of B1-Adrenergic Receptor Functional Dynamics and Allosteric Coupling from Backbone NMR. *Nat. Commun.*, **2020**, *11* (1), 2216. <https://doi.org/10.1038/s41467-020-15864-y>.
- [364] Yang, H.-S.; Sun, N.; Zhao, X.; Kim, H. R.; Park, H.-J.; Kim, K.-M.; Chung, K. Y. Role of Helix 8 in Dopamine Receptor Signaling. *Biomol. Ther.*, **2019**, *27* (6), 514–521. <https://doi.org/10.4062/biomolther.2019.026>.
- [365] Wang, J.; Wu, M.; Wu, L.; Xu, Y.; Li, F.; Wu, Y.; Popov, P.; Wang, L.; Bai, F.; Zhao, S.; et al. The Structural Study of Mutation-Induced Inactivation of Human Muscarinic Receptor M4. *IUCrJ*, **2020**, *7* (Pt 2), 294–305. <https://doi.org/10.1107/s2052252520000597>.

- [366] Zhang, X. C.; Zhou, Y.; Cao, C. Proton Transfer during Class-A GPCR Activation: Do the CWxP Motif and the Membrane Potential Act in Concert? *Biophysics Reports*, **2018**, *4* (3), 115–122. <https://doi.org/10.1007/s41048-018-0056-0>.
- [367] Claff, T.; Klapschinski, T. A.; Subhramanyam, U. K. T.; Vaaßen, V. J.; Schlegel, J. G.; Vielmuth, C.; Voß, J. H.; Labahn, J.; Müller, C. E. Single Stabilizing Point Mutation Enables High-Resolution Co-Crystal Structures of the Adenosine A2A Receptor with Preladenant Conjugates. *Angew. Chem. Int. Ed.*, **2022**, *61* (22), e202115545. <https://doi.org/10.1002/anie.202115545>.
- [368] Tsutsumi, N.; Mukherjee, S.; Waghay, D.; Janda, C. Y.; Jude, K. M.; Miao, Y.; Burg, J. S.; Aduri, N. G.; Kossiakoff, A. A.; Gati, C.; et al. Structure of Human Frizzled5 by Fiducial-Assisted Cryo-EM Supports a Heterodimeric Mechanism of Canonical Wnt Signaling. *eLife*, **2020**, *9*, e58464. <https://doi.org/10.7554/elife.58464>.
- [369] Crooks, G. E.; Hon, G.; Chandonia, J.-M.; Brenner, S. E. WebLogo: A Sequence Logo Generator. *Genome Res*, **2004**, *14* (6), 1188–1190. <https://doi.org/10.1101/gr.849004>.
- [370] Matsingos, C.; Howell, L. A.; McCormick, P. J.; Fornili, A. Elucidating the Activation Mechanism of the Proton-Sensing GPR68 Receptor. *bioRxiv*, **2023**, 2023.12.08.570878. <https://doi.org/10.1101/2023.12.08.570878>.
- [371] Hoppe, N. Structural Insights into Orphan G Protein-Coupled Receptors: A Focus on GPR161 and Proton-Sensing GPCRs, University of California, San Francisco, 2023.
- [372] Vaisse, C.; Clement, K.; Durand, E.; Hercberg, S.; Guy-Grand, B.; Froguel, P. Melanocortin-4 Receptor Mutations Are a Frequent and Heterogeneous Cause of Morbid Obesity. *J. Clin. Investig.*, **2000**, *106* (2), 253–262. <https://doi.org/10.1172/jci9238>.
- [373] Deupi, X.; Edwards, P.; Singhal, A.; Nickle, B.; Oprian, D.; Schertler, G.; Standfuss, J. Stabilized G Protein Binding Site in the Structure of Constitutively Active Metarhodopsin-II. *Proc National Acad Sci*, **2012**, *109* (1), 119–124. <https://doi.org/10.1073/pnas.1114089108>.
- [374] Shalaeva, D. N.; Cherepanov, D. A.; Galperin, M. Y.; Vriend, G.; Mulkidjanian, A. Y. G Protein-Coupled Receptors of Class A Harness the Energy of Membrane Potential to Increase Their Sensitivity and Selectivity. *Biochim. Biophys. Acta (BBA) - Biomembr.*, **2019**, *1861* (12), 183051. <https://doi.org/10.1016/j.bbamem.2019.183051>.
- [375] Howard, M. K.; Hoppe, N.; Huang, X.-P.; Macdonald, C. B.; Mehrotra, E.; Grimes, P. R.; Zahm, A. M.; Trinidad, D. D.; English, J. G.; Coyote-Maestas, W.; et al. Molecular Basis of Proton-Sensing by G Protein-Coupled Receptors. **2024**. <https://doi.org/10.1101/2024.04.17.590000>.
- [376] Glitsch, M. D. Helix 8 – Putting a Spring in Mechano-Sensing. *Cell Calcium*, **2020**, *87*, 102192. <https://doi.org/10.1016/j.ceca.2020.102192>.

- [377] Goddard, A. D.; Watts, A. Regulation of G Protein-Coupled Receptors by Palmitoylation and Cholesterol. *Bmc Biol*, **2012**, *10* (1), 27. <https://doi.org/10.1186/1741-7007-10-27>.
- [378] Lam, P. V. N.; Goldman, R.; Karagiannis, K.; Narsule, T.; Simonyan, V.; Soika, V.; Mazumder, R. Structure-Based Comparative Analysis and Prediction of N-Linked Glycosylation Sites in Evolutionarily Distant Eukaryotes. *Genom., Proteom. Bioinform.*, **2013**, *11* (2), 96–104. <https://doi.org/10.1016/j.gpb.2012.11.003>.
- [379] Mochimaru, Y.; Azuma, M.; Oshima, N.; Ichijo, Y.; Satou, K.; Matsuda, K.; Asaoka, Y.; Nishina, H.; Nakakura, T.; Mogi, C.; et al. Extracellular Acidification Activates Ovarian Cancer G-Protein-Coupled Receptor 1 and GPR4 Homologs of Zebra Fish. *Biochem Bioph Res Co*, **2015**, *457* (4), 493–499. <https://doi.org/10.1016/j.bbrc.2014.12.105>.
- [380] Williams, C. H.; Neitzel, L. R.; Karki, P.; Keyser, B. D.; Thayer, T. E.; Wells, Q. S.; Perry, J. A.; Birukova, A. A.; Birukov, K. G.; Hong, C. C. Coupling Metastasis to PH-Sensing GPR68 Using a Novel Small Molecule Inhibitor. *Biorxiv*, **2021**, 612549. <https://doi.org/10.1101/612549>.
- [381] Karki, P.; Ke, Y.; Williams, C. H.; Hong, C. C.; Birukov, K. G.; Birukova, A. A. GPR68 Inhibition with a Novel Group of Opioid Inhibitors Upregulate Endothelial Barrier Function and Protect Against Bacterial Pathogens or Acidosis-induced Inflammation in Lung Endothelium. *Faseb J*, **2022**, *36* (S1). <https://doi.org/10.1096/fasebj.2022.36.s1.r4848>.
- [382] Bamford, S.; Dawson, E.; Forbes, S.; Clements, J.; Pettett, R.; Dogan, A.; Flanagan, A.; Teague, J.; Futreal, P. A.; Stratton, M. R.; et al. The COSMIC (Catalogue of Somatic Mutations in Cancer) Database and Website. *Br. J. Cancer*, **2004**, *91* (2), 355–358. <https://doi.org/10.1038/sj.bjc.6601894>.
- [383] Tate, J. G.; Bamford, S.; Jubb, H. C.; Sondka, Z.; Beare, D. M.; Bindal, N.; Boutselakis, H.; Cole, C. G.; Creatore, C.; Dawson, E.; et al. COSMIC: The Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.*, **2018**, *47* (Database issue), gky1015-. <https://doi.org/10.1093/nar/gky1015>.
- [384] Li, B.; Wang, Y.; Castro, A.; Ng, C.; Wang, Z.; Chaudhry, H.; Agbaje, Z.; Ulloa, G. A.; Yu, Y. The Roles of Two Extracellular Loops in Proton Sensing and Permeation in Human Otop1 Proton Channel. *Commun. Biol.*, **2022**, *5* (1), 1110. <https://doi.org/10.1038/s42003-022-04085-2>.
- [385] Bondar, A.-N. Proton-Binding Motifs of Membrane-Bound Proteins: From Bacteriorhodopsin to Spike Protein S. *Front Chem*, **2021**, *9*, 685761. <https://doi.org/10.3389/fchem.2021.685761>.
- [386] Gu, W.; Helms, V. Tightly Connected Water Wires Facilitate Fast Proton Uptake at The Proton Entrance of Proton Pumping Proteins. *J. Am. Chem. Soc.*, **2009**, *131* (6), 2080–2081. <https://doi.org/10.1021/ja809301w>.

- [387] Shang, Y.; LeRouzic, V.; Schneider, S.; Bisignano, P.; Pasternak, G. W.; Filizola, M. Mechanistic Insights into the Allosteric Modulation of Opioid Receptors by Sodium Ions. *Biochemistry*, **2014**, *53* (31), 5140–5149. <https://doi.org/10.1021/bi5006915>.
- [388] Selvam, B.; Shamsi, Z.; Shukla, D. Universality of the Sodium Ion Binding Mechanism in Class A G-Protein-Coupled Receptors. *Angew. Chem. Int. Ed.*, **2018**, *57* (12), 3048–3053. <https://doi.org/10.1002/anie.201708889>.
- [389] Elemam, N. M.; Youness, R. A.; Hussein, A.; Shihab, I.; Yakout, N. M.; Elwany, Y. N.; Manie, T. M.; Talaat, I. M.; Maghazachi, A. A. Expression of GPR68, an Acid-Sensing Orphan G Protein-Coupled Receptor, in Breast Cancer. *Front. Oncol.*, **2022**, *12*, 847543. <https://doi.org/10.3389/fonc.2022.847543>.
- [390] Pavlova, N. N.; Thompson, C. B. The Emerging Hallmarks of Cancer Metabolism. *Cell Metab.*, **2016**, *23* (1), 27–47. <https://doi.org/10.1016/j.cmet.2015.12.006>.
- [391] Blasi, R. D.; Marbiah, M. M.; Siciliano, V.; Polizzi, K.; Ceroni, F. A Call for Caution in Analysing Mammalian Co-Transfection Experiments and Implications of Resource Competition in Data Misinterpretation. *Nat. Commun.*, **2021**, *12* (1), 2545. <https://doi.org/10.1038/s41467-021-22795-9>.
- [392] Sacks, V.; Marantz, Y.; Aagaard, A.; Checover, S.; Nachliel, E.; Gutman, M. The Dynamic Feature of the Proton Collecting Antenna of a Protein Surface. *Biochimica Et Biophysica Acta Bba - Bioenergetics*, **1998**, *1365* (1–2), 232–240. [https://doi.org/10.1016/s0005-2728\(98\)00073-5](https://doi.org/10.1016/s0005-2728(98)00073-5).
- [393] Gutman, M.; Nachliel, E. The Dynamic Aspects of Proton Transfer Processes. *Biochim. Biophys. Acta (BBA) - Bioenerg.*, **1990**, *1015* (3), 391–414. [https://doi.org/10.1016/0005-2728\(90\)90073-d](https://doi.org/10.1016/0005-2728(90)90073-d).
- [394] Ädelroth, P.; Brzezinski, P. Surface-Mediated Proton-Transfer Reactions in Membrane-Bound Proteins. *Biochim. Biophys. Acta (BBA) - Bioenerg.*, **2004**, *1655* (1–3), 102–115. <https://doi.org/10.1016/j.bbabi.2003.10.018>.
- [395] Marantz, Y.; Nachliel, E.; Aagaard, A.; Brzezinski, P.; Gutman, M. The Proton Collecting Function of the Inner Surface of Cytochrome c Oxidase from Rhodospirillum rubrum. *Proc National Acad Sci*, **1998**, *95* (15), 8590–8595. <https://doi.org/10.1073/pnas.95.15.8590>.
- [396] Frishman, D.; Argos, P. Knowledge-based Protein Secondary Structure Assignment. *Proteins: Struct., Funct., Bioinform.*, **1995**, *23* (4), 566–579. <https://doi.org/10.1002/prot.340230412>.
- [397] Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol Syst Biol*, **2011**, *7* (1), 539–539. <https://doi.org/10.1038/msb.2011.75>.

# Appendices

## Appendix 1: GPR68 constructs

In summary 100 constructs were designed for the investigation of GPR68. The GPR68 constructs have been assigned an abbreviation e.g., C10, where C stands for the word ‘construct’ and the number ‘10’ is the construct number. The constructs C1-C74 and C99-C100 were designed for insect cell expression and the rest of the constructs were designed for HEK cell expression. All the constructs are designed for human GPR68, unless explicitly mentioned. For optimal readability, different colors have been used to highlight various segments of the amino acid sequence within the GPR68 construct sequences shown below. Mutations are marked in red bold font.

### Legend:

HA – Hemagglutinin signal peptide

FLAG – FLAG tag

His10 – Deca histidine tag

TST – Twin-strep-tag

V5 – V5 tag

TEV – Tobacco etch virus protease cleavage site (ENLYFQ/G), ‘/’ represents cleavage site

PreC – PreScission<sup>TM</sup> human rhinovirus (HRV) 3C protease cleavage site (LEVLFQ/GP), ‘/’ represents cleavage site

BRIL – A thermostabilized apocytochrome from *Escherichia coli* b562 with mutations

BRIL $\chi$  – The BRIL fusion with additional linker amino acids derived from Adenosine A<sub>2a</sub> receptor e.g., C17, C30 and C61

T4L – T4 lysozyme (T4L)

mT4L – modified T4 lysozyme (T4L)

PGS – Glycogen synthase from *Pyrococcus abyssi*

RDX – Rubredoxin

GFP – Green fluorescent protein

BLEK –  $\beta$ -Lactamase from *Escherichia coli* K12

Cal A/ Cal B – Rat calcineurin sub-unit A / Rat calcineurin sub-unit B

*The following constructs were designed for insect cell expression:*

*GPR68 C1: HA-FLAG-GPR68 (1-365)-PreC-His10*

MKTIIALSYIFCLVFA DYKDDDDK MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGF PANCLS  
LYFGYLQIKARNELGVYLCNLT VADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYIS  
VGFLCCISVDRYLAV AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE  
HYPIQAWQRAINYYRFLV GFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACF  
LPYHVLLL VRSVWEASCDFAKGVFNAYHFSLLLTSFNCVADPVLYCFVSETTHRDLARLRGACLA  
FLTCSRTRGRAREAYPLGAPEASGKSGAQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA LEVLFQGP  
HHHHHHHHHH

*GPR68 C2: HA-FLAG-GPR68(1-216, GS\_T4L\_GS, 217-365)-PreC-His10*

MKTIIALSYIFCLVFA DYKDDDDK MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGF PANCLS  
LYFGYLQIKARNELGVYLCNLT VADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYIS  
VGFLCCISVDRYLAV AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE  
HYPIQAWQRAINYYRFLV GFLFPICLLASYQGILRAVRRSHGSNIFEMLRIDEGLRLKIYKDTEGY  
YTIGIGHLLTKSPSLNAAKSELDKAIGRNTNGVITKDEAEKLFNQDVDAAVRGILRNAKLKPVYDS  
LDVRRALINMVFQMGETGVAGFTNSLRMLQQKRWDEAAVNLA KSRWYNQTPNRAKRVIITF  
RTGTWDA YGS GTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLL VRSVWEASCDFAKGVFNAYHFS  
LLLTSFNCVADPVLYCFVSETTHRDLARLRGACLAFLTCSRTRGRAREAYPLGAPEASGKSGAQGEE  
PELLTKLHPAFQTPNSPGSGGFPTGRLA LEVLFQGP HHHHHHHHHHH

*GPR68 C3: HA-FLAG-GPR68(1-216, BRIL, 217-365)-Pre-His10*

MKTIIALSYIFCLVFA DYKDDDDK MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGF PANCLS  
LYFGYLQIKARNELGVYLCNLT VADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYIS  
VGFLCCISVDRYLAV AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE  
HYPIQAWQRAINYYRFLV GFLFPICLLASYQGILRAVRRSHADLEDNWETLNDNLK VIEKADNAA  
QVKDALTKMRAAALDAQKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQ  
AAAEQLKTRNAYIQKYLGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLL VRSVWEASCDFAKGV  
FNAYHFSLLLTSFNCVADPVLYCFVSETTHRDLARLRGACLAFLTCSRTRGRAREAYPLGAPEASG  
KSGAQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA LEVLFQGP HHHHHHHHHHH

*GPR68 C4: HA-FLAG-GPR68(1-216, PGS, 217-365)-PreC-His10*

MKTIIALSYIFCLVFA DYKDDDDK MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGF PANCLS  
LYFGYLQIKARNELGVYLCNLT VADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYIS  
VGFLCCISVDRYLAV AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE  
HYPIQAWQRAINYYRFLV GFLFPICLLASYQGILRAVRRSHGIDCSFWNESYLTGSRDERKKSLLS  
KFGMDEGVTFMFIGRFD RGQKGV DVLLKAIEILSSKKEFQEMRFIIIGKGDPELEGWARSLEEKHG  
NVKVITEMLSRE FVRELYGSVDFVIIPSYFEPFGLVALEAMCLGAIPIASAVGGLRDIITNETGILVKA  
GDPGELANAILKALELSRSDL SKFRENCKRAMSFGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLL  
LVRSVWEASCDFAKGVFNAYHFSLLLTSFNCVADPVLYCFVSETTHRDLARLRGACLAFLTCSRTR  
GRAREAYPLGAPEASGKSGAQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA LEVLFQGP HHHHHH  
HHHH

*GPR68 C5: HA-FLAG-GPR68(1-216,GS\_mT4L\_GS, 217-365)-PreC-His10*

MKTIALSIFCLVFA DYKDDDDK MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGF PANCLS  
LYFGYLQIKARNELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYIS  
VGF LCCISVDRYLAV AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE  
HYPIQAWQRAINYYRFLVGFLFPICLLASYQGILRAVRRSHGSMNIFEMLRIDE GGGSGGDEAEKL  
FNQDVDAAVRGILRNAKLPVYDSLDAVRRRAALINMVFQMGETGVAGFTNSLRMLQQKRWDEA  
AVNLAKSRWYNQTPNRAKR VITTFRTGTWDA YSGTQKSRKDQIQRLVLSTVVIFLACFLPYHVL  
LLVRSVWEASCDFAKGVFNAYHFSLLLSFNVCVADPVLVCFVSETTHRD LARLRGACLAFLTC SRT  
GRAREAYPLGAPEASGKSGAQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA LEVLFQGP HHHHHH  
HHHH

*GPR68 C6: HA-FLAG-GPR68(1-214, RDX, 220-365)-PreC-His10*

MKTIALSIFCLVFA DYKDDDDK MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGF PANCLS  
LYFGYLQIKARNELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYIS  
VGF LCCISVDRYLAV AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE  
HYPIQAWQRAINYYRFLVGFLFPICLLASYQGILRAVRRMKKYTCTVCGYIYNPEDGDPDNGVNP  
GTDFK DIPDDWVCP LCGV GKDQFEEVEE KSRKDQIQRLVLSTVVIFLACFLPYHVL L VRSVWEAS  
CDFAKGVFNAYHFSLLLSFNVCVADPVLVCFVSETTHRD LARLRGACLAFLTC SRTGRAREAYPLG  
APEASGKSGAQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA LEVLFQGP HHHHHHHHHH

*GPR68 C7: HA-FLAG-GPR68(BRIL\_GGS-PreC-9-216, GS\_T4L\_GS, 217-365)-TEV-His10*

MKTIALSIFCLVFA DYKDDDDK ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY  
LGGSLLEVLFQGPSSMSCTIDHTIHQTLAPVVYVTVLVVGF PANCLS LYFGYLQIKARNELGVYLCN  
LTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGF LCCISVDRYLAV AHPFRF  
HQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFL  
FPICLLASYQGILRAVRRSHGSMNIFEMLRIDEGLRLKIYKDTEGYTIGIHLTKSPSLNAAKSELD  
KAIGRNTNGVITKDEAEKLFNQDVDAAVRGILRNAKLPVYDSLDAVRRRAALINMVFQMGETGV  
AGFTNSLRMLQQKRWDEAAVNLAKSRWYNQTPNRAKR VITTFRTGTWDA YSGTQKSRKDQIQ  
RLVLSTVVIFLACFLPYHVL L VRSVWEASCDFAKGVFNAYHFSLLLSFNVCVADPVLVCFVSETT  
HRDLARLRGACLAFLTC SRTGRAREAYPLGAPEASGKSGAQGEEPELLTKLHPAFQTPNSPGSGGF  
PTGRLA LEVLFQGP HHHHHHHHHH

*GPR68 C8: HA-FLAG-GPR68(BRIL\_GGS-PreC-9-365)-TEV-His10*

MKTIALSIFCLVFA DYKDDDDK ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY  
LGGSLLEVLFQGPSSMSCTIDHTIHQTLAPVVYVTVLVVGF PANCLS LYFGYLQIKARNELGVYLCN  
LTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGF LCCISVDRYLAV AHPFRF  
HQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFL  
FPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVL L VRSVWEASCDFAK  
GVFNAYHFSLLLSFNVCVADPVLVCFVSETTHRD LARLRGACLAFLTC SRTGRAREAYPLGAPEAS  
GKSGAQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA ENLYFQGHHHHHHHHHH

*GPR68 C9: HA-FLAG-GPR68(9-216,GS\_T4L\_GS, 217-365)-PreC-His10*

MKTIALSIFCLVFA DYKDDDDK SSMSC TIDHTIHQTLAPVVYVTVLVVGF PANCLS LYFGYLQIK  
ARNELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGF LCCISV  
DRYLAV AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQ  
RAINYYRFLVGFLFPICLLASYQGILRAVRRSHGSMNIFEMLRIDEGLRLKIYKDTEGYTIGIHLTK  
KSPSLNAAKSELDKAIGRNTNGVITKDEAEKLFNQDVDAAVRGILRNAKLPVYDSLDAVRRRAAL  
INMVFQMGETGVAGFTNSLRMLQQKRWDEAAVNLAKSRWYNQTPNRAKR VITTFRTGTWDA YG

SGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAYHFSLLLSFNCA  
DPVLYCFVSETTHRDARLRGACLAFLTCRTRAREAYPLGAPEASGKSGAQGEEPELLTKLHPA  
FQTPNSPGSGGFPTGRLA **LEVLFQGP**HHHHHHHHHH

*GPR68 C10: HA-FLAG-GPR68(15-216, GS\_T4L\_GS, 217-365)-PreC-His10*

**MKTI**ALS**YIFCLVFA****DYKDDDDK**IDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELG  
VYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGLCCISVDRYLAV  
AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAINYYR  
FLVGLFPICLLASYQGILRAVRRSH**GSNIFEMLRIDEGLRLKIYKDTEGYTIGIGHLLTKSPSLNA**  
**AKSELDKAIGRNTNGVITKDEAEKLFNQDVDAAVRGILRNAKLKPVYDSLDAVRRRAALINMVFQ**  
**MGETGVAGFTNSLRMLQQRWDEAAVNLA**KSRWYNQTPNRAKR**VITTFRTGTWDA**YGS**GTQKS**  
RKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAYHFSLLLSFNCA  
DPVLYCFVSETTHRDARLRGACLAFLTCRTRAREAYPLGAPEASGKSGAQGEEPELLTKLHPAFQTPNS  
PGSGGFPTGRLA **LEVLFQGP**HHHHHHHHHH

*GPR68 C11: HA-FLAG-GPR68(21-216, GS\_T4L\_GS, 217-365)-PreC-His10*

**MKTI**ALS**YIFCLVFA****DYKDDDDK**QTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCN  
LTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGLCCISVDRYLAVAH  
PFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAINYYR  
FLVGLFPICLLASYQGILRAVRRSH**GSNIFEMLRIDEGLRLKIYKDTEGYTIGIGHLLTKSPSLNAKSELD**  
**KAIGRNTNGVITKDEAEKLFNQDVDAAVRGILRNAKLKPVYDSLDAVRRRAALINMVFQMGETGV**  
**AGFTNSLRMLQQRWDEAAVNLA**KSRWYNQTPNRAKR**VITTFRTGTWDA**YGS**GTQKSRKDQIQ**  
RLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAYHFSLLLSFNCA  
DPVLYCFVSETTHRDARLRGACLAFLTCRTRAREAYPLGAPEASGKSGAQGEEPELLTKLHPAFQTPNS  
PGSGGFPTGRLA **LEVLFQGP**HHHHHHHHHH

*GPR68 C12: HA-FLAG-GPR68(1-216, GS\_T4L\_GS, 217-347)-PreC-His10*

**MKTI**ALS**YIFCLVFA****DYKDDDDK**MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSL  
LYFGYLQIKARNELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYIS  
VGLCCISVDRYLAVAH  
PFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFE  
HYPIQAWQRAINYYRFLVGLFPICLLASYQGILRAVRRSH**GSNIFEMLRIDEGLRLKIYKDTEGY**  
**YTIGIGHLLTKSPSLNAKSELDKAIGRNTNGVITKDEAEKLFNQDVDAAVRGILRNAKLKPVYDS**  
**LDAVRRRAALINMVFQMGETGVAGFTNSLRMLQQRWDEAAVNLA**KSRWYNQTPNRAKR**VITTF**  
**RTGTWDA**YGS**GTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAYHFS**  
LLLSFNCA  
DPVLYCFVSETTHRDARLRGACLAFLTCRTRAREAYPLGAPEASGKSGAQGEE  
PELLTKLHPA **LEVLFQGP**HHHHHHHHHH

*GPR68 C13: HA-FLAG-GPR68(1-216, GS\_T4L\_GS, 217-313)-PreC-His10*

**MKTI**ALS**YIFCLVFA****DYKDDDDK**MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSL  
LYFGYLQIKARNELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYIS  
VGLCCISVDRYLAVAH  
PFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFE  
HYPIQAWQRAINYYRFLVGLFPICLLASYQGILRAVRRSH**GSNIFEMLRIDEGLRLKIYKDTEGY**  
**YTIGIGHLLTKSPSLNAKSELDKAIGRNTNGVITKDEAEKLFNQDVDAAVRGILRNAKLKPVYDS**  
**LDAVRRRAALINMVFQMGETGVAGFTNSLRMLQQRWDEAAVNLA**KSRWYNQTPNRAKR**VITTF**  
**RTGTWDA**YGS**GTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAYHFS**  
LLLSFNCA  
DPVLYCFVSETTHRDARLRGACLAFLTCRTRAREAYPLGAPEASGKSGAQGEE  
PELLTKLHPA **LEVLFQGP**HHHHHHHHHH

*GPR68 C14: HA-FLAG-GPR68(1-216, GS\_T4L\_GS, 217-310)-PreC-His10*

MKTIIALSYIFCLVFA DYKDDDDK MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLS  
LYFGYLQIKARNELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYIS  
VGFLCCISVDRYLAVAHPPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE  
HYPIQAWQRAINYYRFLVGFLFPICLLASYQGILRAVRRSHGSNIFEMLRIDEGLRLKIYKDTEGY  
YTIGIGHLLTKSPSLNAAKSELDKAIGRNTNGVITKDEAEKLFNQDVDAAVRGILRNAKLKPVYDS  
LDVRRRAALINMVFQMGETGVAGFTNSLRMLQQKRWDEAAVNLAKSRWYNQTPNRAKRVIITF  
RTGTWDAYSGTQKSRKDQIQLVLSVTVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAYHFS  
LLTSFNVCVADPVLYCFVSETTHRDLARLRGACLAFLTC LEVLFQGP HHHHHHHHHH

*GPR68 C15: HA-FLAG-GPR68(1-216, GS\_T4L\_GS, 217-305)-PreC-His10*

MKTIIALSYIFCLVFA DYKDDDDK MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLS  
LYFGYLQIKARNELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYIS  
VGFLCCISVDRYLAVAHPPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE  
HYPIQAWQRAINYYRFLVGFLFPICLLASYQGILRAVRRSHGSNIFEMLRIDEGLRLKIYKDTEGY  
YTIGIGHLLTKSPSLNAAKSELDKAIGRNTNGVITKDEAEKLFNQDVDAAVRGILRNAKLKPVYDS  
LDVRRRAALINMVFQMGETGVAGFTNSLRMLQQKRWDEAAVNLAKSRWYNQTPNRAKRVIITF  
RTGTWDAYSGTQKSRKDQIQLVLSVTVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAYHFS  
LLTSFNVCVADPVLYCFVSETTHRDLARLRGACL LEVLFQGP HHHHHHHHHH

*GPR68 C16: HA-FLAG-GPR68(1-214, BRIL, 219-365)-PreC-His10*

MKTIIALSYIFCLVFA DYKDDDDK MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLS  
LYFGYLQIKARNELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYIS  
VGFLCCISVDRYLAVAHPPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE  
HYPIQAWQRAINYYRFLVGFLFPICLLASYQGILRAVRRADLEDNWTLNLDNLKVIKADNAAQ  
VKDALTKMRAAALDAQKATPPKLEDKSPDSEPMKDFRHGFDILVGQIDDALKLANEGKVKEAQA  
AAEQLKTRNAYIQKYLQKSRKDQIQLVLSVTVIFLACFLPYHVLLLVRSVWEASCDFAKGVFN  
AYHFSLLTSFNVCVADPVLYCFVSETTHRDLARLRGACLAFLTC SRTGRAREAYPLGAPEASGKSG  
AQGEPELLTKLHPAFQTPNSPGSGGFPTGRLA LEVLFQGP HHHHHHHHHH

*GPR68 C17: HA-FLAG-GPR68(1-211, ARRQL\_BRIL\_ERARSTL, 221-365)-PreC-His10*

BRIL fusion is connected to TM5 and TM6 by two short peptides (ARRQL and ERARSTL) derived from TM5 and TM6 of A<sub>2A</sub> adenosine receptor.

MKTIIALSYIFCLVFA DYKDDDDK MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLS  
LYFGYLQIKARNELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYIS  
VGFLCCISVDRYLAVAHPPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE  
HYPIQAWQRAINYYRFLVGFLFPICLLASYQGILRAARRQLADLEDNWTLNLDNLKVIKADNAAQ  
QVKDALTKMRAAALDAQKATPPKLEDKSPDSEPMKDFRHGFDILVGQIDDALKLANEGKVKEAQA  
AAEQLKTRNAYIQKYL ERARSTL SRKDQIQLVLSVTVIFLACFLPYHVLLLVRSVWEASCDF  
KGVFNAYHFSLLTSFNVCVADPVLYCFVSETTHRDLARLRGACLAFLTC SRTGRAREAYPLGAPEA  
SGKSGAQGEPELLTKLHPAFQTPNSPGSGGFPTGRLA LEVLFQGP HHHHHHHHHH

*GPR68 C18: HA-FLAG-GPR68(BRIL\_GGS-PreC-9-216, PGS, 217-365)-TEV-His10*

MKTIIALSYIFCLVFA DYKDDDDK ADLEDNWTLNLDNLKVIKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSEPMKDFRHGFDILVGQIDDALKLANEGKVKEAQA AAEQLKTRNAYIQKY  
L GGS LEVLFQGP SSMSC TIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCN  
LTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRF  
HQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE HYPIQAWQRAINYYRFLVGFL  
FPICLLASYQGILRAVRRSHGIDCSFWNESYLTGSRDERKKSLLSKFGMDEGVTFMFIGRFRDGRQK  
GVDVLLKAI EILSSKKEFQEMRFIIIGKGDPELEGWARSLEEKHGNVKVITEMLSREVFVRELYGSVD

FVIIPSYFEPFGLVALEAMCLGAIPIASAVGGLRDIITNETGILVKAGDPGELANAILKALELSRSDLS  
KFRENCKKRAMSFSGTQKSRKDQIQLVLSTVVIFLACFLPYHVLLLVRVSVWEASCDFAKGVFNA  
YHFSLLLTSFNCVADPVLVYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSGA  
QGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA **ENLYFQG** **HHHHHHHHHH**

*GPR68 C19: HA-FLAG-GPR68(BRIL\_GS-PreC-9-216, PGS, 217-365)-TEV-His10*  
(Mutation: N86<sup>23,49</sup>Q)

**MKTI**ALS**YIFCLVFA** **DYKDDDDK**ADLEDN**WETLNDNLK**VIEKAD**NAAQVKDAL**TKMRAAALDA  
QKATPPK**LEDKSPD**SP**EMKDFR**HGFDILVGQID**DALKLANEGKVKEA**QAAAEQLK**TTR**NA**YIQY**  
**LGG**S**LEVLFQGP**SSMSCTIDHTI**HQTLAPV**VYVTVLVVGF**PANCL**SLYFGYLQIKAR**NELG**VYLCN  
LTVADLFYICSLPFWLQYVLQHD**Q**WSHGDLSCQVCGILLYENIYISVGF**LCCISVDRYLAV**AHPFRF  
HQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFL  
FPICLLAS**YQGILRAVRRSH**GIDCSFWNESYLTGSRDERK**KSLLSKFGMDEG**VTFM**FIGR**FDRGQK  
GVDVLLKAI**EILSSKKEFQEMR**FI**IGKGDPELEG**WARSLEEKHG**NVKVIT**EMLSREFVRELYGSVD  
FVIIPSYFEPFGLVALEAMCLGAIPIASAVGGLRDIITNETGILVKAGDPGELANAILKALELSRSDLS  
KFRENCKKRAMSFSGTQKSRKDQIQLVLSTVVIFLACFLPYHVLLLVRVSVWEASCDFAKGVFNA  
YHFSLLLTSFNCVADPVLVYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSGA  
QGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA **ENLYFQG** **HHHHHHHHHH**

*GPR68 C20: HA-FLAG-GPR68(15-216, GS\_T4L\_GS, 217-365)-PreC-TST*

**MKTI**ALS**YIFCLVFA** **DYKDDDDK**IDHTI**HQTLAPV**VYVTVLVVGF**PANCL**SLYFGYLQIKAR**NELG**  
VYLCNLT**VADLFYICSLPFWLQYVLQHD**NWSHGDLSCQVCGILLYENIYISVGF**LCCISVDRYLAV**  
AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYR  
FLVGFLFPICLLAS**YQGILRAVRRSH**G**SNIFEMLRIDEGLRLKIYKDTEGY**YTIGIGHLLTK**SPSLNA**  
AK**SELDK**AI**GRNTNGVITKDEAEKLF**NQDVDA**AVRGILRNAK**LKP**VYD**SLDAV**RR**AALIN**MVFQ**  
MGETGVAGFTNSLRMLQ**QKRWDEAAVN**LAKSRWYNQTP**NR**AKRVITTFRTGTWDA**YGS**GTQKS  
RKDQIQRLVLSTVVIFLACFLPYHVLLLVRVSVWEASCDFAKGVFNAYHFSLLLTSFNCVADPVLVY  
FVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSGAQGEEPELLTKLHPAFQTPNS  
PGSGGFPTGRLA **LEVLFQGP** **WSHPQFEKGGGSGGGSSAW**SH**PQFEK**

*GPR68 C21: HA-FLAG-GPR68(15-216, GS\_T4L\_GS, 217-347)-PreC-His10*

**MKTI**ALS**YIFCLVFA** **DYKDDDDK**IDHTI**HQTLAPV**VYVTVLVVGF**PANCL**SLYFGYLQIKAR**NELG**  
VYLCNLT**VADLFYICSLPFWLQYVLQHD**NWSHGDLSCQVCGILLYENIYISVGF**LCCISVDRYLAV**  
AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYR  
FLVGFLFPICLLAS**YQGILRAVRRSH**G**SNIFEMLRIDEGLRLKIYKDTEGY**YTIGIGHLLTK**SPSLNA**  
AK**SELDK**AI**GRNTNGVITKDEAEKLF**NQDVDA**AVRGILRNAK**LKP**VYD**SLDAV**RR**AALIN**MVFQ**  
MGETGVAGFTNSLRMLQ**QKRWDEAAVN**LAKSRWYNQTP**NR**AKRVITTFRTGTWDA**YGS**GTQKS  
RKDQIQRLVLSTVVIFLACFLPYHVLLLVRVSVWEASCDFAKGVFNAYHFSLLLTSFNCVADPVLVY  
FVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSGAQGEEPELLTKLHPA **LEVLFQ**  
**GP** **HHHHHHHHHH**

*GPR68 C22: HA-FLAG-GPR68(15-216, GS\_T4L\_GS, 217-313)-PreC-His10*

**MKTI**ALS**YIFCLVFA** **DYKDDDDK**IDHTI**HQTLAPV**VYVTVLVVGF**PANCL**SLYFGYLQIKAR**NELG**  
VYLCNLT**VADLFYICSLPFWLQYVLQHD**NWSHGDLSCQVCGILLYENIYISVGF**LCCISVDRYLAV**  
AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYR  
FLVGFLFPICLLAS**YQGILRAVRRSH**G**SNIFEMLRIDEGLRLKIYKDTEGY**YTIGIGHLLTK**SPSLNA**  
AK**SELDK**AI**GRNTNGVITKDEAEKLF**NQDVDA**AVRGILRNAK**LKP**VYD**SLDAV**RR**AALIN**MVFQ**  
MGETGVAGFTNSLRMLQ**QKRWDEAAVN**LAKSRWYNQTP**NR**AKRVITTFRTGTWDA**YGS**GTQKS  
RKDQIQRLVLSTVVIFLACFLPYHVLLLVRVSVWEASCDFAKGVFNAYHFSLLLTSFNCVADPVLVY  
FVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSGAQGEEPELLTKLHPA **LEVLFQGP** **HHHHHHHHHH**

*GPR68 C23: HA-FLAG-GPR68(15-216, GS\_T4L\_GS, 217-305)-PreC-His10*

MKTIALSIFCLVFA DYKDDDDKIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELG  
VYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAV  
AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHIEEVIEDENQHRVCFEHYPIQAWQRAINYYR  
FLVGFLFPICLLASYQGILRAVRRSHGSNIFEMLRIDEGLRLKIYKDEGYTIGIGHLLTKSPSLNA  
AKSELDKAIGRNTNGVITKDEAEKLFNQDVDAAVRGILRNAKLPVYDSLDAVRRRAALINMVFQ  
MGETGVAGFTNSLRMLQQRWDEAAVNLAKSRYWYQTPNRAKRVTTFRTGTWDAVYSGTQKS  
RKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAYHFSLLTTSFNVCVADPVLYC  
FVSETTHRDLARLRGACL LEVLFQGP HHHHHHHHHH

*GPR68 C24: HA-FLAG-GPR68(15-216, PGS, 217-347)-PreC-His10*

MKTIALSIFCLVFA DYKDDDDKIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELG  
VYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAV  
AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHIEEVIEDENQHRVCFEHYPIQAWQRAINYYR  
FLVGFLFPICLLASYQGILRAVRRSHGIDCSFWNESYLTGSRDERKKSLLSKFGMDEGVTFMFIGR  
FDRGQKGVVDVLLKAIEILSSKKEFQEMRFIIIGKGDPELEGWARSLEEKHGNVKVITEMLSREFVRE  
LYGSVDFVIIPSYFEPFGLVALEAMCLGAIPASAVGGLRDIITNETGILVKAGDPGELANAILKALEL  
SRSDLSKFRENCKKRAMSFSGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAK  
GVFNAYHFSLLTTSFNVCVADPVLYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEAS  
GKSGAQGEEPELLTKLHPA LEVLFQGP HHHHHHHHHH

*GPR68 C25: HA-FLAG-GPR68(15-216, PGS, 217-313)-PreC-His10*

MKTIALSIFCLVFA DYKDDDDKIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELG  
VYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAV  
AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHIEEVIEDENQHRVCFEHYPIQAWQRAINYYR  
FLVGFLFPICLLASYQGILRAVRRSHGIDCSFWNESYLTGSRDERKKSLLSKFGMDEGVTFMFIGR  
FDRGQKGVVDVLLKAIEILSSKKEFQEMRFIIIGKGDPELEGWARSLEEKHGNVKVITEMLSREFVRE  
LYGSVDFVIIPSYFEPFGLVALEAMCLGAIPASAVGGLRDIITNETGILVKAGDPGELANAILKALEL  
SRSDLSKFRENCKKRAMSFSGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAK  
GVFNAYHFSLLTTSFNVCVADPVLYCFVSETTHRDLARLRGACLAFLTCRT LEVLFQGP HHHHHH  
HHHH

*GPR68 C26: HA-FLAG-GPR68(15-216, PGS, 217-305)-PreC-His10*

MKTIALSIFCLVFA DYKDDDDKIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELG  
VYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAV  
AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHIEEVIEDENQHRVCFEHYPIQAWQRAINYYR  
FLVGFLFPICLLASYQGILRAVRRSHGIDCSFWNESYLTGSRDERKKSLLSKFGMDEGVTFMFIGR  
FDRGQKGVVDVLLKAIEILSSKKEFQEMRFIIIGKGDPELEGWARSLEEKHGNVKVITEMLSREFVRE  
LYGSVDFVIIPSYFEPFGLVALEAMCLGAIPASAVGGLRDIITNETGILVKAGDPGELANAILKALEL  
SRSDLSKFRENCKKRAMSFSGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAK  
GVFNAYHFSLLTTSFNVCVADPVLYCFVSETTHRDLARLRGACL LEVLFQGP HHHHHHHHHH

*GPR68 C27: HA-FLAG-GPR68(BRIL\_GGS-PreC-15-347)-TEV-His10*

MKTIALSIFCLVFA DYKDDDDKADLEDNWTLNLDNLKVIKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDPEMDFRHFDFILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY  
LGGSL LEVLFQGP IDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVADL  
FYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAV AHPFRFHQFRTL  
KAAVGVSVVIWAKELLTSIYFLMHIEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPICLL  
ASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAY  
HFSLLTTSFNVCVADPVLYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSGAQ  
GEEPELLTKLHPA ENLYFQG HHHHHHHHHH

*GPR68 C28: HA-FLAG-GPR68(BRIL\_GGS-PreC-15-313)-TEV-His10*

MKTIIALSYIFCLVFA**DYKDDDDK**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAALDA  
QKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGG**SEVLFQGP**IDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVADL  
FYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFRTL  
KAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPICLL  
ASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAY  
HFSLLTTSFNVCVADPVLYCFVSETTHRDARLRGACLAFLTCSRT**ENLYFQG**HHHHHHHHHHH

*GPR68 C29: HA-FLAG-GPR68(BRIL\_GGS-PreC-15-305)-TEV-His10*

MKTIIALSYIFCLVFA**DYKDDDDK**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAALDA  
QKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGG**SEVLFQGP**IDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVADL  
FYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFRTL  
KAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPICLL  
ASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAY  
HFSLLTTSFNVCVADPVLYCFVSETTHRDARLRGACL**ENLYFQG**HHHHHHHHHHH

*GPR68 C30: HA-FLAG-GPR68(15-211, ARRQL\_BRIL\_ERARSTL, 221-365)-PreC-His10*

BRIL fusion is connected to TM5 and TM6 by two short peptides (ARRQL and ERARSTL) derived from TM5 and TM6 of A<sub>2A</sub> adenosine receptor.

MKTIIALSYIFCLVFA**DYKDDDDK**IDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELG  
VYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRYLAV  
AHPPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYR  
FLVGFLFPICLLASYQGILRA**ARRQL**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAL  
**DAQKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQ**  
**KYLERARSTL**SRKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAYHFSLLTTSF  
NCVADPVLYCFVSETTHRDARLRGACLAFLTCSRTGRAREAYPLGAPEASGKSGAQGEEPELLT  
KLHPAFQTPNSPGSGGFPTGRLA**LEVLFQGP**HHHHHHHHHHH

*GPR68 C31: HA-FLAG-GPR68(1-214, BRIL, 219-347)-PreC-His10*

MKTIIALSYIFCLVFA**DYKDDDDK**MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCL  
LYFGYLQIKARNELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYIS  
VGFLCCISVDRYLAVAHPPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE  
HYPIQAWQRAINYYRFLVGFLFPICLLASYQGILRAVRR**ADLEDNWETLNDNLKVIEKADNAAQ**  
**VKDALT**TKMRAALDA**QKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQA**  
**AAEQLKTRNAYIQKYL**QKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFN  
AYHFSLLTTSFNVCVADPVLYCFVSETTHRDARLRGACLAFLTCSRTGRAREAYPLGAPEASGKSG  
AQGEEPELLTKLHPA**LEVLFQGP**HHHHHHHHHHH

*GPR68 C32: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-365)-TEV-His10*

MKTIIALSYIFCLVFA**DYKDDDDK**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAALDA  
QKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGG**SEVLFQGP**MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRY  
LAVAHPPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRSV  
WEASCDFAKGVFNAYHFSLLTTSFNVCVADPVLYCFVSETTHRDARLRGACLAFLTCSRTGRARE  
AYPLGAPEASGKSGAQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA**ENLYFQG**HHHHHHHHHHH

GPR68 C33: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10

MKTIIALSYIFCLVFA**DYKDDDDK**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSEPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY  
LGGS**LEVLFQGP**MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGGFPANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRY  
LAVAHPPRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVSTVVIFLACFLPYHVLLLVRVSV  
WEASCDFAKGVFNAYHFSLLTSFNVCVADPVLVYCFVSETTHRDLARLRGACLAFLTCRSRTGRARE  
AYPLGAPEASGKSGAQGEEPELLTKLHPA**ENLYFQG**HHHHHHHHHH

GPR68 C34: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-313)-TEV-His10

MKTIIALSYIFCLVFA**DYKDDDDK**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSEPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY  
LGGS**LEVLFQGP**MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGGFPANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRY  
LAVAHPPRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVSTVVIFLACFLPYHVLLLVRVSV  
WEASCDFAKGVFNAYHFSLLTSFNVCVADPVLVYCFVSETTHRDLARLRGACLAFLTCRSRT**ENLYFQ**  
**QG**HHHHHHHHHH

GPR68 C35: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-305)-TEV-His10

MKTIIALSYIFCLVFA**DYKDDDDK**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSEPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY  
LGGS**LEVLFQGP**MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGGFPANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRY  
LAVAHPPRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVSTVVIFLACFLPYHVLLLVRVSV  
WEASCDFAKGVFNAYHFSLLTSFNVCVADPVLVYCFVSETTHRDLARLRGACL**ENLYFQG**HHHHH  
HHHHH

GPR68 C36: HA-FLAG-GPR68(BRIL\_GS-PreC-9-216, PGS, 217-365)-TEV-His10  
(Mutation: N86<sup>23,49</sup>Q, F241<sup>6,48</sup>W)

MKTIIALSYIFCLVFA**DYKDDDDK**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSEPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY  
LGGS**LEVLFQGP**SSMSCTIDHTIHQTLAPVVYVTVLVVGGFPANCLSLYFGYLQIKARNELGVYLCN  
LTVADLFYICSLPFWLQYVLQHD**Q**WSHGDLSQVCGILLYENIYISVGFLCCISVDRYLAVAHPPRF  
HQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFL  
FPICLLASYQGILRAVRRSHGIDCSFVNESYLTGSRDERKKSLLSKFGMDEGVTFMFIGRFDRGQK  
GVDVLLKAIEILSSKKEFQEMRFIIIGKGDPELEGWARSLEEKHGNVKVITEMLSREVFRELYGSVD  
FVIIPSYFEPFGLVALEAMCLGAIPIASAVGGLRDIITNETGILVKAGDPGELANAILKALELSRSDLS  
KFRENCKKRAMSFS**GTQKSRKDQIQRLVSTVVIFLAC****W**LPHYHVLLLVRVSVWEASCDFAKGVFN  
AYHFSLLTSFNVCVADPVLVYCFVSETTHRDLARLRGACLAFLTCRSRTGRAREAYPLGAPEASGKSGA  
QGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA**ENLYFQG**HHHHHHHHHH

GPR68 C37: HA-FLAG-GPR68(BRIL\_GS-PreC-9-216, PGS, 217-365)-TEV-His10  
(Mutation: N86<sup>23,49</sup>Q, I226<sup>6,33</sup>W)

MKTIIALSYIFCLVFA**DYKDDDDK**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSEPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY  
LGGS**LEVLFQGP**SSMSCTIDHTIHQTLAPVVYVTVLVVGGFPANCLSLYFGYLQIKARNELGVYLCN  
LTVADLFYICSLPFWLQYVLQHD**Q**WSHGDLSQVCGILLYENIYISVGFLCCISVDRYLAVAHPPRF  
HQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFL

FPICLLASYQGILRAVRRSHGIDCSFWNESYLTGSRDERKKSLLSKFGMDEGVTFMFIGRFDRGQK  
GVDVLLKAIEILSSKKEFQEMRFIIIGKDPELEGWARSLEEKHGNVKVITEMLSREFVRELYGSVD  
FVIIPSYFEPFGLVALEAMCLGAIPIASAVGGLRDIITNETGILVKAGDPGELANAILKALELSRSDLS  
KFRENCKKRAMSFSGTQKSRKDQDQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNA  
YHFSLLLTSFNCVADPVLVYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSGA  
QGEEPELLTKLHPAFQTPNSPGSGGFPTGRLAENLYFQGHHHHHHHHHH

*GPR68 C38: HA-FLAG-GPR68(BRIL\_GS-PreC-9-216, PGS, 217-365)-TEV-His10*  
(Mutation: N86<sup>23.49</sup>Q, G110<sup>3.41</sup>W)

MKTIIALSYIFLVFA<sup>DK</sup>ADLEDNWTLNLDNLK<sup>VIE</sup>KADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGGSL<sup>EVLFQGP</sup>SSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCN  
LTVADLFYICSLPFWLQYVLQHDQ<sup>W</sup>SHGDLSCQVCGILLYENIYISV<sup>W</sup>FLCCISVDRYLAVAHPRF  
FHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGF  
LPICLLASYQGILRAVRRSHGIDCSFWNESYLTGSRDERKKSLLSKFGMDEGVTFMFIGRFDRGQ  
KGVLDVLLKAIEILSSKKEFQEMRFIIIGKDPELEGWARSLEEKHGNVKVITEMLSREFVRELYGSV  
DFVIIPSYFEPFGLVALEAMCLGAIPIASAVGGLRDIITNETGILVKAGDPGELANAILKALELSRSDLS  
SKFRENCKKRAMSFSGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNA  
YHFSLLLTSFNCVADPVLVYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSGA  
QGEEPELLTKLHPAFQTPNSPGSGGFPTGRLAENLYFQGHHHHHHHHHH

*GPR68 C39: HA-FLAG-GPR68(BRIL\_GS-PreC-9-216, PGS, 217-365)-TEV-His10*  
(Mutation: N86<sup>23.49</sup>Q, Y205<sup>5.58</sup>A)

MKTIIALSYIFCLVFA<sup>DK</sup>ADLEDNWTLNLDNLK<sup>VIE</sup>KADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGGSL<sup>EVLFQGP</sup>SSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCN  
LTVADLFYICSLPFWLQYVLQHDQ<sup>W</sup>SHGDLSCQVCGILLYENIYISVGLFCCISVDRYLAVAHPRF  
FHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGF  
FPICLLAS<sup>A</sup>QGILRAVRRSHGIDCSFWNESYLTGSRDERKKSLLSKFGMDEGVTFMFIGRFDRGQK  
GVDVLLKAIEILSSKKEFQEMRFIIIGKDPELEGWARSLEEKHGNVKVITEMLSREFVRELYGSVD  
FVIIPSYFEPFGLVALEAMCLGAIPIASAVGGLRDIITNETGILVKAGDPGELANAILKALELSRSDLS  
KFRENCKKRAMSFSGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNA  
YHFSLLLTSFNCVADPVLVYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSGA  
QGEEPELLTKLHPAFQTPNSPGSGGFPTGRLAENLYFQGHHHHHHHHHH

*GPR68 C40: HA-FLAG-GPR68(BRIL\_GS-PreC-9-216, PGS, 217-365)-TEV-His10*  
(Mutation: N86<sup>23.49</sup>Q, D67<sup>2.50</sup>N)

MKTIIALSYIFCLVFA<sup>DK</sup>ADLEDNWTLNLDNLK<sup>VIE</sup>KADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGGSL<sup>EVLFQGP</sup>SSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCN  
LTVAN<sup>L</sup>LFYICSLPFWLQYVLQHDQ<sup>W</sup>SHGDLSCQVCGILLYENIYISVGLFCCISVDRYLAVAHPRF  
FHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGF  
FPICLLASYQGILRAVRRSHGIDCSFWNESYLTGSRDERKKSLLSKFGMDEGVTFMFIGRFDRGQK  
GVDVLLKAIEILSSKKEFQEMRFIIIGKDPELEGWARSLEEKHGNVKVITEMLSREFVRELYGSVD  
FVIIPSYFEPFGLVALEAMCLGAIPIASAVGGLRDIITNETGILVKAGDPGELANAILKALELSRSDLS  
KFRENCKKRAMSFSGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNA  
YHFSLLLTSFNCVADPVLVYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSGA  
QGEEPELLTKLHPAFQTPNSPGSGGFPTGRLAENLYFQGHHHHHHHHHH

GPR68 C41: HA-FLAG-GPR68(BRIL\_GS-PreC-9-216, PGS, 217-365)-TEV-His10  
(Mutation: N86<sup>23.49</sup>Q, D282<sup>7.49</sup>N)

MKTIIALSYIFCLVFA DYKDDDDK ADLEDN WETLNDNLKVIEKADNAAQVKDALTKMRAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGGSEVLFQGPSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCN  
LTVADLFYICSLPFWLQYVLQHDQ WSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPRF  
HQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFL  
FPICLLASYQGILRAVRRSHGIDCSFWNESYLTGSRDERKKSLLSKFGMDEGVTFMFIGRFDGRGQK  
GVDVLLKAIILSSKKEFQEMRFIIIGKDPELEGWARSLEEKHGNVKVITEMLSRELVRELYGSVD  
FVIIPSYFEPFGLVALEAMCLGAIPIASAVGGLRDIITNETGILVKAGDPGELANAILKALELSRSDLS  
KFRENCKKRAMSFSGTQKSRKDQIQRLVSTVVIPLACFLPYHVLLLVRSVWEASCDFAKGVFNA  
YHFSLLLTSFNCVANPVLVCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSGA  
QGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA ENLYFQG HHHHHHHHHH

GPR68 C42: HA-FLAG-GPR68(BRIL\_GS-PreC-9-216, PGS, 217-365)-TEV-His10  
(Mutation: N86<sup>23.49</sup>Q, T293<sup>8.50</sup>F)

MKTIIALSYIFCLVFA DYKDDDDK ADLEDN WETLNDNLKVIEKADNAAQVKDALTKMRAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGGSEVLFQGPSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCN  
LTVADLFYICSLPFWLQYVLQHDQ WSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPRF  
HQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFL  
FPICLLASYQGILRAVRRSHGIDCSFWNESYLTGSRDERKKSLLSKFGMDEGVTFMFIGRFDGRGQK  
GVDVLLKAIILSSKKEFQEMRFIIIGKDPELEGWARSLEEKHGNVKVITEMLSRELVRELYGSVD  
FVIIPSYFEPFGLVALEAMCLGAIPIASAVGGLRDIITNETGILVKAGDPGELANAILKALELSRSDLS  
KFRENCKKRAMSFSGTQKSRKDQIQRLVSTVVIPLACFLPYHVLLLVRSVWEASCDFAKGVFNA  
YHFSLLLTSFNCVADPVLVCFVSETFHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSGA  
QGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA ENLYFQG HHHHHHHHHH

GPR68 C43: HA-FLAG-GPR68(15-214, BRIL, 219-347)-PreC-His10

MKTIIALSYIFCLVFA DYKDDDDK IDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELG  
VYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAV  
AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYR  
FLVGFLFPICLLASYQGILRAVRRADLEDN WETLNDNLKVIEKADNAAQVKDALTKMRAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LQKSRKDQIQRLVSTVVIPLACFLPYHVLLLVRSVWEASCDFAKGVFNAYHFSLLLTSFNCVADP  
VLVCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSGAQGEEPELLTKLHPA LE  
VLFQGP HHHHHHHHHH

GPR68 C44: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-365)-TEV-His10  
(Mutation: L63<sup>2.46</sup>A)

MKTIIALSYIFCLVFA DYKDDDDK ADLEDN WETLNDNLKVIEKADNAAQVKDALTKMRAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGGSEVLFQGPMGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN  
ELGVYLCNATVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRY  
LAVAHPRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVSTVVIPLACFLPYHVLLLVRSV  
WEASCDFAKGVFNAYHFSLLLTSFNCVADPVLVCFVSETTHRDLARLRGACLAFLTCRTGRARE  
AYPLGAPEASGKSGAQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA ENLYFQG HHHHHHHHHH

GPR68 C45: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-365)-TEV-His10  
(Mutation: R119<sup>3.50</sup>A)

MKTIIALSYIFCLVFA**DYKDDDDK**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY  
LGG**S****LEVLFQGP**MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGGFPANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVD**A**Y  
LAVAHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVSV  
WEASCDFAKGVFNAYHFSLLLTSFNCAVADPVLYCFVSETTHRDARLRGACLAFLTCRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPAFQTPNSPGSGGFPTGRLA**ENLYFQGH**HHHHHHHHHH

GPR68 C46: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10  
(Mutation: C304A, C310A)

MKTIIALSYIFCLVFA**DYKDDDDK**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY  
LGG**S****LEVLFQGP**MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGGFPANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVD**R**Y  
LAVAHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVSV  
WEASCDFAKGVFNAYHFSLLLTSFNCAVADPVLYCFVSETTHRDARLRGA**A**LAFL**A**SRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPA**ENLYFQGH**HHHHHHHHHH

GPR68 C47: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10  
(Mutation: G46<sup>1.57</sup>A)

MKTIIALSYIFCLVFA**DYKDDDDK**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY  
LGG**S****LEVLFQGP**MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGGFPANCLSLY**F**AYLQIKARN  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVD**R**Y  
LAVAHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVSV  
WEASCDFAKGVFNAYHFSLLLTSFNCAVADPVLYCFVSETTHRDARLRGACLAFLTCRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPA**ENLYFQGH**HHHHHHHHHH

GPR68 C48: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10  
(Mutation: L60<sup>2.43</sup>I)

MKTIIALSYIFCLVFA**DYKDDDDK**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY  
LGG**S****LEVLFQGP**MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGGFPANCLSLYFGYLQIKARN  
ELGVY**I**CNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVD**R**YL  
AVAHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN  
YRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVSVW  
EASCDFAKGVFNAYHFSLLLTSFNCAVADPVLYCFVSETTHRDARLRGACLAFLTCRTGRARE**A**  
PLGAPEASGKSGAQGEPELLTKLHPA**ENLYFQGH**HHHHHHHHHH

GPR68 C49: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10  
(Mutation: L112<sup>3.43</sup>A)

MKTIIALSYIFCLVFA**DYKDDDDK**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY  
LGG**S****LEVLFQGP**MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGGFPANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGF**A**CCISVD**R**Y

LVAHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQLVLSTVVIFLACFLPYHVLLLVRV  
WEASCDFAKGVFNAYHFSLLTSFNCAVADPVLVCFVSETTHRDLARLRGACLAFLTCRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPA **ENLYFQG** **HHHHHHHHHH**

*GPR68 C50: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10*  
(Mutation: I115<sup>3.46</sup>A)

**MKTIALSIFCLVFA** **DYKDDDDK** **ADLEDNWETLNDNLK** **VIEKADNAAQVKDAL** **TKMRAAALDA**  
**QKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY**  
**LGGS** **LEVLFQGP** **MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN**  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFCC **A**SVDRY  
LVAHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQLVLSTVVIFLACFLPYHVLLLVRV  
WEASCDFAKGVFNAYHFSLLTSFNCAVADPVLVCFVSETTHRDLARLRGACLAFLTCRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPA **ENLYFQG** **HHHHHHHHHH**

*GPR68 C51: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10*  
(Mutation: K223<sup>6.30</sup>E)

**MKTIALSIFCLVFA** **DYKDDDDK** **ADLEDNWETLNDNLK** **VIEKADNAAQVKDAL** **TKMRAAALDA**  
**QKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY**  
**LGGS** **LEVLFQGP** **MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN**  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFCCISVDRY  
LVAHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSR **E**DQIQLVLSTVVIFLACFLPYHVLLLVRV  
WEASCDFAKGVFNAYHFSLLTSFNCAVADPVLVCFVSETTHRDLARLRGACLAFLTCRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPA **ENLYFQG** **HHHHHHHHHH**

*GPR68 C52: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10*  
(Mutation: T64<sup>2.47</sup>A)

**MKTIALSIFCLVFA** **DYKDDDDK** **ADLEDNWETLNDNLK** **VIEKADNAAQVKDAL** **TKMRAAALDA**  
**QKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY**  
**LGGS** **LEVLFQGP** **MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN**  
ELGVYLCNLT **A**VADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFCCISVDRY  
LVAHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQLVLSTVVIFLACFLPYHVLLLVRV  
WEASCDFAKGVFNAYHFSLLTSFNCAVADPVLVCFVSETTHRDLARLRGACLAFLTCRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPA **ENLYFQG** **HHHHHHHHHH**

*GPR68 C53: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10*  
(Mutation: V109<sup>3.40</sup>A)

**MKTIALSIFCLVFA** **DYKDDDDK** **ADLEDNWETLNDNLK** **VIEKADNAAQVKDAL** **TKMRAAALDA**  
**QKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYIQKY**  
**LGGS** **LEVLFQGP** **MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN**  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYIS **A**GFLCCISVDRY  
LVAHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQLVLSTVVIFLACFLPYHVLLLVRV  
WEASCDFAKGVFNAYHFSLLTSFNCAVADPVLVCFVSETTHRDLARLRGACLAFLTCRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPA **ENLYFQG** **HHHHHHHHHH**

GPR68 C54: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10  
(Mutation: E103<sup>3.34</sup>A)

MKTIIALSYIFLVFA DYKDDDDK ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGGSEVLFQGP MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLY ANIYISVGFLCCISVDRY  
LAVAHPRFRHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVLSVVIPLACFLPYHVLLLVRVSV  
WEASCDFAKGVFNAYHFSLLTSFNVCVADPVLVCFVSETTHRDLARLRGACLAFLTCRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPA ENLYFQGH HHHHHHHHHH

GPR68 C55: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10  
(Mutation: E149<sup>4.53</sup>A)

MKTIIALSYIFCLVFA DYKDDDDK ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGGSEVLFQGP MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRY  
LAVAHPRFRHQFRTLKAAVGVSVVIWAK ALLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVLSVVIPLACFLPYHVLLLVRVSV  
WEASCDFAKGVFNAYHFSLLTSFNVCVADPVLVCFVSETTHRDLARLRGACLAFLTCRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPA ENLYFQGH HHHHHHHHHH

GPR68 C56: HA-FLAG-GPR68(BRIL\_GGS-PreC-9-347)-TEV-His10

MKTIIALSYIFCLVFA DYKDDDDK ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGGSEVLFQGP SSMSCCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCN  
LTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRYLAVAHPRFR  
HQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFL  
FPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVLSVVIPLACFLPYHVLLLVRVSVWEASCDFAK  
GVFNAYHFSLLTSFNVCVADPVLVCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEAS  
GKSGAQGEPELLTKLHPA ENLYFQGH HHHHHHHHHH

GPR68 C57: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10  
(Mutation: T233<sup>6.40</sup>V)

MKTIIALSYIFCLVFA DYKDDDDK ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGGSEVLFQGP MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRY  
LAVAHPRFRHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVLS VVVIPLACFLPYHVLLLVRVSV  
WEASCDFAKGVFNAYHFSLLTSFNVCVADPVLVCFVSETTHRDLARLRGACLAFLTCRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPA ENLYFQGH HHHHHHHHHH

GPR68 C58: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10  
(Mutation: V230<sup>6.37</sup>A)

MKTIIALSYIFCLVFA DYKDDDDK ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGGSEVLFQGP MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRY  
LAVAHPRFRHQFRTLKAAVGVSVVIWAKELLTSIYFLMHVEEVIEDENQHRVCFEHYPIQAWQRAIN

YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLALSTVVIFLACFLPYHVLLLVRVSV  
WEASCDFAKGVFNAYHFSLLTSFNVCVADPVLVYCFVSETTHRDLARLRGACLAFLTCRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPAENLYFQGHHHHHHHHHH

*GPR68 C59: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10*  
(Mutation: V234<sup>6.41A</sup>)

MKTIALSIFCLVFA DYKDDDDKADLEDNWTLNNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGGSEVLFQGMGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLLCCISVDRY  
LAVAHPRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTAVIFLACFLPYHVLLLVRVSV  
WEASCDFAKGVFNAYHFSLLTSFNVCVADPVLVYCFVSETTHRDLARLRGACLAFLTCRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPAENLYFQGHHHHHHHHHH

*GPR68 C60: HA-FLAG-GPR68(1-214, QL\_BRIL, 220-347)-PreC-His10*

MKTIALSIFCLVFA DYKDDDDKIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELG  
VYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLLCCISVDRYLAV  
AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRRAINYYR  
FLVGFLFPICLLASYQGILRAVRRQLADLEDNWTLNNDNLKVIEKADNAAQVKDALTKMRAAAL  
DAQKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQ  
KYLKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVSVWEASCDFAKGVFNAYHFSLLTSFNVCVAD  
PVLVYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSGAQGEPELLTKLHPALE  
EVLFQGHHHHHHHHHH

*GPR68 C61: HA-FLAG-GPR68(15-214, QL\_BRIL, 221-347)-PreC-His10*

MKTIALSIFCLVFA DYKDDDDKIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELG  
VYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLLCCISVDRYLAV  
AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRRAINYYR  
FLVGFLFPICLLASYQGILRAVRRQLADLEDNWTLNNDNLKVIEKADNAAQVKDALTKMRAAAL  
DAQKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQ  
KYLKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVSVWEASCDFAKGVFNAYHFSLLTSFNVCVADP  
VLVYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSGAQGEPELLTKLHPALE  
VLFQGHHHHHHHHHH

*GPR68 C62: HA-FLAG-GPR68(1-211, ARRQL\_BRIL\_ERARSTL, 221-365)-PreC-His10*  
(Mutation S108<sup>3.39R</sup>)

BRIL fusion is connected to TM5 and TM6 by two short peptides (ARRQL and ERARSTL) derived from TM5 and TM6 of A<sub>2A</sub> adenosine receptor.

MKTIALSIFCLVFA DYKDDDDKMGNTADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSL  
LYFGYLQIKARNELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYIR  
VGFLLCCISVDRYLAVAHPRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE  
HYPIQAWQRRAINYYRFLVGFLFPICLLASYQGILRAARRQLADLEDNWTLNNDNLKVIEKADNAA  
QVKDALTKMRAAALDAQKATPPKLEDKSPDPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQ  
AAAEQLKTRNAYIQKYLERARSTLSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVSVWEASCDFA  
KGVFNAYHFSLLTSFNVCVADPVLVYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEA  
SGKSGAQGEPELLTKLHPAFQTPNSPGSGGFPTGRLEVLFQGHHHHHHHHHH

*GPR68 C63: HA-FLAG-GPR68(1-211, ARRQL\_BRIL\_ERARSTL, 221-319)-PreC-His10*

BRIL fusion is connected to TM5 and TM6 by two short peptides (ARRQL and ERARSTL) derived from TM5 and TM6 of A<sub>2A</sub> adenosine receptor.

MKTIALSIFCLVFA DYKDDDDK MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGF PANCLS  
LYFGYLQIKARNELGVYLCNLT VADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYIS  
VGFLCCISVDRYLAV AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE  
HYPIQAWQRAINYYRFLVGF LFPICLLLAS YQGILRA ARRQLADLEDN WETLNDNLK VIEKADNAA  
QVKDALTKMRAAALDA QKATPPKLEDKSPD SPEMKDFRHGFDILVG QIDDALKLANEGKVKEAQ  
AAAEQLKTRNAYIQYLERARSTL SRKDQIQRLV LSTVVIFLACFLPYHVLLLR SVWEASCDFA  
KGVFNAYHFSLLLSFN CVADPVLYCFVSE TTHRDLARLRGACLAFLTC SRTGRAREA LEVLFQGP  
HHHHHHHHHH

*GPR68 C64: HA-FLAG-GPR68(15-214, BRIL, 219-319)-PreC-His10*

MKTIALSIFCLVFA DYKDDDDK IDHTIHQTLAPVVYVTVLVVGF PANCLS LYFGYLQIKARNELG  
VYLCNLT VADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAV  
AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE HYPIQAWQRAINYYR  
FLVGF LFPICLLLAS YQGILRA VRRADLEDN WETLNDNLK VIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPD SPEMKDFRHGFDILVG QIDDALKLANEGKVKEAQAAAEQLKTRNAYIQY  
LQKSRKDQIQRLV LSTVVIFLACFLPYHVLLLR SVWEASCDFAKGVFNAYHFSLLLSFN CVADP  
VLYCFVSE TTHRDLARLRGACLAFLTC SRTGRAREA LEVLFQGP HHHHHHHHHH

*GPR68 C65: HA-FLAG-GPR68[1-214, BLEK (19-377), 221-347]-PreC-His10*

MKTIALSIFCLVFA DYKDDDDK MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGF PANCLS  
LYFGYLQIKARNELGVYLCNLT VADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYIS  
VGFLCCISVDRYLAV AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE  
HYPIQAWQRAINYYRFLVGF LFPICLLLAS YQGILRA VRRAAPQQINDIVHRTITPLIEQQKIPGMAV  
AVIYQGKPPYYFTWGY ADIAKKQPVTQQT LFELGSVSKTFTGVLGGDAIARGEIKLSDPTTKYWPEL  
TAKQWNGITLLHLATYTAGGLPLQVPDEVKSSDLLRFYQNWQPAWAPGTQRLYANSSIGLFGAL  
AVKPSGLSFEQAMQTRVFQPLKLNHTWINVPPAEKNYA WGYREGKAVHVSFGALDAEAYGVKS  
TIEDMARVWQSNLPLDINEKTLQQGIQAQSR YWQTGDMYQGLGWEMLDWPVNPDSIINGS DN  
KIALAARPVKAITPPTPAVRASWVHKTGATGGFGSYVAFIPEKELGIVMLANKNYPNPARVDAAW  
QILNALQSRKDQIQRLV LSTVVIFLACFLPYHVLLLR SVWEASCDFAKGVFNAYHFSLLLSFN CV  
ADPVLYCFVSE TTHRDLARLRGACLAFLTC SRTGRAREAYPLGAPEASGKSGAQGEPELLTKLHP  
A LEVLFQGP HHHHHHHHHH

*GPR68 C66: HA-FLAG-GPR68[BRIL\_GGS-PreC-9-214, BLEK (19-377), 221-347]-TEV-His10*

MKTIALSIFCLVFA DYKDDDDK ADLEDN WETLNDNLK VIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPD SPEMKDFRHGFDILVG QIDDALKLANEGKVKEAQAAAEQLKTRNAYIQY  
LGGSL EVLFQGP SSMSC TIDHTIHQTLAPVVYVTVLVVGF PANCLS LYFGYLQIKARNELGVYLCN  
LTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAV AHPFRF  
HQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE HYPIQAWQRAINYYRFLVGF  
LFPICLLLAS YQGILRA VRRAAPQQINDIVHRTITPLIEQQKIPGMAVAVIYQGKPPYYFTWGYADIAK  
KQPVTQQTLFELGSVSKTFTGVLGGDAIARGEIKLSDPTTKYWPELTAKQWNGITLLHLATYTAGG  
LPLQVPDEVKSSDLLRFYQNWQPAWAPGTQRLYANSSIGLFGALAVKPSGLSFEQAMQTRVFQPL  
KLNHTWINVPPAEKNYA WGYREGKAVHVSFGALDAEAYGVKSTIEDMARVWQSNLPLDINE  
KTLQQGIQAQSR YWQTGDMYQGLGWEMLDWPVNPDSIINGS DNKIALAARPVKAITPPTPAVRA  
SWVHKTGATGGFGSYVAFIPEKELGIVMLANKNYPNPARVDAAWQILNALQSRKDQIQRLV LSTV  
VIFLACFLPYHVLLLR SVWEASCDFAKGVFNAYHFSLLLSFN CVADPVLYCFVSE TTHRDLARL  
RGACLAFLTC SRTGRAREAYPLGAPEASGKSGAQGEPELLTKLHPA ENLYFQG HHHHHHHHHH

GPR68 C67: HA-FLAG-GPR68(15-214, CalB(16-170), 219-319)-CalA(14-371)-PreC-His10

MKTIIALSYIFCLVFA DYKDDDDK IDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELG  
VYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRYLAV  
AHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYR  
FLVGF LFPICLLASYQGILRAVRR DADEIKRLGKRFKL DLNSGSLSV EEFMSLP ELQQNPLVQR  
VIDIFDTGNGEVDFKEFIEGVSQFSVKGDKEQKLRFAFRYDMDKDG YISNGELFQVLKMMVGN  
NLKDTQLQQIVDKTIINADKDG DGRISFE EFC AVVGGLDIHKM VVDVQKSRKDQIQRLVLSTVVI  
FLACFLPYHVLLLVR SVWEASCDFAKGVFNAYHFSLLTSFN CVADPVL YCFVSETTHRDLARLR  
GACLAFLTCSRTGRAREA TDRVVKAVPFP SHRLTAKEVFDNDGKPRVDILKAHLMKEGRLEESV  
ALRIITEGASILRQEKNLDDIDAPVTVCGDIHGQFFDLMKLFEVGGSPANTRYLFLGDYVDRGYFSI  
ECVLYLWALKILYPKTLFLLRGNHECRHLTEYFTFKQECKIKYSERVYDACMDAFDCLPLAALMN  
QQFLCVHGGLSPEINTLDDIRKLRDFKEPPAYGPMCDILWSDPLEDFGNEKTQEHTHTNTVRGCSY  
FYSYPAVCDFLQHNLLSILRAHEAQDAGYRMYRKSQTTGFPSLITIFSAPNYLDVYNNKA AVLKY  
ENNVMNIRQFNCSPHYWLPNFMVFTWSLPFVGEKVTEMLVNVLNI LEVLFQGP HHHHHHHHHH  
H

GPR68 C68: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10  
(Mutation F76<sup>2.59</sup>C)

MKTIIALSYIFCLVFA DYKDDDDK ADLEDN WETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGGSL E V L F Q G P MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSLP CWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRY  
LAVAHPRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGF LFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVI FLACFLPYHVLLLVR SV  
WEASCDFAKGVFNAYHFSLLTSFN CVADPVL YCFVSETTHRDLARLRGACLAFLTCSRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPA ENLYFQGH HHHHHHHHHH

GPR68 C69: HA-FLAG-GPR68(BRIL\_GGS-PreC-9-319)-TEV-GFP-His10  
(Mutation N86<sup>23.49</sup>Q)

MKTIIALSYIFCLVFA DYKDDDDK ADLEDN WETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGGSL E V L F Q G P SSMSC TIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCN  
LTVADLFYICSLPFWLQYVLQHD Q WSHGDLSQVCGILLYENIYISVGFLCCISVDRYLAVAHPRF  
HQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGF  
LFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVI FLACFLPYHVLLLVR SVWEASCDFAK  
GVFNAYHFSLLTSFN CVADPVL YCFVSETTHRDLARLRGACLAFLTCSRTGRAREA ENLYFQGH  
GEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLT LKFICTTGKLPVPWP TLVTTLYGVO  
CFSRYPDHMKRHDFK SAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDG  
NILGRHKLEYNYN SHNVYITADKQKNGIKANFKIRHNIEDG SVQLADHYQONTPIGDGPVLLPDNH  
LSTQSALS KDPNEKRDMVLL E FVTAAGITHGMDELYK HHHHHHHHHH

GPR68 C70: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10  
(Mutation C113L, C114A)

MKTIIALSYIFCLVFA DYKDDDDK ADLEDN WETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSPEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGGSL E V L F Q G P MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFL LAISVDRY  
LAVAHPRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGF LFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVI FLACFLPYHVLLLVR SV  
WEASCDFAKGVFNAYHFSLLTSFN CVADPVL YCFVSETTHRDLARLRGACLAFLTCSRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPA ENLYFQGH HHHHHHHHHH

*GPR68 C71: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10  
(Mutation S108<sup>3.39</sup>K)*

MKTIALSIFCLVFA**DYKDDDDK**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGG**SLEVLFGP**MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYIKVGFLLCCISVDYR  
LVAHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVSTVVIFLACFLPYHVLLLVRSV  
WEASCDFAKGVFNAYHFSLLTSFNCAVADPVLVCFVSETTHRDLARLRGACLAFLTCRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPA**ENLYFQG**HHHHHHHHHH

*GPR68 C72: HA-FLAG-GPR68(BRIL\_GGS-PreC-1-347)-TEV-His10  
(Mutation L74<sup>2.57</sup>P)*

MKTIALSIFCLVFA**DYKDDDDK**ADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDA  
QKATPPKLEDKSPDSPMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKY  
LGG**SLEVLFGP**MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN  
ELGVYLCNLTVADLFYICSPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLLCCISVDYR  
LVAHPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAIN  
YYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQRLVSTVVIFLACFLPYHVLLLVRSV  
WEASCDFAKGVFNAYHFSLLTSFNCAVADPVLVCFVSETTHRDLARLRGACLAFLTCRTGRARE  
AYPLGAPEASGKSGAQGEPELLTKLHPA**ENLYFQG**HHHHHHHHHH

*GPR68 C73: HA-FLAG-GPR68(9-50, BRIL, 55-319)-TEV-His10*

MKTIALSIFCLVFA**DYKDDDDK**SSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIA  
DLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDAQKATPPKLEDKSPDSPMKDFRHG  
FDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKYLELGVYLCNLTVADLFYICSLPFWL  
QYVLQHDNWSHGDLSQVCGILLYENIYISVGFLLCCISVDYR LVAHPFRFHQFRTLKAAVGVSVV  
IWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPICLLASYQGILRA  
VRRSHGTQKSRKDQIQRLVSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAYHFSLLTSF  
NCVADPVLVCFVSETTHRDLARLRGACLAFLTCRTGRAREA**ENLYFQG**HHHHHHHHHH

*GPR68 C74: HA-FLAG-GPR68(9-122, BRIL, 135-319)-TEV-His10*

MKTIALSIFCLVFA**DYKDDDDK**SSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIK  
ARNELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLLCCISV  
DRYLAADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDAQKATPPKLEDKSPDSPM  
KDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTRNAYIQKYLKAAVGVSVVIWAKEL  
LTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPICLLASYQGILRAVRRSHG  
TQKSRKDQIQRLVSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAYHFSLLTSFNCAVADP  
VLVCFVSETTHRDLARLRGACLAFLTCRTGRAREA**ENLYFQG**HHHHHHHHHH

*GPR68 C99: HA-FLAG-Zebrafish\_GPR68(GFP\_GGS-TEV-9-209, BRIL, 219-309)-PreC-  
TST (Mutation: N9Q)*

MKTIALSIFCLVFA**DYKDDDDK**KGEELFTGVVPIVLDGDVNGHKFSVSGEGEGDATYGKLT  
KFICTTGKLPVPWPTLVTTLYGVQCFSRYPDHMKRHDFKSAPEGYVQERTISFKDDGNYKTR  
AEVKFEGDTLVNRIELKIDFKEDGNILGHKLEYNYNNSHNVIYITADKQKNGIKANFKIRHNIEDGS  
VOLADHYQQNTPIGDGPVLLPDNHVLSQSALS KDPNEKRDMVLEFVTAAGITHGMDELYKG  
G**SENLYFQG**QCTINHDIHQYLFVAYLLVLLVGLPANAYSLYHAWLQKARNELGVYLLNLTISDL  
LYLGSPLWLQYIFQGDNWSGSEWLCQCGFLLYENIYVSIGFLCCISIDRYLAVVYPRFSAFRTV  
RAATLVSTVWLKELAVGVVFFLHKELSRDKHNQSVCFEHYPMKTWEYQINYYRFYIGFLPLGI  
LSVSYFRVLRVADLEDNWETLNDNLKVIEKADNAAQVKDALTKMRAAALDAQKATPPKLEDK



TLKAAVGVSVVIWAK<sup>Q</sup>LLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPIC  
LLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVWEASCDFAKGVF  
NAYHFSLLLTSFNCAVPVLYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKS  
GAQGEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C79: GPR68(1-365) Mutation(H20<sup>1.31</sup>F)*  
*(HEK293 expression for cellular assays)*

MGNITADNSSMCTIDHTI<sup>F</sup>QTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVA  
DLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
TLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPICL  
LLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVWEASCDFAKGVFN  
AYHFSLLLTSFNCAVPVLYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSG  
AQGEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C80: GPR68(1-365) Mutation(H269<sup>7.36</sup>F)*  
*(HEK293 expression for cellular assays)*

MGNITADNSSMCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVA  
DLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
TLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPICL  
LLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVWEASCDFAKGVFN  
AY<sup>F</sup>FSLLLTSFNCAVPVLYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSG  
AQGEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C81: GPR68(1-365) Mutation(E103<sup>3.34</sup>Q)*  
*(HEK293 expression for cellular assays)*

MGNITADNSSMCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVA  
DLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLY<sup>Q</sup>NIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
RTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPIC  
LLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVWEASCDFAKGVF  
NAY<sup>F</sup>FSLLLTSFNCAVPVLYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKS  
GAQGEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C82: GPR68(1-365) Mutation(E103<sup>3.34</sup>A)*  
*(HEK293 expression for cellular assays)*

MGNITADNSSMCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVA  
DLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLY<sup>A</sup>NIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
TLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPICL  
LLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVWEASCDFAKGVFN  
AYHFSLLLTSFNCAVPVLYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSG  
AQGEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C83: GPR68(1-365) Mutation(K148<sup>4.52</sup>N)*  
*(HEK293 expression for cellular assays)*

MGNITADNSSMCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVA  
DLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
TLKAAVGVSVVIWA<sup>N</sup>ELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPICL  
LLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVWEASCDFAKGVFN  
AYHFSLLLTSFNCAVPVLYCFVSETTHRDLARLRGACLAFLTCRTGRAREAYPLGAPEASGKSG  
AQGEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C84: GPR68(1-365) Mutation(R189<sup>5.42</sup>N)*  
(HEK293 expression for cellular assays)

MGNITADNSSMCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVA  
DLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
TLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYN<sup>N</sup>FLVGF<sup>L</sup>FPICL  
LLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVWEASCDFAKGVFN  
AYHFSLLLTSFNCVADPVLVYCFVSETTHRDARLRGACLAFLTCSRTGRAREAYPLGAPEASGKSG  
AQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C85: GPR68(1-365) Mutation(H245<sup>6.52</sup>N)*  
(HEK293 expression for cellular assays)

MGNITADNSSMCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVA  
DLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
TLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGF<sup>L</sup>FPICL  
LLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPY<sup>D</sup>VLLLVRVWEASCDFAKGVFN  
AYHFSLLLTSFNCVADPVLVYCFVSETTHRDARLRGACLAFLTCSRTGRAREAYPLGAPEASGKSG  
AQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C86: GPR68(1-365) Mutation(N186<sup>6.52</sup>N)*  
(HEK293 expression for cellular assays)

MGNITADNSSMCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVA  
DLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
TLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAI<sup>A</sup>YYRFLVGF<sup>L</sup>FPICL  
LLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVWEASCDFAKGVFN  
AYHFSLLLTSFNCVADPVLVYCFVSETTHRDARLRGACLAFLTCSRTGRAREAYPLGAPEASGKSG  
AQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C87: GPR68(1-365) Mutation(S271<sup>7.38</sup>V)*  
(HEK293 expression for cellular assays)

MGNITADNSSMCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVA  
DLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
TLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGF<sup>L</sup>FPICL  
LLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVWEASCDFAKGVFN  
AYHF<sup>V</sup>LLLTSFNCVADPVLVYCFVSETTHRDARLRGACLAFLTCSRTGRAREAYPLGAPEASGKSG  
AQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C88: GPR68(1-365) Mutation(E174<sup>45.52</sup>Q)*  
(HEK293 expression for cellular assays)

MGNITADNSSMCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVA  
DLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
TLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCF<sup>Q</sup>HYPYPIQAWQRAINYYRFLVGF<sup>L</sup>FPIC  
LLLASYQGILRAVRRSHGTQKSRKDQIQRLVLSTVVIFLACFLPYHVLLLVRVWEASCDFAKGVFN  
NAYHFSLLLTSFNCVADPVLVYCFVSETTHRDARLRGACLAFLTCSRTGRAREAYPLGAPEASGKS  
GAQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C89: GPR68(1-365) Mutation(D85<sup>ECL1</sup>N, D91<sup>3.22</sup>N)  
(HEK293 expression for cellular assays)*

MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVA  
DLFYICSLPFWLQYVVLQHNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
TLKAAVGVSVVIWAKELLTSIYFLMH<sup>N</sup>EEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPICL  
LLASYQGILRAVRRSHGTQKSRKDQIQLVLSTVVIFLACFLPYHVLLLVRVWEASCDFAKGVFN  
AYHFSLLLTSFNCVADPVLVYCFVSETTHRDLARLRGACLAFLTCSRTGRAREAYPLGAPEASGKSG  
AQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C90: GPR68(1-365) Mutation(E160<sup>4.64</sup>Q, E161<sup>ECL2</sup>Q, E164<sup>ECL2</sup>Q, D165<sup>ECL2</sup>N,  
E166<sup>ECL2</sup>Q) (HEK293 expression for cellular assays)*

MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVA  
DLFYICSLPFWLQYVVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
TLKAAVGVSVVIWAKELLTSIYFLMH<sup>QQVIQ</sup>QNRVCFEHYPIQAWQRAINYYRFLVGFLFPICL  
LLLASYQGILRAVRRSHGTQKSRKDQIQLVLSTVVIFLACFLPYHVLLLVRVWEASCDFAKGVFN  
NAYHFSLLLTSFNCVADPVLVYCFVSETTHRDLARLRGACLAFLTCSRTGRAREAYPLGAPEASGKS  
GAQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C91: GPR68(1-365) Mutation(E255<sup>ECL3</sup>Q, D259<sup>ECL3</sup>N)  
(HEK293 expression for cellular assays)*

MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVA  
DLFYICSLPFWLQYVVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
TLKAAVGVSVVIWAKELLTSIYFLMH<sup>QQVIQ</sup>QNRVCFEHYPIQAWQRAINYYRFLVGFLFPICL  
LLASYQGILRAVRRSHGTQKSRKDQIQLVLSTVVIFLACFLPYHVLLLVRVW<sup>QASCN</sup>FAKGVFN  
AYHFSLLLTSFNCVADPVLVYCFVSETTHRDLARLRGACLAFLTCSRTGRAREAYPLGAPEASGKSG  
AQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C92: GPR68(1-365) Mutation(E55<sup>2.38</sup>A)  
(HEK293 expression for cellular assays)*

MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARN<sup>A</sup>LGVYLCNLTVA  
ADLFYICSLPFWLQYVVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
RTLKAAVGVSVVIWAKELLTSIYFLMH<sup>EEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPICL</sup>  
LLASYQGILRAVRRSHGTQKSRKDQIQLVLSTVVIFLACFLPYHVLLLVRVWEASCDFAKGVFN  
NAYHFSLLLTSFNCVADPVLVYCFVSETTHRDLARLRGACLAFLTCSRTGRAREAYPLGAPEASGKS  
GAQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C93: GPR68(1-365) Mutation(T233<sup>6.40</sup>V)  
(HEK293 expression for cellular assays)*

MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVA  
DLFYICSLPFWLQYVVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
TLKAAVGVSVVIWAKELLTSIYFLMH<sup>EEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPICL</sup>  
LLASYQGILRAVRRSHGTQKSRKDQIQLVLS<sup>V</sup>VVIFLACFLPYHVLLLVRVWEASCDFAKGVFN  
AYHFSLLLTSFNCVADPVLVYCFVSETTHRDLARLRGACLAFLTCSRTGRAREAYPLGAPEASGKSG  
AQGEEPELLTKLHPAFQTPNSPGSGGFPTGRLA

*GPR68 C94: GPR68(Met-9-365-V5)*  
*(HEK293 expression for cellular assays)*

MSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVADLFYICS  
LPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFRTLKAAV  
GVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPICLLASYQ  
GILRAVRRSHGTQKSRKDQIQLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAYHFSL  
LLTSFNVCVADPVLVCFVSETTHRDLARLRGACLAFLTCSRTGRAREAYPLGAPEASGKSGAQGEEP  
ELLTKLHPAFQTPNSPGSGGFPTGRLA **GKPIPNNLLGLDST**

*GPR68 C95: GPR68(Met-15-365-V5)*  
*(HEK293 expression for cellular assays)*

MIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVADLFYICSLPFWLQ  
YVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFRTLKAAVGVSVVI  
WAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPICLLASYQGILRAV  
RRSHGTQKSRKDQIQLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFNAYHFSLLTSFN  
CVADPVLVCFVSETTHRDLARLRGACLAFLTCSRTGRAREAYPLGAPEASGKSGAQGEEPPELLTKL  
HPAFQTPNSPGSGGFPTGRLA **GKPIPNNLLGLDST**

*GPR68 C96: GPR68(Met-V5-1-319)*  
*(HEK293 expression for cellular assays)*

M **GKPIPNNLLGLDST**GNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKAR  
NELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDR  
YLAVAHPPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAI  
NYYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQLVLSTVVIFLACFLPYHVLLLVRS  
VWEASCDFAKGVFNAYHFSLLTSFNVCVADPVLVCFVSETTHRDLARLRGACLAFLTCSRTGRAR  
EA

*GPR68 C97: GPR68(Met-V5-1-305)*  
*(HEK293 expression for cellular assays)*

M **GKPIPNNLLGLDST**GNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKAR  
NELGVYLCNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDR  
YLAVAHPPFRFHQFRTLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAI  
NYYRFLVGFLFPICLLASYQGILRAVRRSHGTQKSRKDQIQLVLSTVVIFLACFLPYHVLLLVRS  
VWEASCDFAKGVFNAYHFSLLTSFNVCVADPVLVCFVSETTHRDLARLRGACL

*GPR68 C98: GPR68(1-365-V5)*  
*(HEK293 expression for cellular assays)*

MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFANCLSLYFGYLQIKARNELGVYLCNLTVA  
DLFYICSLPFWLQYVLQHDNWSHGDLSQVCGILLYENIYISVGFLCCISVDRYLAVAHPPFRFHQFR  
TLKAAVGVSVVIWAKELLTSIYFLMHEEVIEDENQHRVCFEHYPIQAWQRAINYYRFLVGFLFPICL  
LLASYQGILRAVRRSHGTQKSRKDQIQLVLSTVVIFLACFLPYHVLLLVRSVWEASCDFAKGVFN  
AYHFSLLTSFNVCVADPVLVCFVSETTHRDLARLRGACLAFLTCSRTGRAREAYPLGAPEASGKSG  
AQGEEPPELLTKLHPAFQTPNSPGSGGFPTGRLA **GKPIPNNLLGLDST**

## Appendix 2: Supplementary figures

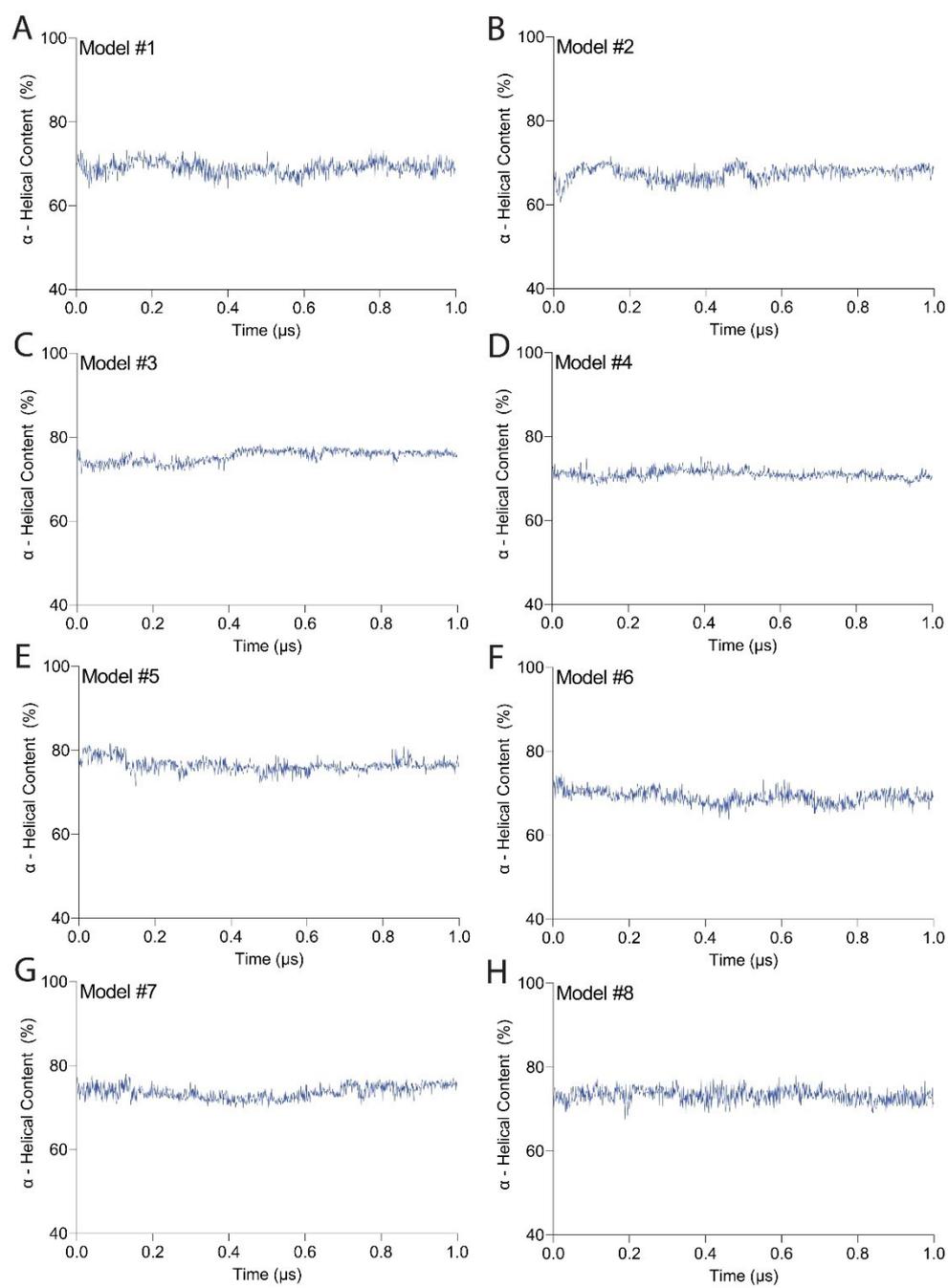


Figure S1: The time series of percentage  $\alpha$ -helical content for simulations #1 - #8 were computed using Stride <sup>[396]</sup>.

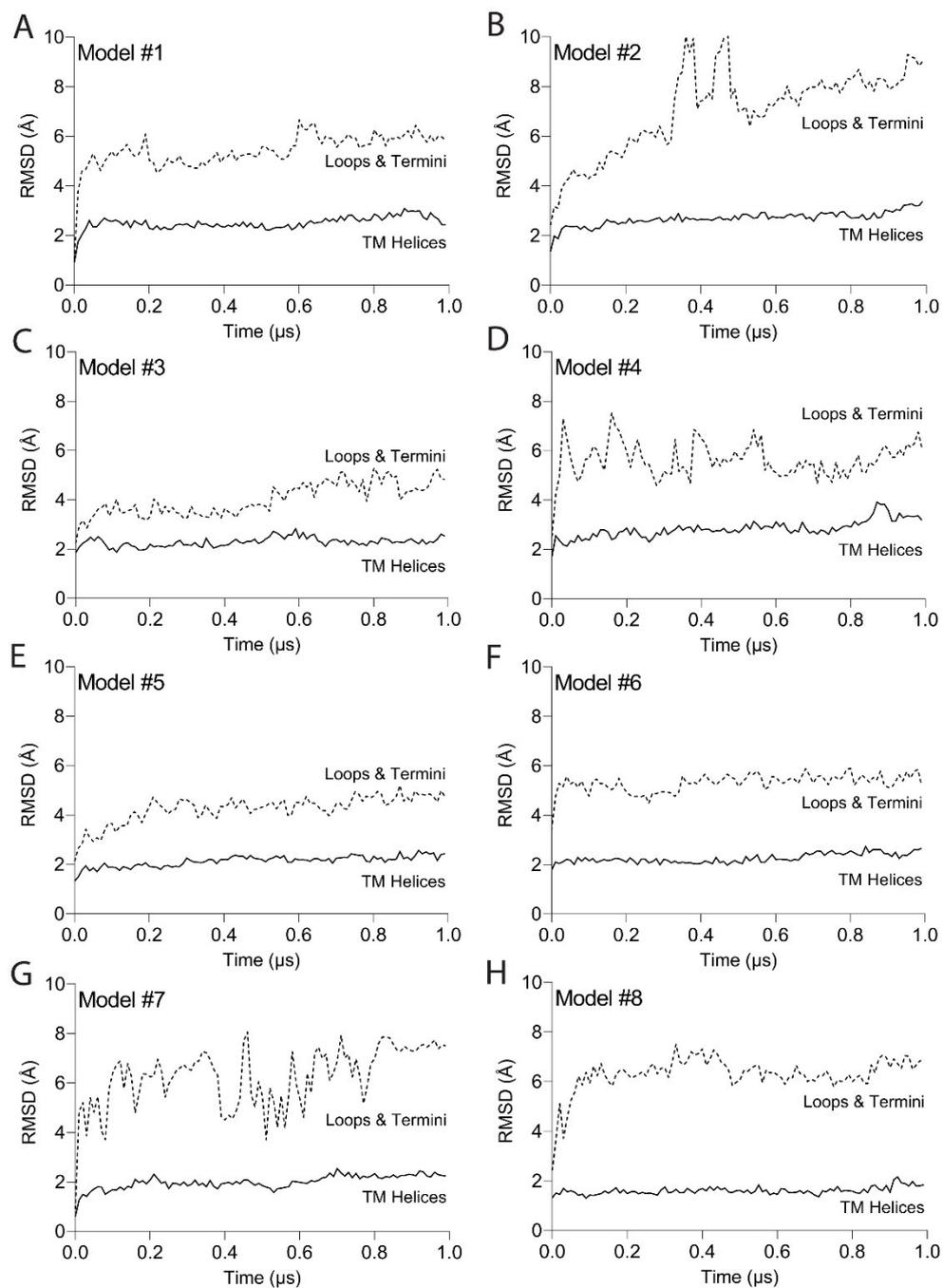


Figure S2: The panels A-H show  $\text{Ca}$  RMSD profiles computed for simulations #1 - #8. For clarity, we read coordinates each 1 ns.

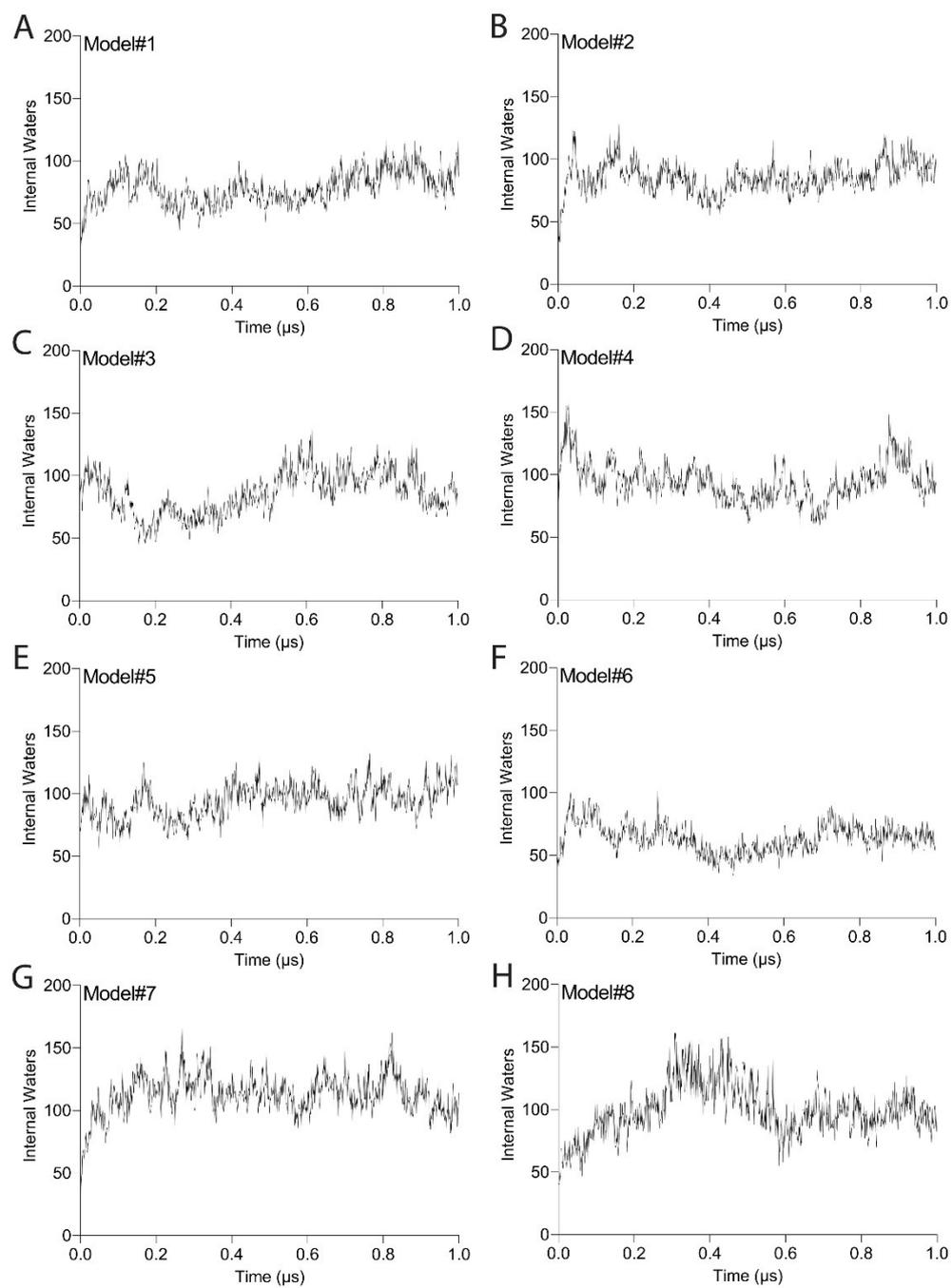


Figure S3: Number of internal water molecules that visit the inter-helical region of GPR68 structural models during simulations #1 - #8.

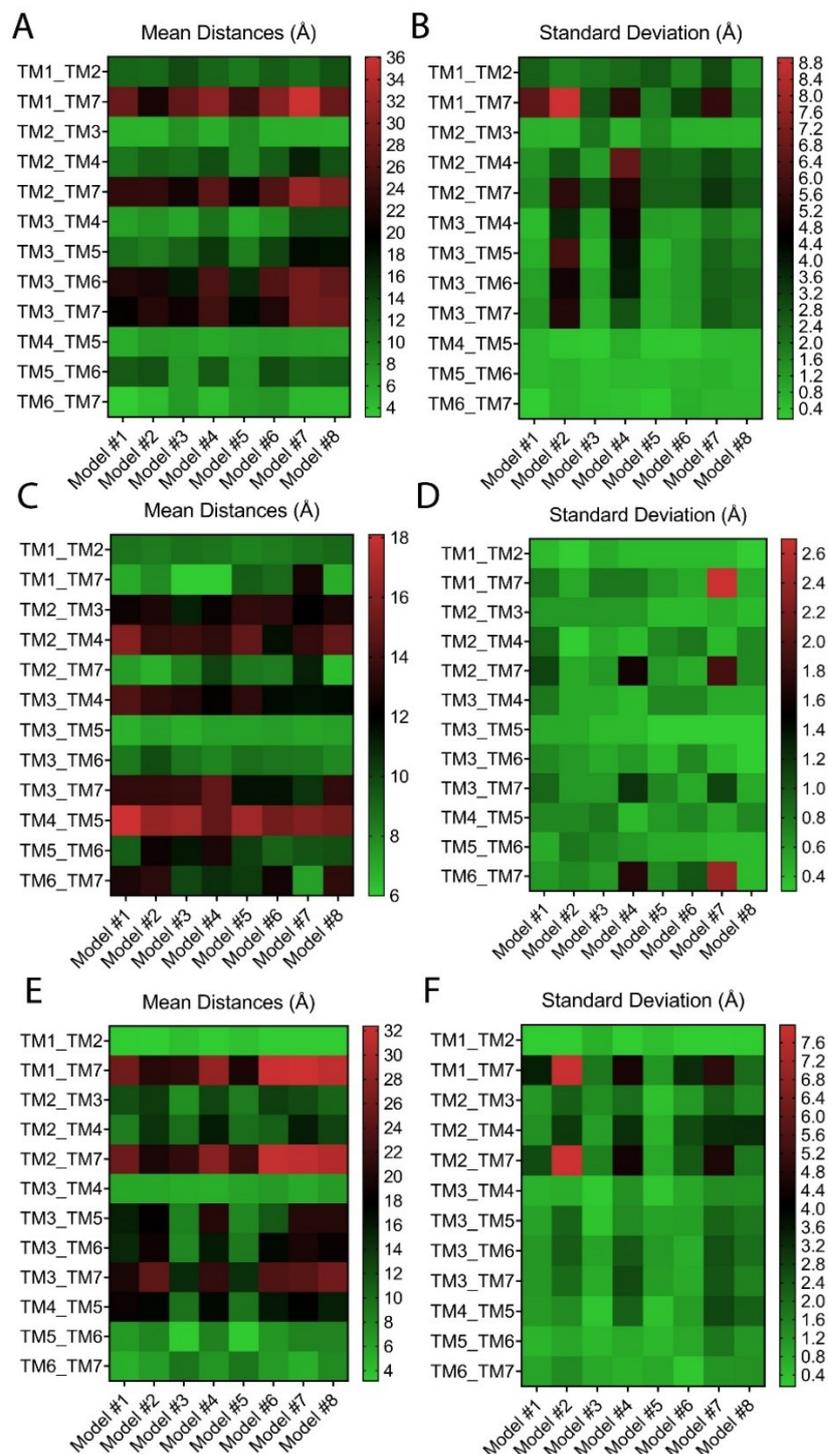


Figure S4: Inter-helical distances computed from simulations of Models #1 - #8. (A-F) Inter-helical distances and corresponding standard deviations measured at the extracellular tips of TM helices (panels A, B), at the central plane (panels C, D), and at the cytoplasmic ends of the TM helices (panels E, F). Mean values and standard deviations were computed every co-ordinate set from the last 600ns of each simulation.

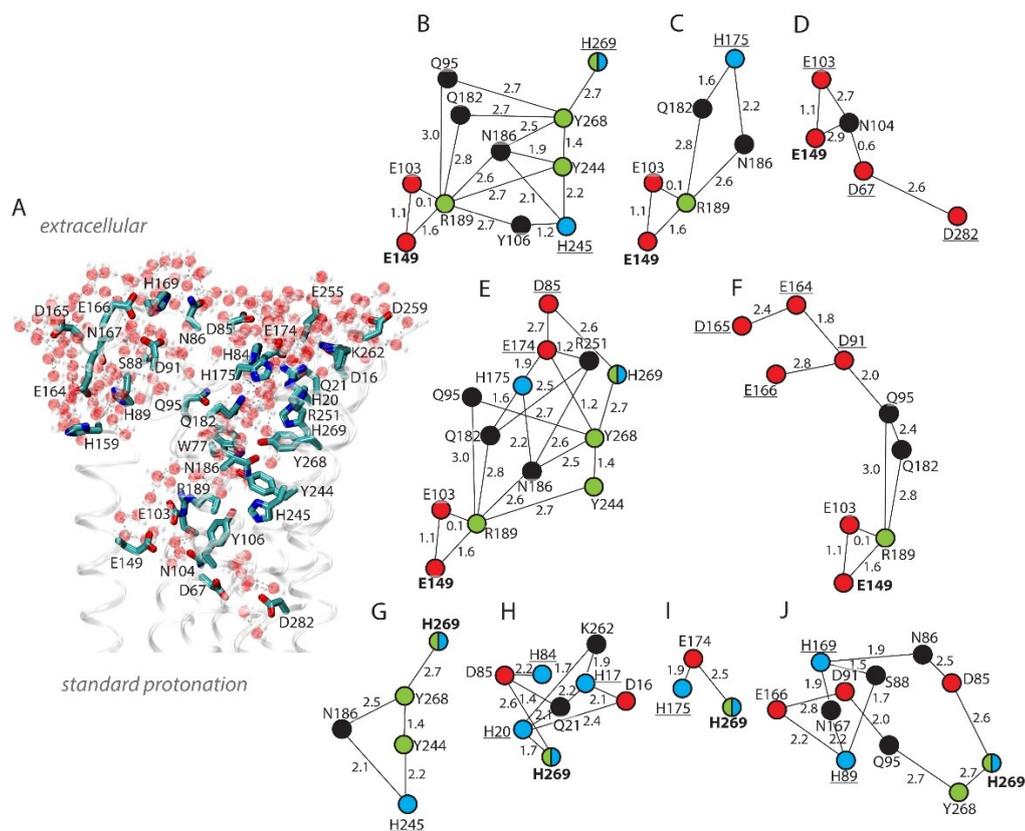


Figure S5: Water molecules mediate the dynamic H-bond networks of GPR68. (A) Molecular graphics illustrating water molecules within 3.5 Å of the amino acid residues shown in panels B-J. (B-J) H-bond clusters corresponding to main text Figure 39 B-J, here shown with the average number of water molecules for each edge representing a direct-H bond or a water-mediated H-bond. The figure has been reused from (Kapur et al., 2023).

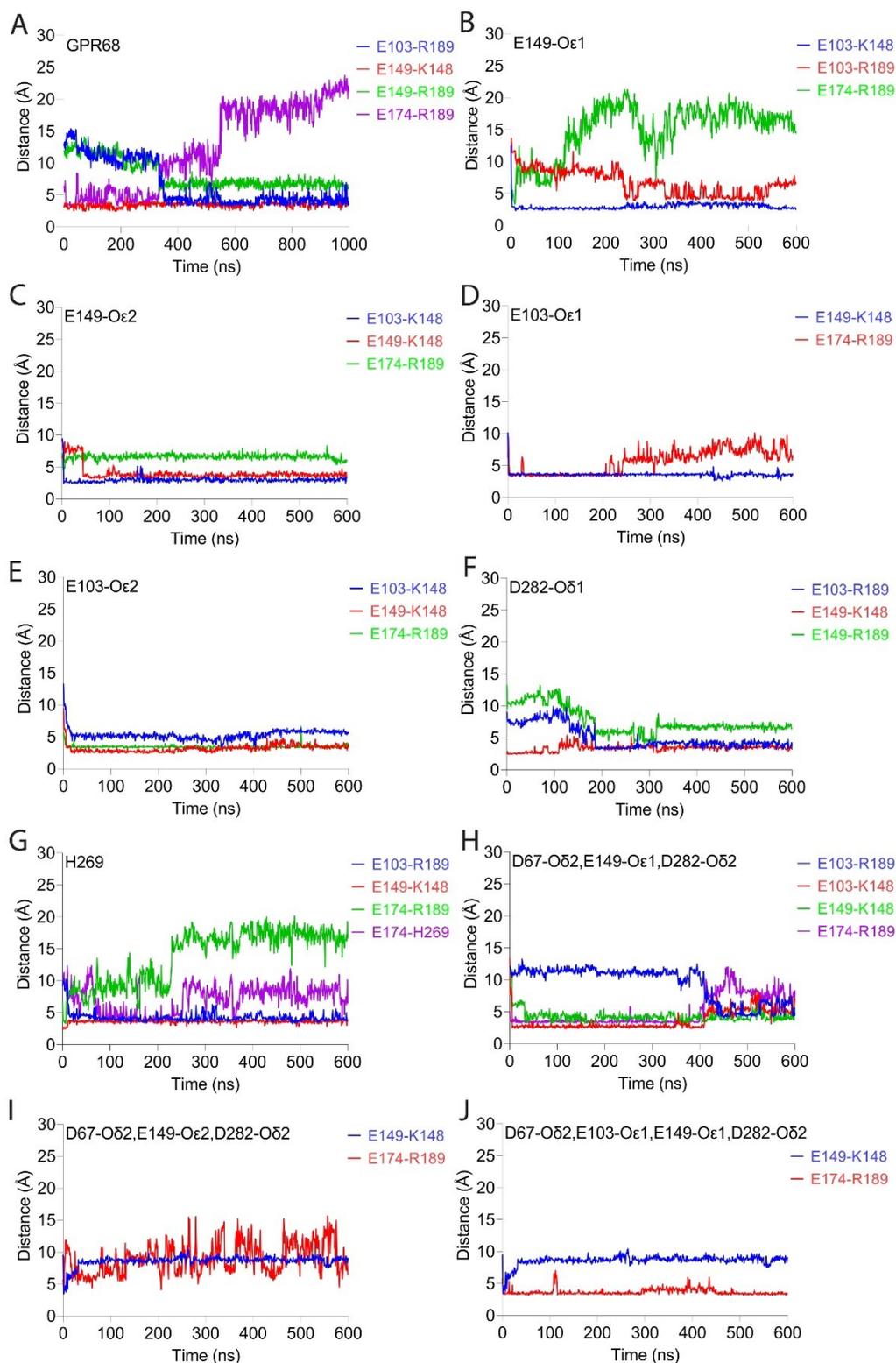


Figure S6: Charged-pair interactions of E103<sup>3,34</sup>, E149<sup>4,53</sup>, E174<sup>45,52</sup>, K148<sup>4,52</sup>, and R189<sup>5,42</sup>. (A-J) Time series for GPR68 with standard protonation states (panel A) and with non-standard protonations (B-J). Calculations were done with VMD using default salt-bridge criteria.

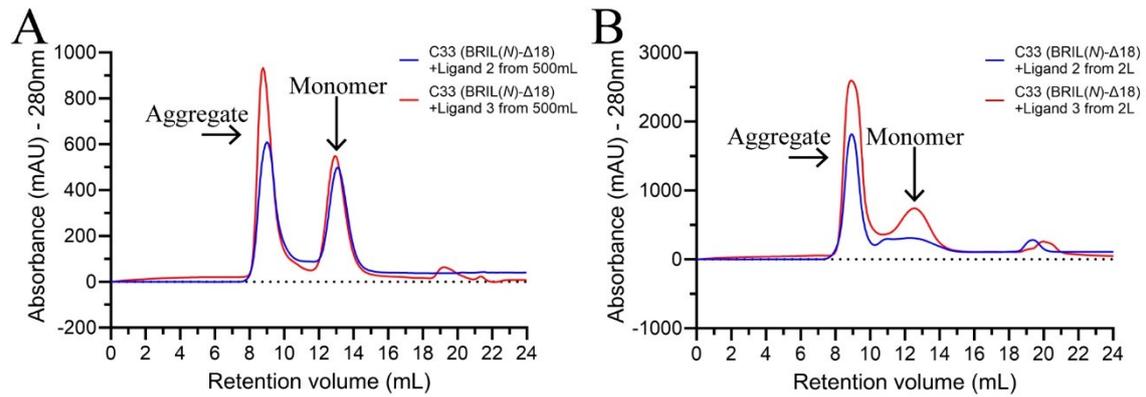


Figure S7: Impact of scaling up purification from 500mL to 2L using a representative GPR68 construct, C33. Size exclusion chromatograms from the pooled eluates after affinity purification (FLAG) from a 500 mL pellet purification (panel A) and 2 L pellet purification (panel B) of C33 purified using a Superdex 200 increase 10/300 GL column in the presence of DDM/CHS at pH 6.0. All purifications were performed in the presence of either Ligand 2 or Ligand 3.

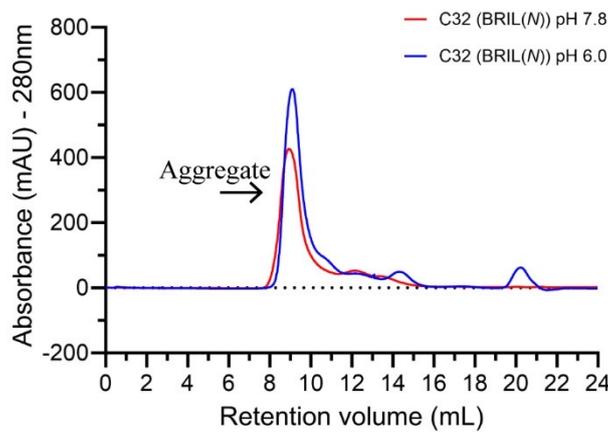


Figure S8: Effect of pH on GPR68 construct, C32. Size exclusion chromatograms from the pooled eluates after affinity purification (TALON or FLAG) from a 500 mL pellet purification using a Superdex 200 increase 10/300 GL column in the presence of DDM/CHS at pH 6.0 or pH 7.8.

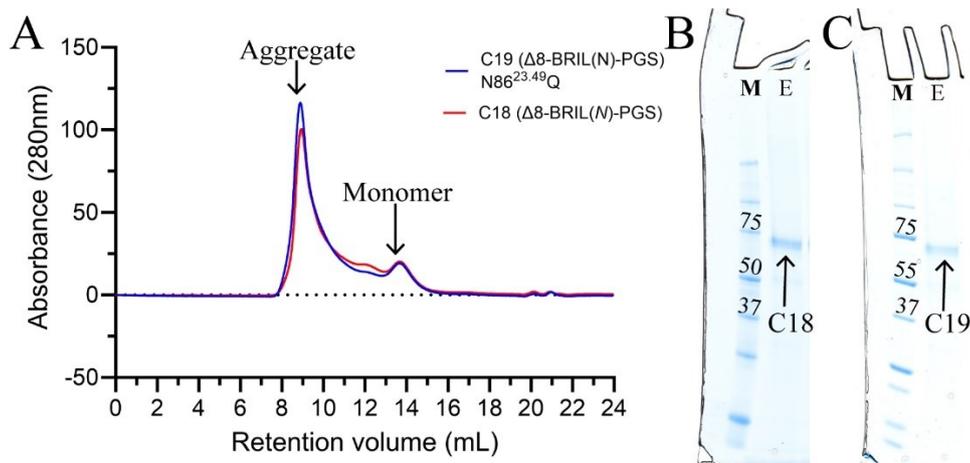


Figure S9: Effect of mutating third putative N-glycosylation site on GPR68 construct. Size exclusion chromatograms from the pooled eluates after affinity purification (TALON) from a 500 mL pellet purification using a Superdex 200 increase 10/300 GL column in the presence of DDM/CHS at pH 7.8 for C18 and C19 (panel A). The SDS PAGE Gels from the TALON eluates for C18 (panel B) and C19 (panel C).

**A****Protein View: NB1284**

22133\_Anti-BRIL FAB\_Heavy Chain and Light Chain

Database: DK10  
 Score: 2967  
 Monoisotopic mass ( $M_r$ ): 47626  
 Calculated pI: 8.82

Sequence similarity is available as an [NCBI BLAST search of NB1284 against nr](#).  
 Sequence similarity is available as an [bioRS search of NB1284 against nr](#).

**Search parameters**

MS data file: 22133\_GFR68\_11.mgf  
 Enzyme: Trypsin: cuts C-term side of KR unless next residue is P.  
 Fixed modifications: [Carbamidomethyl \(C\)](#)  
 Variable modifications: [Gln->pyro-Glu \(N-term Q\)](#), [Oxidation \(M\)](#)

**Protein sequence coverage: 68%**Matched peptides shown in **bold red**.

```

1  EVQLVESGGG LVQPGQSLRL SCAASQFNVV DFLSMWRQA FGRDLEWVAY
51  ISSSSGSTSY ADSVKGRFTI SADTSRNTAY LQMNLSRAED TAVYYCARWG
101 YWFGDFWWEA FDYWGQDTLV TVSSASTRGP SVFPLAPSSK STSGGTAALG
151 CLVKDYFPEP VIVSNNSGAL TSGVHTFAAV LQSSGLYSLS SVIVFPSSSL
201 GTQTYICNVN HKPSNTKVDK KVEPKSDIQM TQSPSSLSAS VGDRVTITCR
251 ASQSVSSAVA WYQQRKQKAP KLLIYSASSL YSGVPSRFSG SRSGDTFTLT
301 ISSLQPEDEFA TYCQQQLYLY SLVTFQGTTR VEIKRTVAAP SVFIFFPSSDS
351 QLRSQTASVV CLLNIFYFRE AKVQWQVNA LQSGNSQESV TEQDSKSDSTY
401 SLSSTLTLSK ADVEKRVVYA CEVTHQGLSS PVTRKSFNRG

```

**B****Protein View: NB1642**

22133 Anti-BRIL FAB Nanobody

Database: Ristw  
 Score: 593  
 Monoisotopic mass ( $M_r$ ): 17670  
 Calculated pI: 7.01

Sequence similarity is available as an [NCBI BLAST search of NB1642 against nr](#).  
 Sequence similarity is available as an [bioRS search of NB1642 against nr](#).

**Search parameters**

MS data file: 22133\_GFR68\_12.mgf  
 Enzyme: Trypsin: cuts C-term side of KR unless next residue is P.  
 Fixed modifications: [Carbamidomethyl \(C\)](#)  
 Variable modifications: [Gln->pyro-Glu \(N-term Q\)](#), [Oxidation \(M\)](#)

**Protein sequence coverage: 52%**Matched peptides shown in **bold red**.

```

1  MKYLLPTAAA GLLLLAAQPA MAAHHHHHHG SSGENLYFQG SQVQLQESGG
51  GLVQPGGSLR LSCAASGRIT SRVAMSWFRQ APGKEREFVA VARRSGDGAF
101 YADSVQGRFT VSRDARNTV YLQMNLSRKE DTAVYYCAID SDTFYSGSYD
151 YWGQGTQVTV SS

```

Figure S10: Mass Spectrometry, MALDI-TOF results showing successful identification of Anti-BRIL FAB (panel A) and Anti-BRIL-FAB nanobody (panel B). These results are from the corresponding SDS-PAGE gels shown for C61 purification shown in Figure 32B.

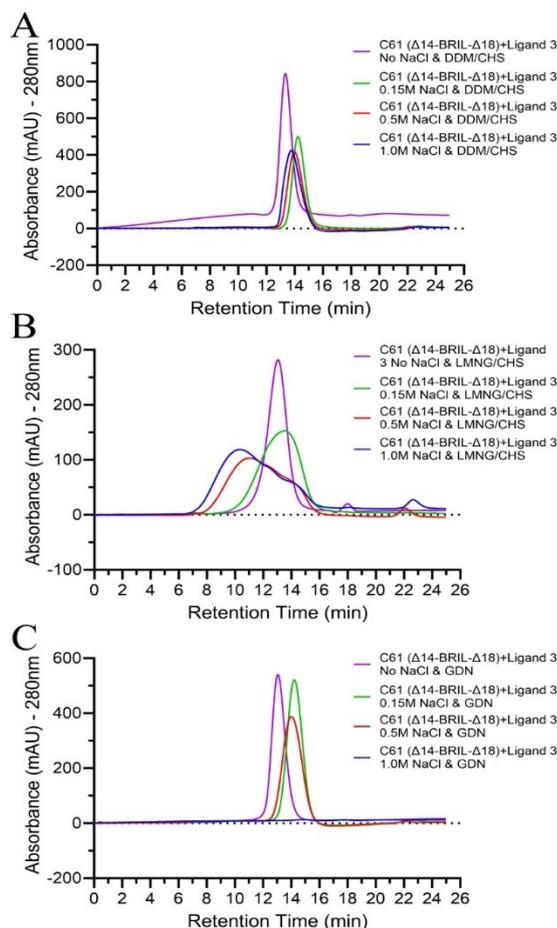


Figure S11: Effect of variable salt concentration (NaCl) on GPR68 construct, C61, ranging from no salt to 1M NaCl in DDM/CHS (panel A), LMNG/CHS (panel B) and GDN (panel C). Size exclusion chromatograms from the pooled eluates after affinity purification (TALON) from a 500 mL pellet purification using a Superose 6 increase 5/150 GL column at pH 7.4.

**A****Protein View: NB1788**

24040\_GPR68\_C99\_Zebrafish GPR68

Database: Ristw  
 Score: 541  
 Monoisotopic mass ( $M_r$ ): 81020  
 Calculated pI: 6.07

Sequence similarity is available as an [NCBI BLAST search of NB1788 against nr](#).  
 Sequence similarity is available as a [BioRS search of NB1788 against nr](#).

**Search parameters**

MS data file: 24040\_GPR68\_BK1\_011.mqf  
 Enzyme: Trypsin: cuts C-term side of KR unless next residue is P.  
 Fixed modifications: Carbamidomethyl (C)  
 Variable modifications: Gln->pyro-Glu (N-term Q), Oxidation (M)

**Protein sequence coverage: 17%**Matched peptides shown in **bold red**.

1 MCTIIALSYI FCIVFADVFD DQKYSRELF TRVVFVLEL DQDVNSRFFS  
 51 VSSGSGSAT VSRILTEPIC TQRLEFVWF TIVTLLIVQ QKSRIFQRM  
 101 VRSCLFFEMAM RQGVQERTI SPSKDNQKRT RANVVEKDT LVRRIELGKI  
 151 DFREKQNLG KRLVYDNRH NYVTADQKQ NGIKANFFIR RHIEDGSVCL  
 201 ADHYQNTPI GQDFVLDFN NYLSTQALS KDKNFKDRH VLEFVTAAG  
 251 ITRHDELYK GSENLVFGI QCTIHDHQQ YLFSVAYLL LVOLFANAY  
 301 SLYRMLQLK ARNELGVYLL NLTISDLYL GSLPLMLQYI FQGDMSGSE  
 351 MLCGLDFEL VENTVYRIPF LCCDIDRVL AVYVFFRFA FRTVRAATIV  
 401 STYMKRELA WQVFFLHRE LRSQDRGQV QREHFRHTW EYLIPIRFI  
 451 IDFLEFLDI SVYVFRULA VADLEDNEMV LKNDVYVIEK ANNAQQRWA  
 501 LTRMAAALD AQATVFRKQ DMSRDSRPMK DFRHGFELV QIDRALKLA  
 551 NQRVWRAGA AARQLRTRH AYIQVYLKQ RIKYLVSTI VFLVCFSEY  
 601 NIFLVRVTF ERDGNFESI FNYVHFKLL TSPNVADFA LYCFISESAQ  
 651 KGIQKADAC TRVFCCELE VLFQKPMRH QFEKGGSGG GSGGSSAMSH  
 701 PQEEK

**B****Protein View: NB1641**

22133\_Anti-BRIL FAB\_Heavy Chain and Light Chain

Database: Ristw  
 Score: 779  
 Monoisotopic mass ( $M_r$ ): 47626  
 Calculated pI: 8.82

Sequence similarity is available as an [NCBI BLAST search of NB1641 against nr](#).  
 Sequence similarity is available as a [BioRS search of NB1641 against nr](#).

**Search parameters**

MS data file: 24040\_GPR68\_BK2\_009.mqf  
 Enzyme: Trypsin: cuts C-term side of KR unless next residue is P.  
 Fixed modifications: Carbamidomethyl (C)  
 Variable modifications: Gln->pyro-Glu (N-term Q), Oxidation (M)

**Protein sequence coverage: 51%**Matched peptides shown in **bold red**.

1 EVQLVESGGG LVQPGQSLRL SCAASGFNVV DFLSHLVQA FORKLEWVAY  
 51 ISSSSGTSY ADSVKRRTI SADTSKNTAY LQMSLRARL TAVYCARWG  
 101 YWFGFHWKA FDVWQQGLV TVSSASTKQV SVFPLAPSK STSGQTAALG  
 151 CLVWDVFFEP VIVMNSGAL TSGVHTFPAV LQSSGLYSLS SVIVVSSSL  
 201 GTQTYICHVN HRFNTRVQR KVEPKSDIGM TQSPSSLSAS VGRVITICR  
 251 ASQVSSAVA WYQKPKQKAP KLLIYSASSL YSGVSRFSG SRSDTFTLI  
 301 ISSIQEDFA TYQCQYLYL SLVTFQQTQ VEIKRVAAP SVFIFPSSDS  
 351 QLFSSIASV CLLNPFYFAE RVCWRVNDNA LQSGNSQSV TQDSKSDSTY  
 401 SLSSTLTSK ADIEKRVVYA CEVTRQGLSS PVYKSFNRG

**C****Protein View: NB1642**

22133 Anti-BRIL FAB Nanobody

Database: Ristw  
 Score: 412  
 Monoisotopic mass ( $M_r$ ): 17670  
 Calculated pI: 7.01

Sequence similarity is available as an [NCBI BLAST search of NB1642 against nr](#).  
 Sequence similarity is available as a [BioRS search of NB1642 against nr](#).

**Search parameters**

MS data file: 24040\_GPR68\_BK1\_007.mqf  
 Enzyme: Trypsin: cuts C-term side of KR unless next residue is R.  
 Fixed modifications: Carbamidomethyl (C)  
 Variable modifications: Gln->pyro-Glu (N-term Q), Oxidation (M)

**Protein sequence coverage: 35%**Matched peptides shown in **bold red**.

1 NYTLFARA GILLIAQFA NGRNRRHS SSGKILFQD SQVQLQESG  
 51 GIVPQGSLS LKCAASRTI SRYANRFRQ AGRSREFA VSRSDGAF  
 101 YADRVRRTY VSRDARTV YLDNRLEPE DAVYICAD SPTFSSSD  
 151 WQVQIVTV SS

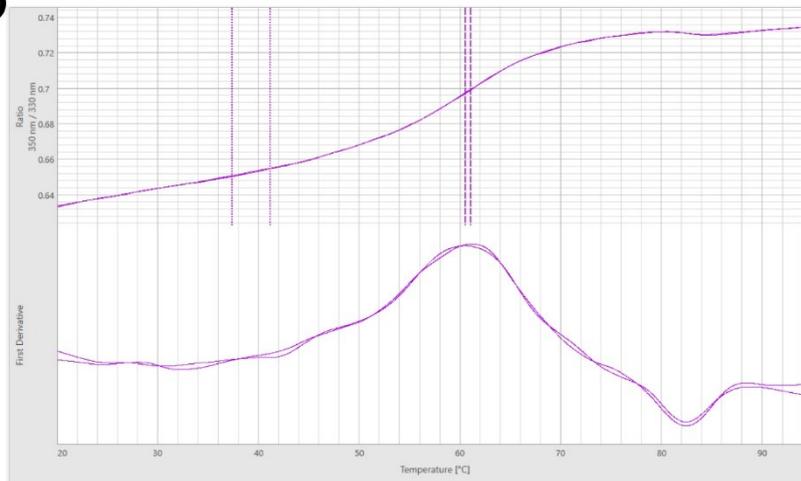
**D**

Figure S12: Mass Spectrometry or MALDI-TOF results showing successful identification of zebrafish GPR68, C99 (panel A), Anti-BRIL FAB (panel B) and Anti-BRIL-FAB nanobody (panel C). The nanoDSF profile highlights the 350/330 nm ratio and the first derivative of this ratio for C99 when purified in presence of Ligand 3 (panel D). The melting temperature for this construct was about ~60-62°C (panel D). These results are from the corresponding SDS-PAGE gels shown for C99 purification with FAB and NB complex in presence of Ligand 3 shown in Figure 34 B,C.



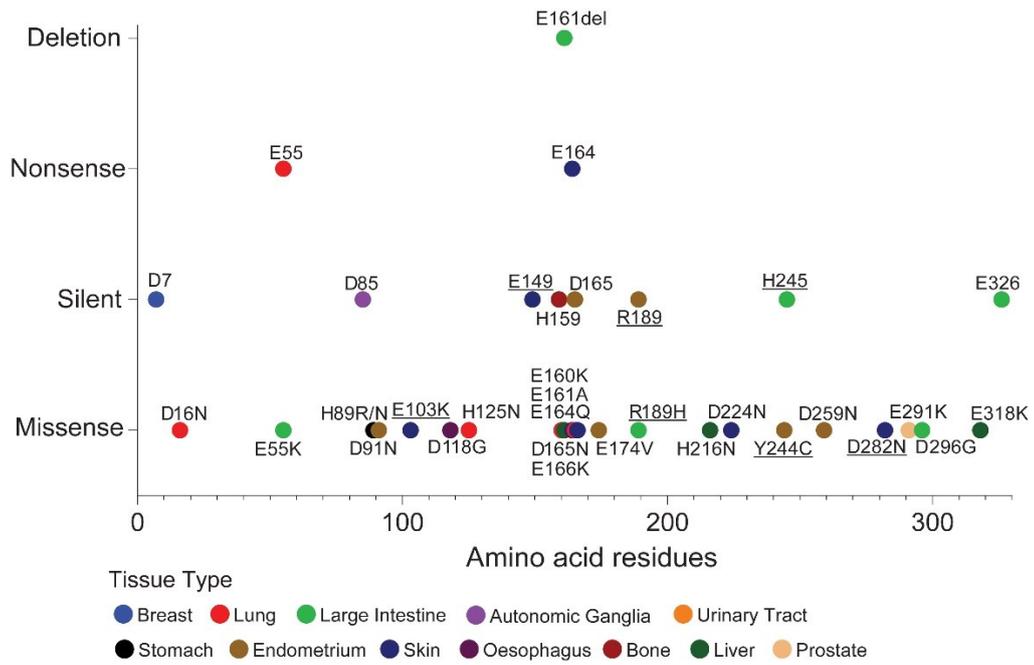


Figure S14: Select amino acid residues with implications in cancer-associated mutations of GPR68. The figure summarizes the result of our COSMIC database search on 13.10.2022. The underlined amino acid residues are a part of the dynamic water-mediated H-bond network of GPR68 (Figure 39).

### Appendix 3: Supplementary tables

#	Paths	Joint Occupancy (%)
<b>Panel B</b>		
1	E149-R189-N186-H245	48
2	E149-R189-Y244-H245	42
3	E149-R189-Y106-H245	28
4	E149-R189-N186-Y268-H269	12
5	E149-R189-Y244-Y268-H269	18
6	E149-R189-Q182-Y268-H269	8
7	E149-R189-Q95-Y268-H269	4
<b>Panel C</b>		
8	E149-R189-N186-H175	38
9	E149-R189-Q182-H175	26
<b>Panel D</b>		
10	E149-E103	100
11	E149-N104-D67	31
12	E149-E103-N104-D67	45
13	E149-N104-D67-D282	14
14	E149-E103-N104-D67-D282	21
<b>Panel E</b>		
15	E149-R189-Y244-Y268-E174	21
16	E149-R189-N186-Y268-E174	17
17	E149-R189-N186-R251-E174	25
18	E149-R189-N186-H175-E174	28
19	E149-R189-Q182-Y268-E174	11
20	E149-R189-Q182-R251-E174	16
21	E149-R189-Q182-H175-E174	16
22	E149-R189-Q95-Y268-E174	3
23	E149-R189-Y244-Y268-H269-D85	8
24	E149-R189-Y244-Y268-E174-D85	6
25	E149-R189-N186-Y268-H269-D85	5
26	E149-R189-N186-Y268-E174-D85	4
27	E149-R189-N186-R251-E174-D85	10
28	E149-R189-N186-H175-E174-D85	13
29	E149-R189-Q182-Y268-H269-D85	4
30	E149-R189-Q182-R251-E174-D85	7
31	E149-R189-Q182-H175-E174-D85	8

<b>32</b>	E149-R189-Q95-Y268-E174-D85	2
<b>Panel F</b>		
<b>33</b>	E149-R189-Q95-D91	17
<b>34</b>	E149-R189-Q182-Q95-D91	25
<b>35</b>	E149-R189-Q95-D91-E166	7
<b>36</b>	E149-R189-Q182-Q95-D91-E166	11
<b>37</b>	E149-R189-Q95-D91-E164	12
<b>38</b>	E149-R189-Q182-Q95-D91-E164	17
<b>39</b>	E149-R189-Q95-D91-E164-D165	7
<b>40</b>	E149-R189-Q182-Q95-D91-E164-D165	10
<b>Panel G</b>		
<b>41</b>	H269-Y268-Y244-H245	16
<b>42</b>	H269-Y268-N186-H245	11
<b>Panel H</b>		
<b>43</b>	H269-H20	73
<b>44</b>	H269-D85-H84	30
<b>45</b>	H269-H20-D16-H17	21
<b>46</b>	H269-D85-Q21-H17	19
<b>47</b>	H269-H20-Q21-H17	24
<b>48</b>	H269-H20-K262-H17	8
<b>Panel I</b>		
<b>49</b>	H269-E174-H175	17
<b>Panel J</b>		
<b>50</b>	H269-D85-N86-H169	4
<b>51</b>	H269-D85-N86-H169-S88-H89	1
<b>52</b>	H269-D85-N86-H169-N167-H89	1
<b>53</b>	H269-Y268-Q95-D91-E166-H89	2

Table S1: Joint occupancy values of H-bond paths shown in Figure 39 of the main text.

## List of Figures

FIGURE 1: OVERALL ARCHITECTURE OF A G PROTEIN-COUPLED RECEPTOR (GPCR) HIGHLIGHTING THE SEVEN TRANSMEMBRANE (TM) HELICES WITH THE ALTERNATING EXTRACELLULAR LOOPS (ECL), INTRACELLULAR LOOPS (IL), LIGAND BINDING SITE AND G PROTEIN COUPLING SITE. THIS FIGURE HAS BEEN ADAPTED AND REUSED <sup>[17]</sup> . .....	2
FIGURE 2: OUTLINE OF THE LIFE CYCLE OF A G PROTEIN-COUPLED RECEPTOR, EMPHASIZING THE TRAFFICKING, RECYCLING, AND DEGRADATION VIA UBIQUITINATION OF GPCRS. THIS FIGURE HAS BEEN REUSED <sup>[39]</sup> . .....	4
FIGURE 3: CLASSIFICATION OF G PROTEIN-COUPLED RECEPTOR FAMILY BASED ON THE GRAFS CLASSIFICATION SYSTEM. THIS FIGURE FROM <sup>[43]</sup> HAS BEEN REUSED WITH PERMISSION. ....	5
FIGURE 4: SIGNALING PATHWAYS FOR G PROTEIN-COUPLED RECEPTORS. THIS FIGURE HAS BEEN REUSED <sup>[52]</sup> UNDER THE CREATIVE COMMONS ATTRIBUTION LICENSE. ....	7
FIGURE 5: VARIOUS TYPES OF DOSE-RESPONSE CURVES AND A SPECTRUM OF RELATIVE EFFICACY FOR GPCR LIGANDS. THIS FIGURE HAS BEEN REUSED <sup>[61]</sup> UNDER THE CREATIVE COMMONS ATTRIBUTION LICENSE. ....	8
FIGURE 6: CONSERVED MOTIFS OF CLASS A GPCRS. THE FIGURE HAS BEEN ADAPTED AND REUSED <sup>[85]</sup> UNDER THE CREATIVE COMMONS ATTRIBUTION LICENSE. ....	10
FIGURE 7: MECHANISM OF THE ACID BASE HOMEOSTASIS IN THE HUMAN BODY. THIS FIGURE HAS BEEN ADAPTED AND REUSED <sup>[107]</sup> . .....	13
FIGURE 8: THE PHYSIOLOGICAL AND PATHOLOGICAL PROPERTIES OF THE PROTON SENSING SUB-FAMILY OF CLASS A GPCRS. ABBREVIATIONS: ER: ENDOPLASMIC RETICULUM, CAMP: CYCLIC ADENOSINE MONOPHOSPHATE, DRG: DORSAL ROOT GANGLIA, LPS: LIPOPOLYSACCHARIDE, IL-6: INTERLEUKIN 6, TNFA: TUMOR NECROSIS FACTOR A, TME: TUMOR MICROENVIRONMENT, SMC: SMOOTH MUSCLE CELLS, COX-2: CYCLOOXYGENASE-2, PGI2: PROSTAGLANDIN I2. THE FIGURE HAS BEEN REUSED <sup>[116]</sup> UNDER THE CREATIVE COMMONS ATTRIBUTION LICENSE. ....	14
FIGURE 9: IN VITRO PH ACTIVATION RANGE, PREFERENCES FOR SIGNALING PATHWAYS, AND SIGNALING OF THE PROTON SENSING GPCRS. THE FOLLOWING FIGURE DISPLAYS THE PH ACTIVATION RANGE FOR EACH OF THE FOUR RECEPTORS, NAMELY TDAG8 (GPR65, GREEN), G2A (GPR132, BLUE), GPR4 (YELLOW), AND OGR1 (GPR68, RED). ABBREVIATIONS: AC: ADENYLYL CYCLASE, CAMP: CYCLIC ADENOSINE MONOPHOSPHATE, RHO: RAS HOMOLOGUE, ROCK: RHO-ASSOCIATED COILED-COIL-CONTAINING PROTEIN KINASE, PLC: PHOSPHOLIPASE C, PKC: PROTEIN KINASE C, $[Ca^{+2}]_i$ : INTRACELLULAR CALCIUM CONCENTRATION, RAS: RAT SARCOMA PROTEIN, MAPK: MITOGEN-ACTIVATED PROTEIN KINASE(S), PI3K: PHOSPHOINOSITIDE 3-KINASE, AKT: PROTEIN KINASE B, LPAR: LYSOPHOSPHATIDIC ACID RECEPTOR, S1PR: SPHINGOSINE-1-PHOSPHATE RECEPTOR. THIS FIGURE HAS BEEN REUSED <sup>[116]</sup> UNDER CREATIVE COMMONS ATTRIBUTION LICENSE. ....	15

FIGURE 10: THE NORMALIZED EXPRESSION (NTPM) LEVELS OF THE GPR8 MITOCHONDRIAL RNA (MRNA) IN 55 DIFFERENT TISSUE TYPES. THIS FIGURE HAS BEEN REUSED FROM THE HUMAN PROTEIN ATLAS DATABASE <sup>[124]</sup>. THE WEBSITE ([HTTPS://WWW.PROTEINATLAS.ORG/ENSG00000119714-GPR68/TISSUE](https://www.proteinatlas.org/ENSG00000119714-GPR68/TISSUE)) WAS ACCESSED ON 05.03.2024. .... 19

FIGURE 11: THE SNAKE DIAGRAM OF GPR68 HIGHLIGHTING THE PROPOSED STRUCTURAL ELEMENTS OF GPR68 FOR PH SENSING. THIS FIGURE WAS PREPARED USING THE GPCRDB WEBSITE <sup>[6]</sup>. .... 21

FIGURE 12: 2D STRUCTURES OF GPR68 ANTAGONISTS (OGREMORPHIN) AND ALLOSTERIC MODULATORS (LORAZEPAM, DIAZEPAM, OGERIN, COMPOUND 71 AND CONOPHYLLINE). THE STRUCTURES ARE REPRESENTED WITHOUT EXPLICIT HYDROGENS. .... 23

FIGURE 13: THIS FIGURE ILLUSTRATES COMMONLY USED FUSION PROTEINS WHICH ARE INSERTED IN THE ICL3 OF A GPCR, SHOWN AS A SNAKE DIAGRAM. THIS FIGURE HAS BEEN REUSED <sup>[236]</sup>. .... 26

FIGURE 14: PHASE DIAGRAM OF THE MONOOLEIN/WATER SYSTEM DETERMINED AT 20 °C. IA3D AND PN3M REPRESENT CUBIC PHASES. THE FIGURE HAS BEEN REUSED <sup>[326]</sup>. .... 56

FIGURE 15: WORKING OF SINGLE PARTICLE ELECTRON MICROSCOPY. (A) THE FIGURE DEPICTS THE PROCESS OF IMAGE FORMATION IN AN ELECTRON MICROSCOPE. (B) SCHEMATIC DEPICTING THE FUNDAMENTAL PROCESS OF DATA GATHERING FOR ELECTRON TOMOGRAPHY. WHEN THE SPECIMEN IS INCLINED IN RELATION TO THE ELECTRON BEAM, SEVERAL IMAGES ARE CAPTURED OF THE IDENTICAL AREA. (C) SELECTED PROJECTION VIEWS OF THE PROTEIN SAMPLE IN VITRIFIED FILM WHICH WERE USED TO REBUILD THE THREE-DIMENSIONAL VOLUME BY USING A COLLECTION OF PROJECTION IMAGES THAT CAN BE USED TO CREATE BACK-PROJECTION PROFILES. THE FIGURE HAS BEEN REUSED WITH PERMISSION FROM <sup>[336]</sup>. .... 59

FIGURE 16: HTRF IP-ONE ASSAY PRINCIPLE. THE SNAKE DIAGRAM IN THE CIRCLE REPRESENTS A GPCR AND THE CYAN SPHERE REPRESENTS THE COGNATE AGONIST FOR THE RECEPTOR. THE ABBREVIATIONS USED IN THE FIGURE ARE – PLC (PHOSPHOLIPASE C), IP<sub>3</sub> (INOSITOL, TRIPHOSPHATE), IP<sub>1</sub> (INOSITOL MONOPHOSPHATE), LiCl (LITHIUM CHLORIDE), AND FRET (FÖRSTER RESONANCE ENERGY TRANSFER). THE FIGURE HAS BEEN REUSED <sup>[353]</sup> UNDER CREATIVE COMMONS ATTRIBUTION LICENSE. .... 66

FIGURE 17: PRINCIPLE FOR TIME-RESOLVED FLUORESCENCE. THE FIGURE HAS BEEN REUSED <sup>[353]</sup> UNDER CREATIVE COMMONS ATTRIBUTION LICENSE. .... 67

FIGURE 18: A GENERIC REPRESENTATIVE CONSTRUCT DESIGN STRATEGY FOR HUMAN GPR68 USED IN THIS THESIS. THE CONSERVED CLASS A GPCR DISULFIDE BOND HAS BEEN SHOWN IN YELLOW SOLID LINE AND THE PUTATIVE DISULFIDE BOND HAS BEEN SHOWN IN DOTTED YELLOW LINE. ADDITIONALLY, THE CYSTEINES INVOLVED IN DISULFIDE BOND FORMATION ARE HIGHLIGHTED IN YELLOW. THE COMMONLY USED TRUNCATIONS THROUGHOUT THE THESIS AT N-TERMINUS OR C-TERMINUS ARE MARKED WITH A BLACK SOLID LINE ALONG WITH DELETION OF AMINO ACIDS REPRESENTED WITH GREEK LETTER ‘Δ’ FOLLOWED BY NUMBER OF AMINO ACIDS. THE MUTATED AMINO ACIDS DESCRIBED IN THIS SECTION HAVE BEEN HIGHLIGHTED IN BLUE. CERTAIN REPRESENTATIVE MODIFICATIONS IN THE CONSTRUCTS WERE INSERTION OF FUSION PROTEINS (EITHER AT N-TERMINUS OR ICL3),

TOBACCO ETCH VIRUS (TEV) PROTEASE CLEAVAGE SITE OR PRESSION (PREC) PROTEASE CLEAVAGE SITE, PURIFICATION TAGS LIKE DECA-HISTIDINE TAG (10× HIS) OR FLAG TAG, AND HAEMAGGLUTININ SIGNAL PEPTIDE (HA). ..... 70

FIGURE 19: REPRESENTATIVE GPR68 PCR PRODUCTS AND THE OPTIMAL BACULOVIRUS TITER SCREENING FOR GPR68 CONSTRUCTS. (A) AGAROSE GEL ELECTROPHORESIS OF TWO BACMIDS FROM GPR68 C21, WHERE M REPRESENTS THE DNA MARKER. (B) ANTI-HIS TAG WESTERN BLOT FROM THE SDS-PAGE OF THE OVEREXPRESSED PROTEINS FROM THE CELL PELLET SAMPLES OF GPR68 C1-C9, HERE M REPRESENTS THE PROTEIN MARKER..... 73

FIGURE 20: SCREENING OF DIFFERENT FUSION PROTEINS INSERTED IN THE ICL3 OF GPR68. (A) SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (TALON) FOR C2-C4 AND C6. THE SEC WAS PERFORMED ON SUPERDEX 200 INCREASE 10/300 GL FROM 500 ML CELL PELLET AND ALL THE PURIFICATION STEPS WERE PERFORMED USING DDM/CHS AT PH 7.8. (B-E) SDS PAGE OF THE FIRST THREE TALON ELUATES FOR C2 (PANEL B), C3 (PANEL C), C4 (PANEL D) AND C6 (PANEL E). THE 'M' REPRESENTS THE PROTEIN MARKER, ALSO NOTE THAT THE MARKER USED IN PANEL E IS DIFFERENT FROM PANEL B-D. NOTE THAT THE NUMBERS ON THE MARKER BANDS REPRESENT THE CORRESPONDING MOLECULAR WEIGHT IN KDA FOR THE BAND. .... 75

FIGURE 21: SCREENING OF THE T4L FUSION PROTEIN INSERTED IN THE ICL3 OF GPR68 WITH TRUNCATION AT EITHER TERMINUS. (A) SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (TALON) FOR THE N-TERMINAL TRUNCATION CONSTRUCTS (C9 AND C10). (B) SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (TALON) FOR THE C-TERMINAL TRUNCATION CONSTRUCTS (C12-C15). THE SEC WAS PERFORMED ON SUPERDEX 200 INCREASE 10/300 GL FROM 500 ML CELL PELLET AND ALL THE PURIFICATION STEPS WERE PERFORMED USING DDM/CHS AT PH 7.8. NOTE THAT THE GREEK LETTER Δ REPRESENTS TRUNCATION, IF THE Δ IS BEFORE THE FUSION PROTEIN, IT IS AN N-TERMINAL TRUNCATION E.G., C9 (Δ8-T4L) MEANS THAT THE 8 AMINO ACIDS FROM THE N-TERMINUS ARE TRUNCATED, WHEREAS IF Δ IS AFTER THE FUSION IT REPRESENTS C-TERMINAL TRUNCATION E.G., C12 (T4L- Δ18) MEANS THAT THE 18 AMINO ACIDS FROM THE C-TERMINUS ARE TRUNCATED..... 76

FIGURE 22: SCREENING OF CONSTRUCTS CONTAINING A COMBINATION OF A FUSION PROTEIN INSERTED IN THE ICL3 OF GPR68 ALONG WITH A TRUNCATION AT EITHER TERMINI OR BOTH TERMINI. (A) SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (TALON) FOR THE GPR68 CONSTRUCTS WHICH HAVE EITHER T4L (C21) OR PGS (C24) AS ICL3 FUSION. (B) SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (TALON) FOR THE GPR68 CONSTRUCTS BASED ON ICL3 BRIL INSERTION. THESE BRIL CONSTRUCTS VARY IN INSERTION POINTS IN THE ICL3, ADDITION OF CHIMERIC BRIL (BRILX), OR A COMBINATION WITH TRUNCATION. THE SEC WAS PERFORMED ON SUPERDEX 200 INCREASE 10/300 GL FROM 500 ML CELL PELLET AND ALL THE PURIFICATION STEPS WERE PERFORMED USING DDM/CHS AT PH 7.8. THE TRUNCATION DENOTATION USED IN THIS FIGURE IS AS FORMERLY DESCRIBED IN FIGURE 21. .... 77

FIGURE 23: SCREENING OF THE COMBINATION CONSTRUCTS THAT INCLUDE A BRIL FUSION PROTEIN INSERTED AT THE N-TERMINUS OF THE GPR68 CONSTRUCTS AND IT IS DENOTED BY 'BRIL(N)'. (A) SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (TALON) FOR THE GPR68 CONSTRUCTS WHICH HAVE VARIATIONS IN THE C-TERMINAL TRUNCATION AND A BRIL(N). (B) SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (TALON) FOR THE GPR68 CONSTRUCTS HAVING BRIL(N) AND A COMBINATION OF TERMINI TRUNCATIONS. (C) SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (TALON) FOR THE GPR68 CONSTRUCTS HAVING BRIL(N) AND A COMBINATION OF TERMINI TRUNCATIONS. THE SEC WAS PERFORMED ON SUPERDEX 200 INCREASE 10/300 GL FROM 500 ML CELL PELLET AND ALL THE PURIFICATION STEPS WERE PERFORMED USING DDM/CHS AT PH 7.8. THE TRUNCATION DENOTATION USED IN THIS FIGURE IS AS FORMERLY DESCRIBED FOR FIGURE 21. .... 78

FIGURE 24: SCREENING OF THE COMBINATION CONSTRUCTS (C10, C12 AND C20) IN PRESENCE OF COBALT CHLORIDE. SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (TALON, IF NOT MENTIONED OTHERWISE) FOR GPR68 C10 (PANEL A), GPR68 C20 (PANEL B), AND GPR68 C12 (PANEL C). THE SEC WAS PERFORMED ON SUPERDEX 200 INCREASE 10/300 GL FROM 500 ML CELL PELLET AND ALL THE PURIFICATION STEPS WERE PERFORMED USING DDM/CHS AT PH 7.8. THE TRUNCATION DENOTATION USED IN THIS FIGURE IS AS FORMERLY DESCRIBED FOR FIGURE 21. .... 80

FIGURE 25: SCREENING OF THE VARIOUS COMBINATION CONSTRUCTS IN PRESENCE OF COBALT CHLORIDE. SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (FLAG) FOR GPR68 C7, C21 AND C24 (PANEL A), GPR68 C17, C31 AND C43 (PANEL B), AND GPR68 C1, C32, C33, C34 AND C35 (PANEL C). THE SEC WAS PERFORMED ON SUPERDEX 200 INCREASE 10/300 GL (PANEL A,B) AND ON SUPEROSE 6 INCREASE 10/300GL (PANEL C) FROM 500 ML CELL PELLET AND ALL THE PURIFICATION STEPS WERE PERFORMED USING DDM/CHS AT PH 7.8 (PANEL A,B) OR AT PH 6.8 (PANEL C). THE TRUNCATION DENOTATION USED IN THIS FIGURE IS AS FORMERLY DESCRIBED FOR FIGURE 21. .... 81

FIGURE 26: SCREENING OF THE GPR68 CONSTRUCTS IN PRESENCE OF PROPRIETARY LIGANDS (LIGAND 1, LIGAND 2 AND LIGAND 3). SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION TALON (PANEL A-D) AND FLAG (PANEL E-F) FOR PURIFICATION OF DIFFERENT GPR68 CONSTRUCTS IN LIGAND 1 (PANEL A-D), LIGAND 2 (PANEL E) AND LIGAND 3 (PANEL F). THE SEC WAS PERFORMED ON SUPERDEX 200 INCREASE 10/300 GL FROM 500 ML CELL PELLET AND ALL THE PURIFICATION STEPS WERE PERFORMED USING DDM/CHS AT PH 7.8. THE TRUNCATION DENOTATION USED IN THIS FIGURE IS AS FORMERLY DESCRIBED FOR FIGURE 21..... 82

FIGURE 27: SCREENING OF THE GPR68 POINT MUTATIONS. SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION TALON (PANEL A,B) AND FLAG (PANEL C,D) FOR PURIFICATION OF DIFFERENT GPR68 MUTATION CONSTRUCTS. THE SEC WAS PERFORMED ON

SUPERDEX 200 INCREASE 10/300 GL FROM 500 ML CELL PELLETT (PANEL A,B AND D) OR ON SUPEROSE 6 INCREASE 5/150 GL COLUMN (PANEL C) USING 50 ML CELL PELLETT AND ALL THE PURIFICATION STEPS WERE PERFORMED USING DDM/CHS AT PH 7.8 WITHOUT LIGAND (PANEL A) IN PRESENCE OF LIGAND 3 (PANEL B-D). THE TRUNCATION DENOTATION USED IN THIS FIGURE IS AS FORMERLY DESCRIBED FOR FIGURE 21..... 84

FIGURE 28: CONSTRUCT DESIGN STRATEGY FOR GPCRS USING ARTIFICIAL INTELLIGENCE BASED STRUCTURAL MODELING TOOL, ALPHAFOLD [213]. THE ACRONYMS USED IN THIS FIGURE ARE DESCRIBED AS FOLLOWS FOR THE FRAGMENT ANTIGEN-BINDING REGION (FAB), SINGLE MONOMERIC VARIABLE ANTIBODY DOMAIN DESIGNED TO BIND ANTI-BRIL FAB (NB), BETA-LACTAMASE FROM ESCHERICHIA COLI K12 (BLEK), CALCINEURIN SUB-UNIT A (CALA), CALCINEURIN SUB-UNIT A (CALB) AND PER-RESIDUE LOCAL DISTANCE DIFFERENCE TEST (PLDDT), A GLOBAL MODEL SCORING SYSTEM. (A) REPRESENTATIVE BRIL-FAB-NB COMPLEX. THIS PANEL WAS ADAPTED AND REUSED WITH PERMISSION UNDER CREATIVE COMMONS ATTRIBUTION LICENSE [368]. STRUCTURAL MODELS GENERATED FOR DIFFERENT GPR68 CONSTRUCTS CONTAINING THE FOLLOWING FUSION PROTEINS – BETA LACTAMASE (PANEL B), CALCINEURIN (PANEL C), BRIL FUSIONS BASED (PANEL D-F)..... 85

FIGURE 29: SCREENING OF THE GPR68 BRIL FUSION CONSTRUCTS FOR COMPLEXATION WITH ANTI-BRIL FAB. SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION FLAG AND SEC FOR COMPLEXATION WITH ANTI-BRIL FAB. THE SEC WAS PERFORMED ON UNIX-C SEC-300 4.6/300 COLUMN FOR C17 (PANEL A), C43 (PANEL B), C61 (PANEL C) AND C33 (PANEL D) OR ON SUPEROSE 6 INCREASE 5/150 GL COLUMN FOR C61 (PANEL E,F) ALL THE PURIFICATION STEPS WERE PERFORMED USING DDM/CHS AT DIFFERENT PH VALUES INDICATED IN THE RESPECTIVE FIGURE LEGEND. ALL PURIFICATION WERE PERFORMED IN PRESENCE OF LIGAND 3. THE TRUNCATION DENOTATION USED IN THIS FIGURE IS AS FORMERLY DESCRIBED FOR FIGURE 21..... 87

FIGURE 30: SCREENING OF THE GPR68 BRIL FUSION CONSTRUCTS FOR COMPLEXATION WITH ANTI-BRIL FAB BASED ON NEW STRATEGIES. SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION, FLAG AND SEC FOR COMPLEXATION WITH ANTI-BRIL FAB (BRIL-FAB) AND OR NANOBODY (NB). THE SEC WAS PERFORMED ON SUPEROSE 6 INCREASE 5/150 GL COLUMN FOR TESTING COMPLEXATION WITH ICL1 BRIL FUSION CONSTRUCT C73 (PANEL A), ICL2 BRIL FUSION CONSTRUCT C74 (PANEL B) AND ICL3 BRIL FUSION CONSTRUCT C75, EXPRESSED IN HEK293 CELLS (PANEL C). ALL THE PURIFICATION STEPS WERE PERFORMED USING DDM/CHS AT DIFFERENT PH 6.0. ALL PURIFICATION WERE PERFORMED IN PRESENCE OF LIGAND 3. THE TRUNCATION DENOTATION USED IN THIS FIGURE IS AS FORMERLY DESCRIBED FOR FIGURE 21..... 88

FIGURE 31: SUMMARY OF THE MAIN X-RAY CRYSTALLOGRAPHY TRIALS UNDERTAKEN USING HUMAN GPR68 CONSTRUCTS. SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (FLAG) FROM A 2L EXPRESSION OF C33 PURIFIED USING A SUPERDEX 200 INCREASE 10/300 GL COLUMN IN THE PRESENCE OF DDM/CHS AT PH 7.8 (PANEL A) AND ITS RESPECTIVE SDS-PAGE GEL (PANEL B). SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER

AFFINITY PURIFICATION (FLAG) FROM A 2L EXPRESSION OF C61 PURIFIED USING A SUPERDEX 200 INCREASE 10/300 GL COLUMN IN THE PRESENCE OF DDM/CHS AT PH 6.0 (PANEL C) AND ITS RESPECTIVE SDS-PAGE GEL (PANEL D). THE CRYSTALS FROM THE CRYSTALLIZATION OF C61 HAS BEEN SHOWN AFTER MAXIMUM GROWTH UNDER VISIBLE AND CROSS-POLARIZED LIGHT (PANEL E). ALL PURIFICATIONS WERE PERFORMED IN THE PRESENCE OF LIGAND 3. THE TRUNCATION DENOTATION USED IN THIS FIGURE IS AS FORMERLY DESCRIBED FOR FIGURE 21. THE 'M' REPRESENTS THE PROTEIN MARKER, AND THE NUMBERS ON THE MARKER BANDS REPRESENT THE CORRESPONDING MOLECULAR WEIGHT IN KDA FOR THE BAND. .... 89

FIGURE 32: SUMMARY OF THE REPRESENTATIVE CRYO-EM TRIALS UNDERTAKEN USING HUMAN GPR68 CONSTRUCTS. SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (FLAG) FROM A 2L EXPRESSION AND SDS-PAGE GELS OF THE FINAL AFTER SEC HAS BEEN SHOWN FOR C61 (PANEL A,B), C66 (PANEL C,D) AND C67 (PANEL E,F). THERE CONSTRUCTS WERE PURIFIED USING A SUPERDEX 200 INCREASE 10/300 GL COLUMN IN THE PRESENCE OF DDM/CHS AT PH 6.0. ALL PURIFICATIONS WERE PERFORMED IN THE PRESENCE OF LIGAND 3. THE TRUNCATION DENOTATION USED IN THIS FIGURE IS AS FORMERLY DESCRIBED FOR FIGURE 21. THE 'M' REPRESENTS THE PROTEIN MARKER, AND THE NUMBERS ON THE MARKER BANDS REPRESENT THE CORRESPONDING MOLECULAR WEIGHT IN KDA FOR THE BAND. .... 91

FIGURE 33: SUMMARY OF THE REPRESENTATIVE MICROGRAPHS AND 2D CLASSIFICATION OF THE CRYO-EM TRIALS UNDERTAKEN USING HUMAN GPR68 CONSTRUCTS – C61 (PANEL A,B), C66 (PANEL C,D) AND C67 (PANEL E,F). NOTE THAT THE REPRESENTATIVE 2D CLASSIFICATION HAS BEEN SHOWN AFTER A FEW ITERATIONS. .... 92

FIGURE 34: SUMMARY OF THE REPRESENTATIVE CRYO-EM TRIALS UNDERTAKEN USING ZEBRAFISH GPR68 CONSTRUCT C99. SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (STREP-TAG) FROM A 1.5L EXPRESSION AND SDS-PAGE GELS OF THE FINAL AFTER SEC HAS BEEN SHOWN FOR APOPROTEIN C99 (PANEL A), AND A COMPLEX OF C99 WITH FAB AND NB IN PRESENCE OF LIGAND 3 OR IN THE PRESENCE OF OGREMORPHIN (PANEL B). THE REPRESENTATIVE SDS-PAGE GEL (PANEL C), MICROGRAPH (PANEL D) AND 2D CLASSIFICATION (PANEL E) HIGHLIGHT THAT THE ZEBRAFISH GPR68 C99 WAS COMPLEXED SUCCESSFULLY WITH FAB AND NANOBODY IN THE PRESENCE OF LIGAND 3. THE CONSTRUCTS WERE PURIFIED USING A SUPEROSE 6 INCREASE 10/300 GL COLUMN IN THE PRESENCE OF LMNG/CHS AT PH 7.4. THE TRUNCATION DENOTATION USED IN THIS FIGURE IS AS FORMERLY DESCRIBED FOR FIGURE 21. THE 'M' REPRESENTS THE PROTEIN MARKER, AND THE NUMBERS ON THE MARKER BANDS REPRESENT THE CORRESPONDING MOLECULAR WEIGHT IN KDA FOR THE BAND. .... 94

FIGURE 35: DISTINCT CHARACTERISTICS OF HUMAN GPR68 BASED ON SEQUENCE AND STRUCTURAL MODEL. (A) CUT-AWAY VIEW OF MEMBRANE-EMBEDDED GPR68, BASED ON THE STANDARD PROTONATION SIMULATION. LIPID AND WATER MOLECULES SHOWN ARE WITHIN 10 Å AND 3 Å OF THE PROTEIN RESPECTIVELY. TITRATABLE RESIDUE SIDECHAINS THOUGHT IMPORTANT FOR PROTON

SENSING ARE SHOWN. (B) SEQUENCE CONSERVATION OF TITRATABLE RESIDUE SIDECHAINS THOUGHT IMPORTANT FOR PROTON SENSING IN GPR68 WITH OTHER CLASS A GPCR MEMBERS AND WITH OTHER PROPOSED PROTON SENSING GPCRS (GPR4, GPR65, GPR132, GPR31 AND GPR151). THE AMINO ACID RESIDUE IN GPR68 IS SHOWN AT THE TOP OF THE PANEL, WHILE THE CORRESPONDING NUMBER IN THE BW SCHEME IS SHOWN AT THE BOTTOM OF THE PANEL, IF IT EXISTS. THE HEIGHT OF EACH AMINO ACID RESIDUE SYMBOL INDICATES ITS CONSERVATION AMONG THE CLASS A GPCR SEQUENCES ALIGNED. THE THINNER ONE-LETTER SYMBOLS FOR POSITIONS 1.28, 1.31 AND 2.67 INDICATE GAPS IN SOME OF THE ALIGNED GPCR SEQUENCES. THE SEQUENCE ALIGNMENT WAS BASED ON GPCRDDB [51] AND THE SEQUENCE CONSERVATION CALCULATIONS WERE DONE BY WEBLOGO SERVER [369]. THE FIGURE HAS BEEN REUSED AND ADAPTED [339]. ..... 96

FIGURE 36: THE POSITION OF THE CONSERVED FUNCTIONAL MOTIFS AND THE HIS, ASP, AND GLU SIDECHAINS OF GPR68 ARE ILLUSTRATED. (A) THE AMINO ACID RESIDUE SIDECHAINS OF THE CONSERVED CLASS A GPCR MOTIFS IN GPR68. THE SIDECHAINS OF THE RESIDUES IN EACH MOTIF ARE DISTINCTLY COLORED – THE CFXP TOGGLE SWITCH (ORANGE), PVF MOTIF (GRAY), SODIUM POCKET (GREEN), WHERE THE  $Na^+$  IS SHOWN AS A GREEN SPHERE, DPXXY MOTIF (PINK) AND DRY MOTIF (CYAN). (B) SPATIAL POSITIONING OF THE HIS, ASP, AND GLU SIDECHAINS IN GPR68. THE FIGURE HAS BEEN REUSED [339]. ..... 98

FIGURE 37: DYNAMICS OF THE SODIUM ENTRY IN THE SODIUM-BINDING POCKET OF GPR68 WITH STANDARD PROTONATION STATES (SIM #5). (A) THE TRAJECTORY OF THE SODIUM ION ENTRY IN THE SODIUM POCKET OF GPR68 MODEL #5 IS SHOWN IN PURPLE. INTERNAL WATER MOLECULES ARE SHOWN WITHIN 3.5 Å OF THE  $Na^+$  (YELLOW SPHERE). TIME SERIES OF THE SODIUM ION ENTRY TO THE SODIUM POCKET, WITH THE INSET FIGURE HIGHLIGHTING THE FIRST 100 NS OF THE TRAJECTORY. THE DISTANCE (Å) REPORTED IS BETWEEN THE GEOMETRIC CENTER OF THE CA COORDINATES OF D67<sup>2.50</sup>, N104<sup>3.35</sup>, S108<sup>3.39</sup>, AND D282<sup>7.49</sup>. (C) CLOSE-UP SNAPSHOT OF THE ALLOSTERIC SODIUM ION POCKET ILLUSTRATING THE WATERS AND AMINO ACID SIDECHAINS NEAR THE SODIUM ION. THE FIGURE HAS BEEN REUSED [339]. ..... 100

FIGURE 38: DYNAMIC WATER MEDIATED PROTEIN-WATER H-BOND NETWORK OF GPR68. (A) H-BOND NETWORK OF GPR68. THE GRAPH NODES ARE ASSIGNED COLORS TO ENHANCE CLARITY AND VISUALIZATION – ASP/GLU SIDECHAINS (RED) HIS SIDECHAINS (BLUE), AMINO ACID SIDECHAINS THAT FORM THE SUGGESTED POSITIVE ALLOSTERIC BINDING POCKET (GREEN) AND OTHER H-BONDING SIDECHAINS (BLACK). NOTE THAT H269<sup>7.36</sup> AND E160<sup>4.64</sup> ARE ALSO A PART OF THE ALLOSTERIC BINDING SITE AND ARE THUS SHOWN IN TWO COLORS. FOR CLARITY, ONLY SELECTED NODES ARE LABELED. THE MINIMUM H-BOND OCCUPANCY SHOWN IS 30%. (B) HISTOGRAM OF THE AVERAGE OCCUPANCY OF ALL H-BOND CONNECTIONS. (C) TIME SERIES OF ALL UNIQUE H-BOND CONNECTIONS SAMPLED AT LEAST ONCE DURING THE TRAJECTORY SEGMENT USED FOR DATA ANALYSES. THE FIGURE HAS BEEN REUSED [339]. ..... 101

FIGURE 39: THE WATER-MEDIATED HYDROGEN BOND NETWORK OF GPR68. THE GRAPH COMPUTATION AND COLOR CODES ARE IDENTICAL TO THOSE DESCRIBED IN DETAIL FOR FIGURE 38. ONLY SPECIFIC NODES ARE LABELED FOR THE PURPOSE OF ENHANCING CLARITY. (A) THE HYDROGEN BOND CLUSTER OF E149 EXTRACTED FROM LARGER NETWORK OF GPR68. (B-J) LOCAL HYDROGEN BOND CLUSTERS THAT CONNECT SPECIFIC HISTIDINE, ASPARTIC ACID, AND GLUTAMIC ACID SIDE CHAINS, WHERE THE END NODES ARE HIGHLIGHTED IN BOLD. THE COLOR-CODED NUMBERS IN ITALICS REPRESENT THE JOIN OCCUPANCY (JO) VALUES OF THE RELEVANT HYDROGEN BOND PATHWAYS. THE FIGURE HAS BEEN REUSED <sup>[339]</sup>..... 102

FIGURE 40: PROTONATION-COUPLED DYNAMICS OF THE H-BOND NETWORK OF GPR68. (A-F) CORE NETWORK OF GPR68 WITH PROTONATED E149-OE1 (PANEL A), PROTONATED E149-OE2 (PANEL B), PROTONATED E103-OE1 (PANEL C), PROTONATED E103-OE2 (PANEL D), PROTONATED D282-OΔ1 (PANEL E), AND POSITIVELY CHARGED H269 (PANEL F). NOTE THAT THESE SIMULATIONS LACK TRANSIENT SODIUM ION BINDING AT THE SODIUM-BINDING SITE. THE FIGURE HAS BEEN REUSED <sup>[339]</sup>. ..... 104

FIGURE 41: PROTONATION OF ALL THREE SIDECHAINS OF THE INTERNAL CARBOXYLIC TRIAD ASSOCIATES WITH PERTURBED LOCAL INTERACTIONS. (A-D) CORE PROTEIN-WATER H-BOND NETWORK OF GPR68 WITH PROTONATED D67-OΔ2, E149-OE1, D282-OΔ2 (PANEL A), PROTONATED D67-OΔ2, E149-OE2, D282-OΔ2 (PANEL B), PROTONATED D67-OΔ2, E149-OE1, D282-OΔ2 (PANEL C). (D) DIRECT-H BONDING PARTNERS AND WATERS WITHIN 3.5 Å OF THE INTERNAL CARBOXYLATE TRIAD (D67, E149, D282), WHEN THE RECEPTOR HAS STANDARD PROTONATION AND WHEN THE TRIAD IS D67-OΔ2, E149-OE2, D282-OΔ2 PROTONATED. THE FIGURE HAS BEEN REUSED <sup>[339]</sup>. ..... 105

FIGURE 42: ORIENTATION OF CRUCIAL SELECT RESIDUES IN GPR68 MODELS PREPARED FROM DIFFERENT TEMPLATES (A) GPR68 MODEL BASED ON BOVINE RHODOPSIN TEMPLATE (PDB ID: 1F88). (B) GPR68 MODEL #5. THE FIGURE HAS BEEN REUSED AND ADAPTED <sup>[339]</sup> ..... 106

FIGURE 43: HOMOGENOUS TIME-RESOLVED FLUORESCENCE (HTRF) ASSAY FOR GA<sub>Q/11</sub> COUPLING OF GPR68. HTRF RATIO WAS MEASURED AT THE INDICATED PHS IN THE PRESENCE OF 50 MM LICI AND WAS LATER NORMALIZED FROM STANDARD CURVES TO DERIVE THE AMOUNT OF IP<sub>1</sub> ACCUMULATION. CELLULAR ASSAY WAS PERFORMED FOR NON-TRANSFECTED CELLS, WILD-TYPE (WT) GPR68, AND 16 MUTANTS. NOTE THAT THE MUTANT '5Q/N' REPRESENTS THE MUTANT – E160Q, E161Q, E164Q, D165N AND E166Q. THE BASAL ACTIVITY OF WT GPR68 HAS BEEN MARKED WITH A GRAY SOLID LINE. THE ENTIRE DATASET WAS NORMALIZED TO THE E<sub>MAX</sub> OF WT AS 100% AND BASELINE CORRECTED AGAINST NON-TRANSFECTED CELLS (CONTROL). THE DATA IS A COLLECTION OF THREE INDEPENDENTLY PERFORMED EXPERIMENTS. THE DATA IS REPRESENTED AS MEAN±S.E.M. THE PH VALUES AT WHICH HALF-MAXIMAL ACTIVATION IS OBSERVED (PH<sub>50</sub>) ARE REPORTED SEPARATELY IN TABLE 8. .... 107

FIGURE 44: EXTRACELLULAR AND TRANSMEMBRANE HISTIDINES OF GPR68 THAT WERE MUTATED IN THIS STUDY. (A) SNAKE DIAGRAM OF GPR68 HIGHLIGHTING THE MUTATED EXTRACELLULAR AND

TRANSMEMBRANE HISTIDINES. (B) STRUCTURAL MODEL OF GPR68 HIGHLIGHTING SPATIAL ORIENTATION OF EXTRACELLULAR AND TRANSMEMBRANE HISTIDINES. .... 108

FIGURE 45: BURIED CHARGED RESIDUES OF GPR68 THAT WERE MUTATED IN THIS STUDY. (A) SNAKE DIAGRAM OF GPR68 HIGHLIGHTING THE MUTATED BURIED CHARGED RESIDUES. (B) STRUCTURAL MODEL OF GPR68 HIGHLIGHTING SPATIAL ORIENTATION OF BURIED CHARGED RESIDUES FROM A CO-ORDINATE SNAPSHOT OF STANDARD PROTONATION SIMULATION, WHICH HIGHLIGHTS THE OUTWARD ORIENTATION OF E174<sup>45.52</sup> (C) STRUCTURAL MODEL OF GPR68 HIGHLIGHTING SPATIAL ORIENTATION OF BURIED CHARGED RESIDUES FROM A CO-ORDINATE SNAPSHOT OF PROTONATED SIMULATION OF E149-OE1, WHICH HIGHLIGHTS THE INWARD ORIENTATION OF E174<sup>45.52</sup>. .... 110

FIGURE 46: SURFACE CARBOXYLATES OF GPR68 THAT WERE MUTATED IN THIS STUDY. (A) SNAKE DIAGRAM OF GPR68 HIGHLIGHTING THE MUTATED SURFACE CARBOXYLATES. (B) STRUCTURAL MODEL OF GPR68 HIGHLIGHTING SPATIAL ORIENTATION OF SURFACE CARBOXYLATES. .... 111

FIGURE 47: OTHER RESIDUES WHICH WERE MUTATED IN GPR68 IN THE STUDY. (A) SNAKE DIAGRAM OF GPR68 HIGHLIGHTING THE MUTATED E55, N186, T233 AND S271 RESIDUES. (B) STRUCTURAL MODEL OF GPR68 HIGHLIGHTING SPATIAL ORIENTATION OF E55, N186, T233 AND S271 RESIDUE. .... 112

FIGURE 48: SUMMARY OF PROPOSED PROTON SENSING MECHANISM OF GPR68. THE FIGURE HAS BEEN REUSED FROM <sup>[339]</sup>. .... 121

FIGURE 49: SUMMARY OF THE FUNCTIONAL ANALYSIS OF GPR68 MUTANTS. (A) NORMALIZED EFFICACY MAXIMUM ( $E_{MAX}$ ) FOR GPR68 MUTANTS. THE DATA IS REPRESENTED AS MEAN FROM 3 INDIVIDUAL EXPERIMENTS WITH ERROR BARS REPRESENTING THE STANDARD ERROR. (B) MEAN  $PH_{50}$  VALUES FOR GPR68 MUTANTS SHOWN WITH ERROR BARS OF STANDARD ERROR. THE DOTTED LINE REPRESENTS THE  $PH_{50}$  VALUE OF WT-GPR68. THE SIGNIFICANCE VALUES HAVE BEEN MARKED AS ACCORDING TO THE P-VALUE SCALE: NS ( $P>0.05$ ), \* ( $P\leq 0.05$ ), \*\* ( $P\leq 0.01$ ), \*\*\* ( $P\leq 0.001$ ), AND \*\*\*\* ( $P\leq 0.0001$ ). IN CASES OF MUTANTS WHERE NO SYMBOL IS MARKED IMPLIES \*\*\*\* P-VALUE..... 124

## List of Supplementary Figures

- FIGURE S1: THE TIME SERIES OF PERCENTAGE A-HELICAL CONTENT FOR SIMULATIONS #1 - #8 WERE COMPUTED USING STRIDE <sup>[396]</sup>. ..... LXXIV
- FIGURE S2: THE PANELS A-H SHOW CA RMSD PROFILES COMPUTED FOR SIMULATIONS #1 - #8. FOR CLARITY, WE READ COORDINATES EACH 1 NS. .... LXXV
- FIGURE S3: NUMBER OF INTERNAL WATER MOLECULES THAT VISIT THE INTER-HELICAL REGION OF GPR68 STRUCTURAL MODELS DURING SIMULATIONS #1 - #8. .... LXXVI
- FIGURE S4: INTER-HELICAL DISTANCES COMPUTED FROM SIMULATIONS OF MODELS #1 - #8. (A-F) INTER-HELICAL DISTANCES AND CORRESPONDING STANDARD DEVIATIONS MEASURED AT THE EXTRACELLULAR TIPS OF TM HELICES (PANELS A, B), AT THE CENTRAL PLANE (PANELS C, D), AND AT THE CYTOPLASMIC ENDS OF THE TM HELICES (PANELS E, F). MEAN VALUES AND STANDARD DEVIATIONS WERE COMPUTED EVERY CO-ORDINATE SET FROM THE LAST 600NS OF EACH SIMULATION. .... LXXVII
- FIGURE S5: WATER MOLECULES MEDIATE THE DYNAMIC H-BOND NETWORKS OF GPR68. (A) MOLECULAR GRAPHICS ILLUSTRATING WATER MOLECULES WITHIN 3.5 Å OF THE AMINO ACID RESIDUES SHOWN IN PANELS B-J. (B-J) H-BOND CLUSTERS CORRESPONDING TO MAIN TEXT FIGURE 39 B-J, HERE SHOWN WITH THE AVERAGE NUMBER OF WATER MOLECULES FOR EACH EDGE REPRESENTING A DIRECT-H BOND OR A WATER-MEDIATED H-BOND. THE FIGURE HAS BEEN REUSED FROM (KAPUR ET AL., 2023). ..... LXXVIII
- FIGURE S6: CHARGED-PAIR INTERACTIONS OF E103<sup>3,34</sup>, E149<sup>4,53</sup>, E174<sup>45,52</sup>, K148<sup>4,52</sup>, AND R189<sup>5,42</sup>. (A-J) TIME SERIES FOR GPR68 WITH STANDARD PROTONATION STATES (PANEL A) AND WITH NON-STANDARD PROTONATIONS (B-J). CALCULATIONS WERE DONE WITH VMD USING DEFAULT SALT-BRIDGE CRITERIA. .... LXXIX
- FIGURE S7: IMPACT OF SCALING UP PURIFICATION FROM 500ML TO 2L USING A REPRESENTATIVE GPR68 CONSTRUCT, C33. SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (FLAG) FROM A 500 ML PELLETT PURIFICATION (PANEL A) AND 2 L PELLETT PURIFICATION (PANEL B) OF C33 PURIFIED USING A SUPERDEX 200 INCREASE 10/300 GL COLUMN IN THE PRESENCE OF DDM/CHS AT PH 6.0. ALL PURIFICATIONS WERE PERFORMED IN THE PRESENCE OF EITHER LIGAND 2 OR LIGAND 3. .... LXXX
- FIGURE S8: EFFECT OF PH ON GPR68 CONSTRUCT, C32. SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (TALON OR FLAG) FROM A 500 ML PELLETT PURIFICATION USING A SUPERDEX 200 INCREASE 10/300 GL COLUMN IN THE PRESENCE OF DDM/CHS AT PH 6.0 OR PH 7.8. .... LXXX
- FIGURE S9: EFFECT OF MUTATING THIRD PUTATIVE N-GLYCOSYLATION SITE ON GPR68 CONSTRUCT. SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (TALON) FROM A 500 ML PELLETT PURIFICATION USING A SUPERDEX 200 INCREASE 10/300 GL COLUMN IN THE

PRESENCE OF DDM/CHS AT PH 7.8 FOR C18 AND C19 (PANEL A). THE SDS PAGE GELS FROM THE TALON ELUATES FOR C18 (PANEL B) AND C19 (PANEL C)..... LXXX

FIGURE S10: MASS SPECTROMETRY, MALDI-TOF RESULTS SHOWING SUCCESSFUL IDENTIFICATION OF ANTI-BRIL FAB (PANEL A) AND ANTI-BRIL-FAB NANOBODY (PANEL B). THESE RESULTS ARE FROM THE CORRESPONDING SDS-PAGE GELS SHOWN FOR C61 PURIFICATION SHOWN IN FIGURE 32B. .... LXXXI

FIGURE S11: EFFECT OF VARIABLE SALT CONCENTRATION (NACL) ON GPR68 CONSTRUCT, C61, RANGING FROM NO SALT TO 1M NACL IN DDM/CHS (PANEL A), LMNG/CHS (PANEL B) AND GDN (PANEL C). SIZE EXCLUSION CHROMATOGRAMS FROM THE POOLED ELUATES AFTER AFFINITY PURIFICATION (TALON) FROM A 500 ML PELLETT PURIFICATION USING A SUPEROSE 6 INCREASE 5/150 GL COLUMN AT PH 7.4. .... LXXXI

FIGURE S12: MASS SPECTROMETRY OR MALDI-TOF RESULTS SHOWING SUCCESSFUL IDENTIFICATION OF ZEBRAFISH GPR68, C99 (PANEL A), ANTI-BRIL FAB (PANEL B) AND ANTI-BRIL-FAB NANOBODY (PANEL C). THE NANODSF PROFILE HIGHLIGHTS THE 350/330 NM RATIO AND THE FIRST DERIVATIVE OF THIS RATIO FOR C99 WHEN PURIFIED IN PRESENCE OF LIGAND 3 (PANEL D). THE MELTING TEMPERATURE FOR THIS CONSTRUCT WAS ABOUT ~60-62°C (PANEL D). THESE RESULTS ARE FROM THE CORRESPONDING SDS-PAGE GELS SHOWN FOR C99 PURIFICATION WITH FAB AND NB COMPLEX IN PRESENCE OF LIGAND 3 SHOWN IN FIGURE 34 B,C. .... LXXXII

FIGURE S13: LOCAL ALIGNMENT BETWEEN ZEBRAFISH GPR68 AND HUMAN GPR68 PERFORMED USING CLUSTAL OMEGA (VERSION 1.2.4) <sup>[397]</sup>. THE GREEN RECTANGLES HIGHLIGHT THE EXTRACELLULAR AND TRANSMEMBRANE HISTIDINES AND THE RED RECTANGLES HIGHLIGHT THE BURIED ACIDIC RESIDUES BASED ON HUMAN GPR68 SEQUENCE. .... LXXXIII

FIGURE S14: SELECT AMINO ACID RESIDUES WITH IMPLICATIONS IN CANCER-ASSOCIATED MUTATIONS OF GPR68. THE FIGURE SUMMARIZES THE RESULT OF OUR COSMIC DATABASE SEARCH ON 13.10.2022. THE UNDERLINED AMINO ACID RESIDUES ARE A PART OF THE DYNAMIC WATER-MEDIATED H-BOND NETWORK OF GPR68 (FIGURE 39). .... LXXXIV

## List of Tables

TABLE 1: STRUCTURAL MODELS OF GPR68 USED IN THIS STUDY. ....	61
TABLE 2: SUMMARY OF ALL MD SIMULATIONS PERFORMED IN THIS THESIS. THE DESIGNATION 'STANDARD' DENOTES THAT ALL TITRATABLE SIDECHAINS POSSESS STANDARD PROTONATION, AS PREVIOUSLY DESCRIBED. SIMULATIONS 10-18 INVOLVED THE INCLUSION OF TITRATABLE SIDECHAINS THAT EXHIBITED NON-STANDARD PROTONATION. SPECIFICALLY, THE PROTONATED OXYGEN ATOM OF THE NEUTRAL ASP/GLU IN THE SIMULATIONS 10, 11, 13-18, WHILE SIMULATION 12 HAD A POSITIVELY CHARGED HIS SIDECHAIN (H269). LIPID BILAYERS WERE COMPRISED OF POPC, EXCEPT FOR SIM #9, WHERE GPR68 WITH STANDARD PROTONATION WAS EMBEDDED WITHIN A POPC MEMBRANE SUPPLEMENTED WITH 10% CHOLESTEROL.....	64
TABLE 3: A SUMMARY OF CONSTRUCT FOR WHICH CHROMATOGRAMS ARE SHOWN AND DESCRIBED IN THIS SECTION. THE TRUNCATION OF AMINO ACIDS REPRESENTED WITH GREEK LETTER 'Δ' FOLLOWED BY NUMBER OF AMINO ACIDS. THE FUSION PROTEIN INSERTED IN AN ICL OF GPR68 IS SHOWN WITH THE INSERTION REGION (AMINO ACIDS BETWEEN WHICH FUSION PROTEIN IS INSERTED) SEPARATED BY A DOUBLE HYPHEN (--). THE FUSIONS ARE INSERTED IN ICL3 UNLESS MENTIONED OTHERWISE E.G., C73. NOTE THAT OTHER REDUNDANT DETAILS LIKE CLEAVAGE TAGS, LINKERS AND SIGNAL PEPTIDE ARE NOT DESCRIBED IN THIS TABLE. OTHER SPECIFIC MUTATION CONSTRUCT DETAILS HAVE BEEN DESCRIBED IN TABLE 4. THE ABBREVIATIONS USED AND THE OTHER CONSTRUCT DETAILS HAS BEEN DESCRIBED IN APPENDIX 1.....	74
TABLE 4:COMPILATION OF POTENTIAL THERMOSTABILIZING MUTATIONS FOR HUMAN GPR68. ....	83
TABLE 5: SUMMARY OF THE VARIATIONS ATTEMPTED FOR THE CRYO-EM TRIALS DURING THE WHOLE DURATION OF MY THESIS. ....	93
TABLE 6: SEQUENCE IDENTITY AND SIMILARITY BETWEEN PROTON SENSING GPCRS. THE VALUES WERE COMPILED BASED ON THE INFORMATION FROM GPCRDB. THE TABLE HAS BEEN REUSED <sup>[339]</sup> .....	97
TABLE 7: SUMMARY OF STRUCTURAL PARAMETERS USED TO EVALUATE THE CONFORMATIONAL DYNAMICS OF THE GPR68 HOMOLOGY MODELS WITH STANDARD PROTONATION STATES FOR ALL TITRATABLE SIDECHAINS. WE REPORT, FOR EACH MODEL, THE AVERAGE RMSD VALUES OF THE TM HELICES COMPUTED EVERY 10 PS FROM THE LAST 600 NS OF EACH SIMULATION. THE TABLE HAS BEEN REUSED <sup>[339]</sup> .....	99
TABLE 8:SUMMARY OF ALL THE PH <sub>50</sub> VALUES FOR WT-GPR68 AND MUTANTS TESTED VIA HTRF ASSAY FOR GA <sub>Q/11</sub> SIGNALING PATHWAY. THE SIGNIFICANCE OF THE MEAN PH <sub>50</sub> VALUES WAS EVALUATED USING ONE-WAY ANOVA ANALYSIS IN GRAPHPAD PRISM 10. THE WORDS N.D. AND NS STAND FOR 'NOT DEFINED' AND 'NOT SIGNIFICANT'. THE P-VALUE SCALE: NS (P>0.05), * (P≤0.05), ** (P≤0.01), *** (P≤0.001), AND **** (P≤0.0001). ....	108
TABLE 9: OVERVIEW OF THE EXPERIMENTS OR SIMULATIONS PERFORMED DURING MY DOCTORAL THESIS. ....	114

## List of Supplementary Tables

TABLE S1: JOINT OCCUPANCY VALUES OF H-BOND PATHS SHOWN IN FIGURE 39 OF THE MAIN TEXT. LXXXVI

## **Eidesstattliche Erklärung / Declaration**

Hiermit versichere ich an Eides statt, dass ich die Arbeit, abgesehen von der Beratung durch meinen Betreuer, inhaltlich und formal selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Diese Arbeit ist weder ganz noch teilweise in ein anderes Prüfungsverfahren eingebracht worden und wurde nach den Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft erstellt. Ich versichere ferner, dass ich an keiner anderen Hochschule ein Promotionsvorhaben nicht bestanden habe und mir nie ein akademischer Grad entzogen wurde.

Übersetzt mit DeepL.com (kostenlose Version)

I hereby confirm under an oath that apart from the advice of my supervisor, I have written the thesis independently in terms of its content and form and have not used any sources or aids other than those specified. This thesis has not been submitted in whole or in part for an examination procedure elsewhere and has been prepared in accordance with the rules of good scientific practice of the German Research Foundation. I also assure you that I have not failed a doctoral project at any other higher education institution, and no academic degree has ever been withdrawn from me.

**Bhav Kapur**

*Biberach an der Riß, May 2024*