ESTABLISHING A BIOPHYSICAL ASSAY FOR THE INTERACTIONS OF CALCITONIN FAMILY G PROTEIN-COUPLED RECEPTORS WITH RECEPTOR ACTIVITY MODIFYING PROTEINS

AND

THE ELUCIDATION OF THE STRUCTURE-BASED SIGNALLING OF G PROTEIN-COUPLED RECEPTORS VIA THE NOVEL APPLICATION OF GEOMETRIC MORPHOMETRICS WITH PRINCIPAL COMPONENT ANALYSIS

DANIEL NICHOLAS WISEMAN

Doctor of Philosophy

ASTON UNIVERSITY

March 2024

Aston University

Establishing a Biophysical Assay for the Interactions of Calcitonin Family G Protein-Coupled Receptors with Receptor Activity Modifying Proteins

And

The Elucidation of the Structure-Based Signalling of G Protein-Coupled Receptors via the Novel Application of Geometric Morphometrics with Principal Component Analysis

Daniel Nicholas Wiseman

Doctor of Philosophy

2024

Abstract

GPCRs are implicated in a wide variety of diseases, making them particularly attractive drug targets; critical developments in structural and functional techniques have led to breakthrough discoveries, shifting paradigms of understanding with novel concepts. Select ligands are now known to induce biased signalling in GPCRs, differentially activating intracellular signalling pathways; conformational landscapes of GPCRs are heterogeneous, indicating diversity in the activation transition and intermediate states which likely correlate with biased signalling. Moreover, the majority of GPCRs signal via more than one G protein sub-type, preferentially coupling with ranked selectivity, and are allosterically modulated by a range of factors. In order to develop safe, effective and selective therapeutics against GPCRs, a holistic understanding of these concepts necessitates comprehensive, multidisciplinary approaches, combining enhanced biophysical and biochemical assays, determination and analyses of structure, and complementary computational techniques. This thesis presents the successful establishment of a FRET-based interaction assay for the calcitonin family of GPCRs and the RAMPs, with an average FRET efficiency of 87.78%, average interaction distance of 3.42 nm, and cAMP pEC50s in agreement with the literature. This promising assay will offer novel insights into GPCR-RAMP dynamics, forming the basis of a high-throughput biophysical drug discovery platform. Secondly, the structure-based signalling of GPCRs was explored with the novel application of geometric morphometrics and principal component analysis to resolved structures, consistently and reliably classifying GPCRs by their global shape morphology, supported by PERMANOVA and ANOSIM multivariate statistics. Case study examples of the β2-adrenergic, adenosine 2A, secretin-like and calcium-sensing receptors first proved this concept effective, before exploration of thermostabilisation, fusion proteins, the structural determinants of G protein coupling, and AlphaFold structures. Overall, this thesis provides novel contributions to fully elucidating the structure/function relationship of GPCRs, building a multidisciplinary model of understanding, which will enable the unprecedented discovery and development of safe, effective therapeutics.

Keywords: G protein-coupled receptor (GPCR), receptor activity modifying protein (RAMP), Förster resonant energy transfer (FRET), geometric morphometrics (GM), principal component analysis (PCA).

I would like to dedicate this work to my mother Annabel Wiseman, and grandmother Judith Wiseman, for their unending support and love.

Acknowledgements

I would like to enormously thank my supervisor Professor Alan Goddard, and co-supervisors Dr John Simms and Professor David Poyner for all of their time, support and guidance throughout the course of this degree. I would also like to thank Dr Cathy Slack for guidance and training in the use of *Drosophila melanogaster*. I would also like to thank my line manager Dr Joanne Gough, and head of school Professor Andrew Devitt, for support in managing my time between teaching and research.

In addition, I would also like to thank past and present members of Aston University's AMPL research group, specifically Dr Sarah Routledge for enlightening conversations and training, and Dr Monserrat Roman Lara for contributions to preliminary data collection. Furthermore, I would like to thank placement students Georgia Stevens, Sophia Ijaz, and Nikita Samra for their contributions to various projects.

Finally, I would also like to thank Aston University's College of Health and Life Sciences for funding and supporting this project.

List of Abbreviations

A2AR – adenosine A2A receptor

Ago-NAM-PAM – treble agonist, negative and positive allosteric modulator bound

Ago-PAM – dual agonist and positive allosteric modulator bound

AM – adrenomedullin

AMY – amylin

ANOSIM – one-way analysis of similarity

β1AR – β1-adrenergic receptor

β2AR – β2-adrenergic receptor

Cα – alpha carbon atom

cAMP – cyclic adenosine monophosphate

CASP – critical assessment of protein structure prediction

CaSR – calcium sensing receptor

CFP – cyan fluorescent protein

CGRP(R) – calcitonin gene related peptide (receptor)

CLR – calcitonin receptor-like receptor

CRD – cysteine-rich domain

Cryo-EM – cryogenic electron microscopy

CT(R) – calcitonin (receptor)

CyO – curly derivative of Oster strain

CXCR4 – chemokine receptor type 4

D3R – dopamine receptor D3

EC50 – half-maximal effective concentration

ECD – extracellular domain

ECL – extracellular loop

ERK – extracellular signal-regulated kinases

FLAG – artificial tag DYKDDDDK

FRET – Förster resonance energy transfer

GEF – guanine nucleotide exchange factor

GM – geometric morphometrics

GMR – glass multiple reporter

GST – glutathione s-transferase

Gt – transducin

GPCR – G protein-coupled receptor

HA – haemagglutinin tag YPYDVPDYA

H1R – histamine H1 receptor

IC50 – half-maximal inhibitory concentration

ICL – intracellular loop

M2R – muscarinic acetylcholine receptor M2

M3R – muscarinic acetylcholine receptor M3

ML – machine learning

NAM – negative allosteric modulator

NMR – nuclear magnetic resonance spectroscopy

PAM – positive allosteric modulator

PC(A) – principal component (analysis)

PDB – protein data bank

PEI – polyethylenimine

PERMANOVA – one-way permutational multivariate analysis of variance

RAMP – receptor activity modifying protein

RMSD – root mean square deviation

SMALP – styrene maleic acid lipid particle

StaR® – stabilised receptor technology, Sosei Heptares

TM(D) – transmembrane (domain)

UAS – upstream activation sequence

VFT – venus flytrap module

WFV – white female virgin

YFP – yellow fluorescent protein

Chapter 1: General Introduction

1.1 – The Biological Cell and its Plasma Membrane

1.1.1 – The Fluid Mosaic Model

The Cell Theory was initially proposed in *Micrographia* by Robert Hooke in 1665, helping to lay the preliminary foundations of microscopic biology (Gest, 2004; Hooke, 1665). As technology advanced further into the 1890s, an update to the Cell Theory was published to include the existence of a cell membrane composed of lipids, reviewed in Lombard, 2014. Its importance was not immediately recognised but led to the lipid bilayer hypothesis in 1925, and the subsequent fluid mosaic model in 1972 by Singer and Nicolson (Gorter and Grendel, 1925; Singer and Nicolson, 1972). It is this fluid mosaic model which remains the principal explanation of the cell membrane – contemporary discoveries continuously modernise this theory, however, its basics remain relatively constant (Nicolson and Ferreira de Mattos, 2021, 2022).

According to the fluid mosaic model, illustrated in Figure 1.1.1, the cell membrane is composed of lipid, protein, and carbohydrate macromolecules, forming the cell's protective barrier (Singer and Nicolson, 1972). The primary lipid components called phospholipids are composed of a hydrophilic head and hydrophobic tail, spontaneously forming an amphipathic bilayer in water. Additionally, cholesterol and other sterols can be interspersed between these phospholipids, regulating the fluidity of the bilayer (Casares, Escribá and Rosselló, 2019; Ying Zhang *et al.*, 2020). As protein molecules are also amphipathic in nature, some can integrate with the bilayer intrinsically, whereas others associate with the membrane peripherally. Furthermore, carbohydrates can be linked to both lipids and proteins to form glycolipids and glycoproteins, respectively (Moran *et al.*, 2014). In summary, this lipid-protein mixture forms a 'fluid', Brownian system whereby these molecules can disperse and concentrate by lateral diffusion, expanding the functions of the cell membrane beyond being a mere barrier (Sonnleitner, Schütz and Schmidt, 1999; Ramadurai *et al.*, 2009).

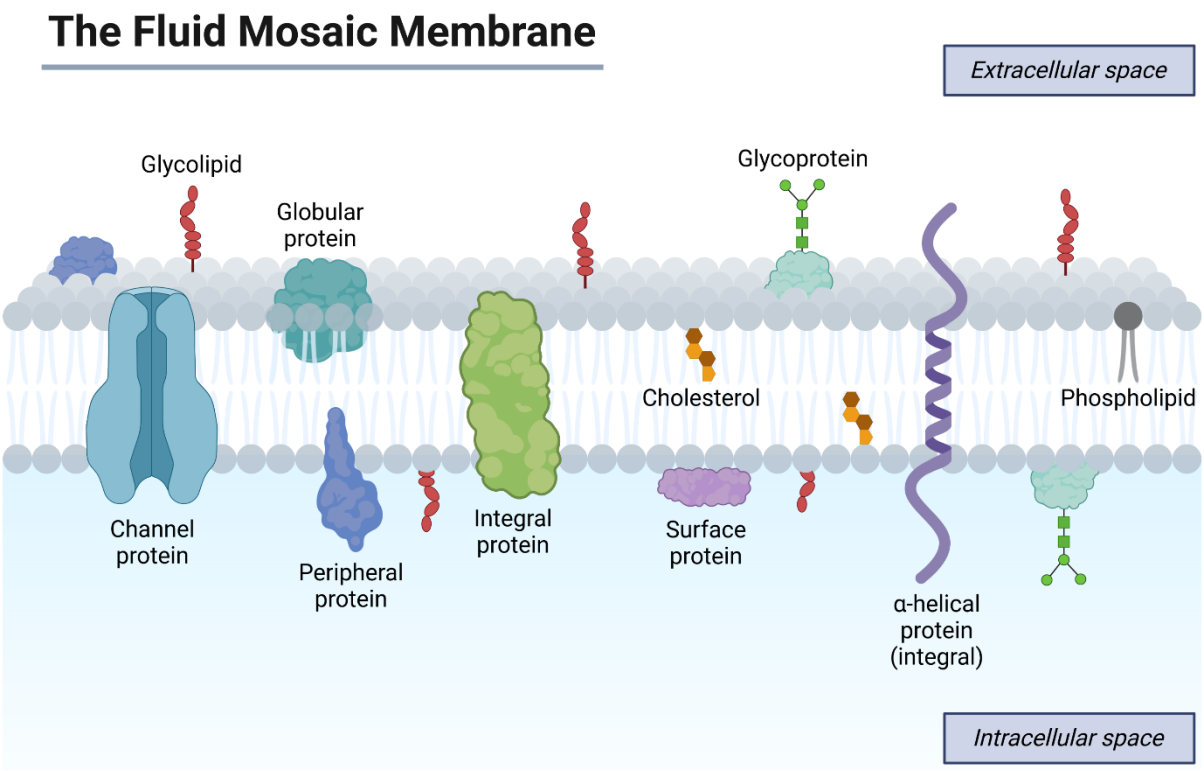

Figure 1.1.1. The Fluid Mosaic Model of the Plasma Membrane. The cell membrane consists of the carbohydrate, lipid and protein macromolecules (approximately 50% lipid, 50% protein by mass) to facilitate and enhance the many functions of the cell's protective barrier. Created with BioRender.com.

The fluid mosaic model is not only widely accepted due to historical experiments, but because the fluid mosaic nature of cell membranes is also supported by modern advancements in fluorescence microscopy and structural biology; the fluid mosaic model therefore continues to be the most widely accepted model as reviewed in Nicolson and Ferreira de Mattos, 2022 (Nicolson and Ferreira de Mattos, 2022). As such, the cell membrane is now understood to be a complex, dynamic system with varied characteristics for consideration, including membrane asymmetry (Scott *et al.*, 2021), curvature (Has, Sivadas and Das, 2022), fluidity (Lipowsky, 2014), lipid rafts (Regen, 2020), and non-bilayer membranes (Verkleij *et al.*, 1984), etc.

Through a combination of passive and active processes, the cell membrane provides selective permeability to regulate the bidirectional transport of material between the cytoplasm and extracellular environment (Yang and Hinner, 2015). For example, some small and uncharged molecules, such as O_2, CO_2, H_2O, and ethanol, can passively move across the membrane by diffusion or osmosis, due to concentration gradients (Cooper, 2000). However, most other molecules require movement facilitated by specific transport or channel proteins embedded within the membrane. Alongside its role in material transport, the cell membrane is also essential to a variety of other processes including the anchoring of the cytoskeleton, interaction with the extracellular matrix, and cell signalling and communication (Nicolson and

Ferreira de Mattos, 2022). Much of these are facilitated by the proteins found within the cell membrane itself.

1.1.2 – The Role of Membrane Proteins

Membrane proteins account for approximately 30% of the human genome and act as a core component of the cell's regulatory and sensory recognition systems; this is achieved by the maintenance of critical ion concentrations and the propagation of signals towards intracellular cascades, for example (Kelly *et al.*, 2009; Fagerberg *et al.*, 2010). Consequently, due to their great importance to the breadth of human health and disease, membrane proteins are the target of over 50% of all drugs (Overington, Al-Lazikani and Hopkins, 2006; Martin and Sawyer, 2019). Despite their importance however, their amphipathic nature and largely hydrophobic transmembrane regions make membrane protein structures typically difficult to elucidate. As such, considerable effort and technological advancement have enabled improved sequence, structural and functional analyses of membrane proteins, furthering fundamental knowledge and application to novel drug design (Kelly *et al.*, 2009; Marinko *et al.*, 2019).

Broadly, there are two categories of membrane protein, as illustrated in Figure 1.1.2: peripheral and integral. Peripheral membrane proteins transiently interact with the lipid bilayer via hydrophobic, electrostatic, and other non-covalent methods (Sahin *et al.*, 2020). As these interactions are transient, temporally relevant proteins such as regulatory subunits and metabolic enzymes are often classed as peripheral membrane proteins, as can G proteins and protein kinases. The second category, integral membrane proteins, are permanently associated with the membrane, either by partially or completely spanning the depth of the bilayer (Bracey, Cravatt and Stevens, 2004). Monotopic proteins such as cyclooxygenases only associate with one side of the membrane, bitopic proteins such as the tumour necrosis factor receptor span the membrane only once, while polytopic proteins span the membrane several times (Garavito and Mulichak, 2003; Tan, Tan and Chung, 2008). The bitopic and polytopic integral proteins, or transmembrane proteins, comprise the majority of this group and give the cell membrane many of its unique functions.

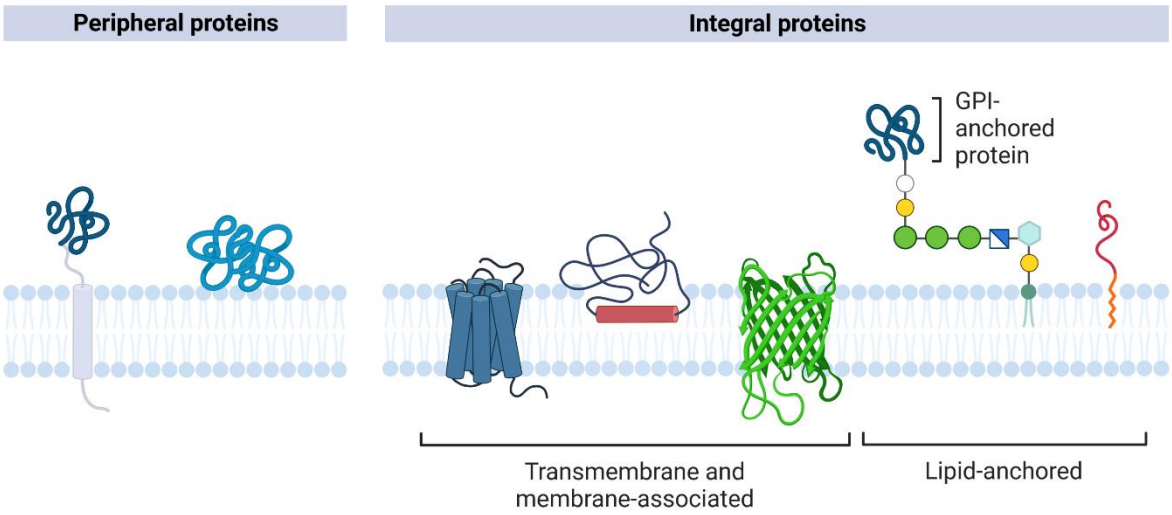

Figure 1.1.2. Peripheral and Integral Membrane Proteins. Peripheral membrane proteins interact with the lipid bilayer transiently, whereas integral membrane proteins associate permanently with the bilayer, spanning some or all of its depth. Created with BioRender.com.

Human transmembrane proteins conform to an alpha-helical structure which contrasts with the topology of the beta-barrel proteins mainly found in the outer membranes of Gram-negative bacteria, chloroplasts, and mitochondria (Taylor *et al.*, 2006). Alpha-helical transmembrane proteins make use of the internal hydrogen bonding of the helical structure, and their helices usually consist of 18 to 30 hydrophobic amino acids flanked by polar residues (Engelman, Steitz and Goldman, 1986; Almén *et al.*, 2009). This makes alpha-helices particularly thermodynamically stable as polar residues interact with water molecules and lipid head groups to stabilise the protein and prevent unfolding within the membrane (von Heijne, 1992). Note that much longer helices do exist, and they often tilt in order to keep the hydrophobic core of the helix within the hydrocarbon region of the membrane (Gofman, Haliloglu and Ben-Tal, 2012). Due to the energetic cost of topological rearrangement, importantly, the stable state of most transmembrane proteins must be achieved alongside translation with aid of the Sec translocon complex (Becker *et al.*, 2009). Once inserted into the bilayer, transmembrane proteins are then subject to the physicochemical properties of their local microenvironment including direct contact with annular lipids, helical tilting and modulation of the membrane thickness (Marinko *et al.*, 2019). Ultimately, the ability of hydrophobic polypeptides to form alpha-helices within lipid bilayers is thought to have been a driving force behind the diverse evolution of transmembrane proteins (Popot and Engelman, 2000; Almén *et al.*, 2009).

As mentioned, transmembrane proteins account for around 30% of the human genome, which corresponds to 6 to 8 thousand members of the human transmembrane proteome (Dobson, Reményi and Tusnády, 2015). Further analyses and deorphanisation of these proteins allow

for in-depth classification into the major membrane protein groups, many of which are fundamentally essential to eukaryotic life, but also of great medical relevance. A few examples of transmembrane proteins will be briefly summarised, highlighting their importance and the need to resolve and study their structures. G protein-coupled receptors (GPCRs) constitute around 15% of all human membrane proteins, are the targets of around 40% of all drugs, and are covered in greater depth in sections 1.2 and 1.3 (von Heijne, 2007).

Aquaporins

Aquaporins are water channels which facilitate the bi-directional movement of water across the cell membrane, down an osmotic gradient; in some cases, they are also able to move other small molecules such as ammonia, CO_2, glycerol, and urea, as well as water (Kruse, Uehlein and Kaldenhoff, 2006; Bill and Hedfalk, 2021). Humans express thirteen aquaporins which are involved in various fluid secretions including bile, cerebrospinal fluid, saliva, sweat, tears and the concentration of urine (Ishibashi, Hara and Kondo, 2009). Consequently, the human aquaporins play a role in equally diverse disease states such as Alzheimer's, brain oedema following traumatic injury or stroke, kidney disease, and Parkinson's, amongst others (Bill and Hedfalk, 2021). Following publication of high-resolution aquaporin structures, their mechanism of action was revealed at the atomic level; they possess six transmembrane helices which form an 'hourglass' model, and each aquaporin assembles as a homotetramer retaining four individual water channels (Kreida and Törnroth-Horsefield, 2015). Water molecules are able to move rapidly through these channels, in such a way that the electrostatic field within the channels does not allow leakage of protons or other ions across the membrane (Fujiyoshi *et al.*, 2002). Finally, as the influence of gene expression is too slow in the context of rapid water flow, research into acute therapeutic mechanisms is focussed upon temporarily modifying the number of available channels at the cell membrane by reversable blockage with inhibitors, and targeting the dynamic translocation of aquaporins between sub-cellular compartments and the cell surface as reviewed in Salman *et al.*, 2022 (Salman *et al.*, 2022).

Ion Channels

Ion channels are transmembrane proteins which regulate cell membrane potentials through the movement of charged ions, such as calcium, chloride, potassium, and sodium, between the cytosol and extracellular space (Rosendo-Pineda, Moreno and Vaca, 2020). They additionally facilitate the movement of ions within intracellular compartments, such as the release of Ca^{2+} from the endoplasmic reticulum as second messengers to activate calmodulin (Taylor and Laude, 2002). These ions move passively through the channel's selective pore, down an electrochemical gradient, which can be one thousand times the rate of other transport

proteins (Kulbacka *et al.*, 2017). Furthermore, the superfamily of ion channels has several mechanisms of activation, including ligand and voltage gating, mechanical stress, and temperature, amongst others, and play essential roles in action potentials, cellular communication, muscle contraction, and cell proliferation to name a few (Kulbacka *et al.*, 2017).

One prominent ion channel is the bacterial potassium ion channel, KcsA; the upper region of this transmembrane protein hosts a selectivity filter which, as the name suggests, imparts K^+ ion selectivity, while the lower region of KcsA controls gating via sensor regions (Kuang, Purhonen and Hebert, 2015). K^+ ions move through this channel due to the presence of oxygen atoms in the selectivity filter which mimic the positions of water oxygens when K^+ is in solution; while Na^+ ions are smaller than K^+ ions, only one in every thousand make it through KcsA, maintaining a high degree of K^+ selectivity (von Heijne, 2007). Importantly, KcsA is thought to function in a similar manner to the human Kv channels, confirmed by functional chimeras between the two (Williamson *et al.*, 2003). In contrast, another ion channel, the nicotinic acetylcholine receptor, is composed of a large funnel which undergoes a conformational change upon ligand binding, and is characteristically found at the post-synaptic neuromuscular junction (Changeux, 2018). The rotation of helices which line the pore of this channel allow cations to move through in a non-selective manner, with Na^+ generally moving into the cell, and K^+ leaving the cell. Finally, as additional ion channel structures are resolved, therapeutic-focussed research can target their relationships with human disease, such as the recent association of the mechanosensitive PIEZO channels to various cancers and more (Yu and Liao, 2021; Delmas, Parpaite and Coste, 2022).

ABC Transporters

Around 50 human ATP-binding cassette, or ABC transporters, have been identified thus far; by utilising the energy from ATP hydrolysis, these transmembrane proteins export and sometimes import small-molecule substrates across the cell membrane (Liu, 2019). Examples include amino acids, peptides, sugars, and others including drugs; indeed, at least eleven ABC transporters including the breast cancer resistance protein (BCRP/*ABCG2*), multidrug resistance-associated proteins (MRPs/*ABCCs*), and P-glycoprotein (P-GP/*ABCB1*) are involved in the development of multidrug resistance (Slot, Molinski and Cole, 2011; Schinkel and Jonker, 2012). Of course, this is clinically important, especially as ABC transporters are expressed in various tissues such as the brain, intestines, kidneys and liver, and mutations in ABCs have been linked to cystic fibrosis, hepatic cholestasis, and retinal degradation, to name a few conditions (Gadsby, Vergani and Csanády, 2006; Tarling, Vallim and Edwards, 2013; Liu, 2019).

Functional ABC transporters are generally composed of two nucleotide-binding domains (NBDs), and two transmembrane domains (TMDs) consisting of six α-helices each (Liu, 2019). While the NBDs seem to be highly conserved throughout the ABC transporters, the TMDs provide substrate specificity. Two molecules of ATP bind to the NBDs, moving them close together which completes the interface for ATP hydrolysis (George and Jones, 2012). This induces a conformational change to switch the substrate binding cavity from facing the extra- to the intracellular space, ready to accept its substrate from the inner bilayer leaflet or cytoplasm, and then back again to release substrate from the cell. This mechanism is therefore one of alternating access and release, though the intricate details of all ABC transporters are not yet fully understood (von Heijne, 2007; Liu, 2019).

From these examples given, transmembrane proteins are governed by two overarching principles which seem to be essential to their structure/function relationship; access to both sides of the cell membrane, and the ability to undergo conformational changes upon activation. In any case, in order to better understand human membrane proteins as individual molecules and components of cellular mechanisms, they must be expressed, solubilised, purified and have their structures resolved to acceptable resolution, which is no simple task.

1.1.3 – Expression and Purification Strategies for Membrane Proteins

The characterisation of membrane protein structural dynamics is of great importance, not only to provide insight into mechanisms of action, but to additionally aid novel structure-based drug design. Given the importance and prevalence of membrane proteins, it may be surprising to realise that their experimentally resolved structures only account for 3.8% of total proteins in the PDB (7,414 out of 197,848 as of November 2022); this number has doubled since 2016 but still pales in comparison to soluble protein structures (Pandey *et al.*, 2016). For GPCRs specifically, the GPCRdb lists 793 structures at the time of writing, of which 423 are human receptors (Kooistra *et al.*, 2021). Much of this challenge arises from the hydrophobicity of GPCRs, and their relatively low levels of endogenous expression; indeed, recombinant expression can be highly variable, even within the same expression system. Structures resolved before the year 2000 often relied upon proteins found in natural abundance, limiting the pool of feasible structural studies; early examples include bacteriorhodopsin (Henderson and Unwin, 1975), aquaporin channels (Cheng *et al.*, 1997), and the G protein-coupled receptor rhodopsin (Palczewski *et al.*, 2000). These experimental bottlenecks have thus fuelled the structural biology field over the last two decades, driving advancements in the methods and technologies used to express, solubilise and purify membrane proteins for structure determination.

Typically, milligram quantities of sample are required for structural studies, and thus heterologous expression of recombinant proteins has become the dominant method, accounting for the source of approximately 70% of resolved membrane protein structures (Bill *et al.*, 2011; Dilworth *et al.*, 2018). A variety of expression hosts are also now available to exploit, each with their advantages and disadvantages, as summarised in Table 1.1.3, which is a qualitative representation of evidence-based opinion; several factors must be considered including expense, toxicity of the target protein when overexpressed, and the ability for post-translational modifications to occur, for example. As such, the choice of host remains to be target protein specific, as there is no standardised approach which is deemed best for all (Grisshammer, 2006; Bernaudat *et al.*, 2011; Thomas and Tate, 2014; Wiseman *et al.*, 2020). Indeed, these experimental choices may impact on the likeness of the produced protein to its wildtype, and care must therefore be taken during these steps to prevent or at least understand any unwanted differences in structural similarity.

Table 1.1.3. A Comparison of Cellular Membrane Protein Expression Systems. Each consideration linked to cellular recombinant expression systems is conveyed as a qualitative summary, linking experimental evidence and personal opinion. Green = positive, amber = moderate, red = negative, relatively. *E. coli* and yeast were historically favoured due to their ease of culture, genetic flexibility and scalability, however, expression lines such as expiCHO, expi293 and S*f*9 can now be grown in litre volumes, producing milligram quantities of protein. *Drosophila* offer a promising compromise but are currently underused in comparison.

	Ease of Culture	Ease of Genetic Manipulation	Expense	Membrane Composition	Post-translational Modification	Scalability
E. coli						
Yeast						
Insect						
Mammalian						
Drosophila						

Escherichia coli

Historically, the expression of membrane proteins in eukaryotes was typically difficult, and thus *Escherichia coli* became a staple host for their expression and purification, for a number of reasons (de Mendoza, Sebé-Pedrós and Ruiz-Trillo, 2014; Hattab *et al.*, 2015; Z. Zhang *et al.*, 2015). *E. coli* is easily cultured, inexpensive, and decades of work have resulted in well-characterised strains which grow rapidly, and are optimised for protein expression, including membrane proteins specifically (Miroux and Walker, 1996; Depping *et al.*, 2022). Transformation of *E. coli* entails relatively simple protocols, and there is an abundance of expression vectors to choose from with a variety of modifiable properties. Indeed, tuning

protein expression can be of great importance in reducing rate-limiting steps such as saturation of the translocon (Wagner *et al.*, 2007, 2008). Additionally, T7 RNA polymerase from the *Enterobacteria* phage λ often drives recombinant expression, which gives a faster and higher yield when compared to native *E. coli* RNA polymerase (Studier, 1991; Iost, Guillerez and Dreyfus, 1992).

Despite these attractive qualities however, there are also significant limitations; as a prokaryotic organism, *E. coli* lacks features which can be essential for the production of high-quality, functional GPCRs (Mathieu *et al.*, 2019). Formation of disulphide bonds are inefficient at best, proteins may not fold correctly, the lipid composition of the membrane differs from eukaryotic cells, and there is a distinct lack of post-translational modification, including glycosylation (Wheatley and Hawtin, 1999; Baneres *et al.*, 2003; Wagner *et al.*, 2007). However, despite these clear limitations, several active GPCRs including cannabinoid, neurotensin and serotonin receptors have been successfully expressed in *E. coli* (Bertin *et al.*, 1992; Grisshammer, Duckworth and Henderson, 1993; Krepkiy *et al.*, 2006; Michalke *et al.*, 2010; Mallipeddi, Zvonok and Makriyannis, 2018). The genetic tractability of *E. coli* means key modifications such as the use of fusion partners and stabilising mutations can be performed with relative ease, and downstream applications such as isotopic labelling for NMR are also relatively easy to conduct (Marley, Lu and Bracken, 2001; Attrill *et al.*, 2009; Petrovskaya *et al.*, 2010; Tapaneeyakorn *et al.*, 2011; Hattab *et al.*, 2015; Mallipeddi, Zvonok and Makriyannis, 2018). In the more recent years since 2010, the use of *E. coli* has dwindled, and eukaryotic expression systems have become ever more favoured (Kesidis *et al.*, 2020).

Yeasts

Next in popularity are yeasts which exhibit many of the advantages of *E. coli*, but with the additional ability to perform some eukaryotic-like post-translational modifications (Carlesso *et al.*, 2022). Yeasts are similarly quick to grow to high density, relatively inexpensive, well characterised, and amenable to genetic manipulation. In addition, yeasts can promote the formation of disulphide bonds in proteins, process signal peptide sequences and perform O- and N-linked glycosylation, all of which may be essential for the correct folding, trafficking and activity of target membrane proteins (Böer *et al.*, 2007). Note that while the early glycosylation process is similar to that in mammalian cells, yeasts are prone to hypermannosylate proteins, leading to potential misfolding (Choi *et al.*, 2003; Vieira Gomes *et al.*, 2018). Another further consideration relates to the replacement of cholesterol with ergosterol in the yeast membrane composition; cholesterol is required for the correct function of some GPCRs, and may also contribute directly to their conformational changes or activity (Hanson *et al.*, 2008; Goddard and Watts, 2012; Goddard *et al.*, 2013). One such example is the μ-opioid receptor which is

constrained to an inactive state in the presence of ergosterol, but active with cholesterol (Lagane *et al.*, 2000). That being said, yeast has been shown to produce humanised sterol compositions, including the production of cholesterol instead of ergosterol, when reprogrammed genetically (Hirz *et al.*, 2013; Emmerstorfer-Augustin *et al.*, 2019).

The baker's yeast *Saccharomyces cerevisiae* and methylotrophic yeast *Pichia pastoris* are the most commonly used expression species; the fission yeast *Schizosaccharomyces pombe* is used less frequently but can outperform other species under certain conditions (Takegawa *et al.*, 2009; Joubert *et al.*, 2010; Byrne, 2015). Similar to *E. coli*, *S. cerevisiae* has a long history of study as a model organism, with many strains available for genetic manipulation, and several examples of its involvement in the production of membrane proteins for structural and functional studies (Newstead *et al.*, 2007; Li *et al.*, 2009). However, despite these successes, *P. pastoris* has become the most frequently and successfully utilised species of yeast (Ramón and Marín, 2011; Ayub *et al.*, 2022). This is undoubtedly due to the ease of integrating stable expression vectors into *P. pastoris*, and its ability to consume methanol as the sole source of carbon which enables high levels of expression under strict control of the *AOX1* promoter (Hollenberg and Gellissen, 1997; Ahmad *et al.*, 2014). *P. pastoris* strains have since been engineered to perform human N-glycosylation, as well as the ability to synthesise cholesterol, which may ultimately benefit GPCR expression (Jacobs *et al.*, 2009; Hirz *et al.*, 2013). Overall, yeasts offer an advantageous expression system, spanning the gap between prokaryotic and eukaryotic approaches, and have contributed significantly to structural and functional studies of GPCRs thus far including the adenosine 2a, β-adrenergic and muscarinic acetylcholine receptors (Weiss *et al.*, 1998; Singh *et al.*, 2008; Asada *et al.*, 2011; Routledge *et al.*, 2016; Carlesso *et al.*, 2022).

Insect Cells

The most commonly used method for GPCR expression, is the advantageous use of the baculovirus system with insect cells (Saarenpää, Jaakola and Goldman, 2015). This is not only a very safe method (as the baculovirus does not infect mammalian cells) but also generally allows proper folding of membrane proteins, due to post-translational modifications which are often almost identical to those in mammalian cells (Milić and Veprintsev, 2015; Vaitsopoulou *et al.*, 2022). The commonly used insect cell lines are S*f*9 and S*f*21 from the *Spodoptera frugiperda* moth, and the Hi5 line from the *Trichoplusia ni* moth (Saarenpää, Jaakola and Goldman, 2015). Expression is achieved by infection with a recombinant *Autographa californica* multi-nucleopolyhedrovirus which drives target protein expression via the polyhedrin promoter (McKenzie and Abbott, 2018). This system not only offers scalable cultures able to grow in serum-free shakers, but can achieve milligram quantities of target

protein following optimisation (Unger and Peleg, 2012; Chen *et al.*, 2013). However, the use of insect cells can be a relatively expensive approach when compared to *E. coli* and yeasts, and there are notable differences between the lipid compositions of insect and mammalian membranes (Milić and Veprintsev, 2015). Insect cell membranes are low in cholesterol, high in phosphatidylinositol, and contain no phosphatidylserine; as mentioned previously, lipid-protein interactions are now known to be an important consideration in the structure and/or function of membrane proteins (Dawaliby *et al.*, 2016). Despite these disadvantages however, the baculovirus/insect cell system remains to be a key overexpression approach for GPCRs, particularly for their crystallisation, ranging from well-characterised to orphan receptors (Boivineau, Haffke and Jaakola, 2020; Chen *et al.*, 2022; Peng *et al.*, 2022).

Mammalian Cells

As a consequence of the disadvantages of prokaryote and lesser eukaryote systems, in some cases, it is simply necessary to use mammalian cell lines to produce functional, properly folded mammalian membrane proteins (Tate, 2001). Though this may seem an obvious solution, there are significant drawbacks which make mammalian cell lines one of the minor contributors to the resolution of GPCR structures, and why *E. coli*, yeasts and insect cells have been so widely favoured thus far. Mammalian cell lines exhibit relatively low expression levels, leading to longer time-frames to reach appropriate protein yields, and an increased expense of culture (Andréll and Tate, 2013). As yields are typically low, constructs are often first optimised for mammalian expression; examples include codon optimisation, and the inclusion of Kozak consensus (GCCACCATGG) and signal peptide sequences which can enhance expression and trafficking to the cell surface (Chakraborty *et al.*, 2015; Zucchelli *et al.*, 2016). However, an increased yield does not necessarily correlate with an increased functional yield, and it is similarly pragmatic to highlight that experimental conditions favourable for structural studies may not also apply to functional studies.

Several cell lines are commonly used including CHO, COS-7 and HEK293 but the level and quality of target protein expression can vary widely, often requiring individual optimisation (Andréll and Tate, 2013). Additionally, expression is either transient or stable; transient transfection can not only vary in the resulting expression level, but is also impermanent (Geisse and Fux, 2009). In contrast, the creation of stably expressing lines can be costly and time consuming, but benefits from reliably higher levels of heterologous protein expression; as such, several cell lines have been engineered which may confer advantages depending on the particular target protein. For example, the tetracycline-inducible HEK293S-TetR cell line was used to crystalise a mutant rhodopsin which was otherwise toxic to constitutive expression (Reeves, Kim and Khorana, 2002), and the HEK293-GnTI^{-} line lacks N-acetyltransferase I

activity which reduces complex glycan formation, leading to better ordered crystal formation for structure determination (Reeves *et al.*, 2002; Xu *et al.*, 2013; Milić and Veprintsev, 2015). Overall, despite the advantages of immortalised mammalian cell lines, their expense of culture and lower yield have largely limited their involvement in GPCR structure determination but may be more suited for functional analyses. In addition, the expression levels of mammalian approaches can remain problematic, despite optimisation.

Drosophila melanogaster

Several other promising but underutilised expression systems are available, one of which involves the eyes of the *Drosophila melanogaster* fruit fly. The unique architecture and properties of the *Drosophila* eye mean the rhabdomere membrane stacks of its photoreceptor cells can provide a large surface area for targeted overexpression and folding of membrane proteins (Kumar and Ready, 1995; Panneels *et al.*, 2011). This is facilitated by the well-established GAL4-UAS system, whereby the yeast transcription factor GAL4 binds to an Upstream Activation Sequence (UAS) to drive expression of a selected transgene (Phelps and Brand, 1998; Panneels *et al.*, 2011). By selecting strains which express the GAL4 transcription factor in the eye, heterologous expression of GPCRs such as the metabotropic glutamate and vasopressin receptors can be restricted to the photoreceptor cells without affecting endogenous membrane protein expression (Panneels *et al.*, 2011). This system has been demonstrated to achieve 0.2 to 0.4 mg of pure target protein from just 1 g of *Drosophila* heads (approximately 10 ml volume of flies), which is a relatively easy and affordable scale to achieve, even in smaller laboratories (Eroglu *et al.*, 2002; Panneels *et al.*, 2011). Alongside these attractive benefits however, *Drosophila* expression requires expertise in fly genetics, transgenesis and culture facilities which may not be easily accessible. In addition, *Drosophila* photoreceptor cells exhibit differences to mammalian cells, including less complex N-glycosylation (Bernaudat *et al.*, 2011; Schiller *et al.*, 2012). Overall, this underused expression system has the potential to rival the most popular systems over time but its limitations and lack of use means only 10 membrane protein structures have been resolved using *Drosophila* thus far, none of which are GPCRs (https://www.rcsb.org/). It is likely that failed attempts have not been reported or published.

Cell Free Expression

Finally, an interesting alternative system is cell-free expression which uses cell lysate to overcome the toxicity of overexpression and the sequestering of protein into inclusion bodies (Klammt *et al.*, 2006; Sobhanifar *et al.*, 2009). *In vitro* translation of target proteins enables careful construction of a folding environment, considering important factors including lipid

composition and membrane curvature, elasticity and lateral pressure. In summary, cell-free expression has been particularly useful for the inclusion of unnatural amino acids and isotopic labelling for NMR, and has successfully expressed several GPCRs for structural and functional studies (Muranaka *et al.*, 2007).

Fusion Proteins

As discussed, expression and purification of membrane proteins is often difficult; one improvement strategy involves the use of a fusion partner (Pandey *et al.*, 2016). There are a wide range of fusion partners available for use, and many examples of their application to GPCR expression. These fusion partners range from relatively larger globular proteins such as glutathione S-transferase, apocytochrome b562 (BRIL) and the HaloTag, to relatively smaller tags including haemagglutinin (HA), FLAG, and polyhistidine sequences. Further examples which typically yield a high degree of purity are the Strep-tag and Avi-tag which rely upon the biotinylation of the target protein, and biotin's high affinity for streptavidin resin (Schmidt and Skerra, 2007; Fairhead and Howarth, 2015). Not only can these additions improve yield through increased expression, some can also improve solubility and enhance subsequent purification by affinity chromatography, for example (Hochuli *et al.*, 1988; Marblestone *et al.*, 2006; Pandey *et al.*, 2014). However, in contrast, larger fusion partners tend to increase the metabolic burden on the cell, can sterically hinder the target protein itself, and can impede the crystallisation process due to the flexibility of linker sequences; importantly, many fusion partners can be designed to be removed by cleavage, eliminating some of these inherent disadvantages (Pandey *et al.*, 2016). As such, the use of fusion partners adds extra steps and complication to the expression process, but the potential benefits can be very significant.

Membrane Solubilisation

Following construct design and the choice of appropriate expression host, techniques for the solubilisation and purification of the target GPCR must also be carefully considered. Traditionally, surfactant detergents have been used to solubilise and extract proteins away from the membrane, through disrupting intra- and intermolecular interactions and increasing the protein's aqueous solubility (Seddon, Curnow and Booth, 2004; Anandan and Vrielink, 2016; Stetsenko and Guskov, 2017). Many detergents are commercially available, and can be screened by their physicochemical properties, critical micelle concentration, and hydrophilic-lipophilic balance, etc. (Arachea *et al.*, 2012). Overall however, the non-ionic alkyl maltopyranoside detergents DM and DDM have been most favoured, and have contributed to the resolution of nearly half of all membrane protein structures (Lyons *et al.*, 2016). In addition,

lauryl maltoside neopentyl glycol (LMNG) has proven particularly effective for the crystallisation of GPCRs in the lipidic cubic phase (LCP), and for the stabilisation of membrane proteins for cryogenic electron microscopy (cryo-EM) (Caffrey, 2015; Hauer *et al.*, 2015). Despite this overwhelming success however, detergents generally remove potentially important annular lipids, can result in the loss of protein-protein interactions, and detergent micelles do not exert the same lateral pressure as the native membrane. Therefore, to prevent GPCR destabilisation or inactivation during solubilisation, alternative methods have been developed to retain the native microenvironment around the membrane protein itself.

One of the popular alternatives uses the amphipathic co-polymer, styrene maleic acid (SMA) which spontaneously forms a lipid particle, or SMALP, when added to lipids (Knowles *et al.*, 2009). On average, this results in a nanoparticle structure approximately 10 nm in diameter which, when added to a biological membrane, has the ability to capture and solubilise membrane proteins and their surrounding lipid environment (Lee *et al.*, 2016; Pollock *et al.*, 2018). GPCRs are, on average, approximately 4 nm in diameter and are thus viable candidates for the SMALP method, especially because of their potential sensitivity to detergents and interactions with annular lipids (Wheatley *et al.*, 2016). The adenosine-2A receptor (A2AR) was one of the first reported successes of SMA solubilisation, producing purified and functional A2AR from both HEK293T and *Pichia* overexpression, without the use of detergent (Jamshad *et al.*, 2015). Since then, several more GPCRs have been purified with SMA such as the calcitonin, ghrelin and melatonin receptors, and SMA also contributed to the high resolution structure of the *Haloquadratum walsbyi* bacteriorhodopsin (Charlton, 2016; Logez *et al.*, 2016; Broecker, Eger and Ernst, 2017).

Additionally, several co-polymers have also been developed including styrene maleimide (SMI) which is acid compatible, and diisobutylene-maleic acid (DIBMA) which can overcome SMA's sensitivity to divalent cations (Oluwole *et al.*, 2017; L. Hall *et al.*, 2018). Overall, receptors solubilised with co-polymers are often far more stable at room temperature, can withstand freeze-thaw cycles, and retain ligand binding capacity when compared to detergent solubilised receptor (Jamshad *et al.*, 2015). This flexibility and utility has therefore made co-polymers a greatly important development in GPCR solubilisation, especially as the co-polymers are compatible with ligand binding and functional assays, *in meso* crystallisation and cryo-EM as well. Indeed, there have been several more solubilisation systems developed to overcome the problems associated with detergents, all of which solubilise membrane proteins successfully, and have additionally proven effective for the cryo-EM format (Sgro and Costa, 2018; Birch *et al.*, 2020).

Bicelles, or bilayered micelles, were the first of these to attempt the incorporation of lipids into a solubilised membrane protein system (Sanders and Prosser, 1998). Typically, phospholipids such as dimyristoylphosphatidylcholine (DMPC) are used to form discoidal bilayers of varying shapes and sizes, which are scaffolded by either detergent or short-chain lipids, such as CHAPS or dihexanoylphosphatidylcholine (DHPC), respectively (Thoma and Burmann, 2021). This incorporation of lipid into bicelles enables them to retain solubilised membrane proteins in a functional state when compared to micelles, however, the lipid composition of bicelles can also influence the molecular dynamics of membrane proteins significantly (Morrison and Henzler-Wildman, 2012; Vestergaard *et al.*, 2015). While the use of bicelles in NMR has become popular, they have also been implemented in the lipidic cubic phase of crystallisation, and were instrumental in the early study of both bacteriorhodopsin and the human β2-adrenergic receptor, for example (Faham and Bowie, 2002; Rasmussen *et al.*, 2007; Thoma and Burmann, 2021).

Since their inception, the bicelle fundamentals have been further developed, leading to the collective nanodiscs, which are also scaffolded lipid bilayer particles. One such class of amphipathic scaffold proteins was specifically engineered and optimised for the self-assembly of bilayer nanodiscs, termed membrane scaffold proteins (MSPs) (Thoma and Burmann, 2021). It was hypothesised then, that the inclusion of a target membrane protein within the self-assembly mechanism would enable its solubilisation into an environment more reminiscent of the native membrane. While bacteriorhodopsin was again used as a model target protein, MSP nanodiscs have been widely used for the structural, biochemical and biophysical analyses of various membrane proteins, and the nanodiscs themselves (Bayburt and Sligar, 2003; Sligar and Denisov, 2021). In addition, membrane proteins have also been solubilised and reconstituted into saposin-lipoprotein (Salipro) nanoparticles, which offers a detergent-free alternative for structural and functional studies (Frauenfeld *et al.*, 2016). And finally, detergents can also be substituted by the amphipols, or amphipathic polymers, which are not only more stable than their detergent counterparts, but can also be chemically labelled or functionalised, forming stable membrane protein complexes (Zoonens and Popot, 2014).

To conclude, why is this topic so important to address? Simply, the choice of expression host, protein modifications and solubilisation/purification techniques may have an impact on the resolved structure's likeness to its natural, wildtype equivalent. It would therefore be logical, pragmatic even, to hypothesise that structures significantly different to the wildtype receptor should be treated with caution until these differences are understood. Researchers should therefore be aware of and carefully consider their experimental methods, before drawing any conclusions from structures or assay results obtained from GPCRs under varying

experimental conditions. Chapters 4 and 5 examine this topic further, in which structural differences between GPCRs have been analysed in depth with geometric morphometrics.

1.1.4 – Structural Techniques for Membrane Proteins

Following the successful expression, solubilisation, and purification of stable target membrane protein, several techniques are available to gain structural information, and ideally elucidate three-dimensional models of the protein itself. X-ray crystallography, cryogenic electron microscopy (cryo-EM), and nuclear magnetic resonance spectroscopy (NMR) are the most common, each with their benefits and drawbacks, and will be briefly reviewed in terms of sample preparation and structure determination.

X-ray Crystallography

The leading method used to solve membrane protein structures is X-ray crystallography, contributing to approximately 80% of structures thus far. It is considered notoriously difficult to generate diffracting crystals for high-resolution analyses, further highlighting the importance of the pre-analytical techniques discussed in section 1.1.3, as milligram quantities of protein are often required for structure determination (Kermani, 2021). In addition, membrane protein crystals are typically fragile and susceptible to damage by X-rays, requiring a pre-crystallisation screening process to discover the best conditions for diffraction (Robert *et al.*, 2017).

Firstly, sample quality, in terms of aggregation and homogeneity, can be assessed by size exclusion chromatography (SEC), though this requires a relatively large amount of protein, is not considered high-throughput and can be quite time consuming (Kwan *et al.*, 2019). An alternative, therefore, is dynamic light scattering (DLS) which requires a smaller amount of protein sample, and can detect small differences in the hydrodynamic radii of protein-detergent complexes, for example (Meyer *et al.*, 2015). In addition, a thermal denaturation assay is a rapid, high-throughput method to assess the thermostability of the solubilised membrane protein (Alexandrov *et al.*, 2008). The thiol-specific dye, *N*-[4-(7-diethylamino-4-methyl-3-coumarinyl)-phenyl] maleimide (CPM), fluoresces upon binding to cysteine residues which are exposed to solvent due to membrane protein unfolding upon temperature increase. In any case, once the target protein sample has been screened for quality, the next step is to create the diffracting crystals themselves.

Vapor diffusion is most commonly used to crystallise membrane proteins, performed in either the sitting- or hanging-drop format (Delmar *et al.*, 2015; Kermani, 2021). Concentrated target protein, buffer and precipitant within a droplet are allowed to gently equilibrate with a larger

reservoir to gradually increase the protein and precipitant concentrations within the droplet. This causes the growth of large, well-ordered crystals within the droplet that can be ideal for diffraction; factors such as salt content, pH, temperature, and precipitant (most commonly PEG) are usually further optimised to find the best conditions for vapor diffusion (Delmar *et al.*, 2015). This method of crystallisation is characterised, however, by inherent disadvantages such as the loss of the protein's lipid environment and lateral pressure, and their occasional anisotropic orientation within the crystal (Wlodawer *et al.*, 2013). Anisotropy causes the non-uniform diffraction of X-rays favouring one direction which often resulted in unusable data, however, subsequent developments in data processing have since reduced this negative outcome (Kermani, 2021).

Consequently, to overcome these disadvantages, *in meso* crystallisation was developed to provide a membrane mimetic environment which has enabled the crystallisation of particularly difficult membrane proteins, including GPCRs (Landau and Rosenbusch, 1996; Ghosh *et al.*, 2015). The lipidic cubic phase does this by combining detergent-solubilised protein with neutral lipids such as monoacylglycerols. Protein-detergent micelles are reconstituted into this three-dimensional lipid bilayer of the mesophase, which then forms a lamellar phase upon addition of the precipitant; crystal growth subsequently takes place within the lamellar phase, under optimised conditions (Caffrey, 2015). Furthermore, an additional technique to further aid the crystallisation process includes the optional use of a soluble protein chaperone such as T4 lysozyme, Fab antibody fragments, or nanobodies to increase the surface area needed for crystal contact and formation (Kermani, 2021).

Two types of crystal can form depending on the organisation of lipid and protein packing and generally correlate with the technique used to create them; type I crystals, often from the lipidic cubic phase, are composed of planar sheets stabilised by both hydrophobic and polar interactions, while type II crystals, frequently from vapor diffusion, rely mainly on polar interactions as protein hydrophobic regions are covered by detergent micelles (Birch *et al.*, 2018). This generally makes type II crystals more fragile, sensitive to radiation damage, and even diffract to relatively low resolution, further highlighting the benefits of *in meso* crystallisation (Carpenter *et al.*, 2008). Once prepared, these crystals are then mounted on a loop, flash frozen in liquid nitrogen, and exposed to a powerful X-ray beam; the angles and intensities of diffracted X-rays are measured as the crystal rotates, gradually building a unique diffraction pattern which can be refined into an electron density map. Both amplitude and phase are required for this, however, X-ray diffraction only provides the former, presenting a significant problem. The fastest method to obtain phase data is to calculate it from previously solved homologous structures, novel structures however would require experimental phasing which can be time-consuming and costly (Parker and Newstead, 2013; Huang *et al.*, 2018).

Finally, the electron density map is refined and fitted to produce the atomic model of the target membrane protein's crystal structure which would then be submitted to the Protein Data Bank (PDB).

Overall, X-ray crystallography has proven to be an invaluable technique in solving the majority of protein structures, despite the many difficulties involved in obtaining a well-diffracting crystal for high resolution analyses.

Cryogenic Electron Microscopy

One potential rival to the resounding success of X-ray crystallography is cryogenic electron microscopy (cryo-EM); EM has proven useful for the determination of membrane protein structures in proteolipid crystals, however, its capability for single particle analysis has been enabling high resolution structures to be determined without the need for well-ordered crystal preparation (Raunser and Walz, 2009; Moraes *et al.*, 2014). Cryo-EM has therefore received much attention from the structural biology field, especially as membrane proteins are able to be preserved in unrestrained states through rapid vitrification. Protein complexes can also be observed in multiple conformations, meaning cryo-EM can be used to study proteins too flexible or heterogeneous for the well-ordered crystals needed for X-ray diffraction. Indeed, the number of structures resolved using cryo-EM has been rising exponentially, and resolutions of 3 Å or less are achieved routinely (Wu and Lander, 2020).

Cryo-EM typically requires less sample than X-ray crystallography, which is of particular use with proteins which are considered relatively low-yield or unstable over time. Despite this, and similarly to X-ray crystallography, the quality of pure protein sample is of utmost importance. While sample purity can be determined through the use of size exclusion chromatography and sodium-dodecyl-sulfate-polyacrylamide gels (SDS-PAGE), sample quality and homogeneity can also be assessed by negative staining transmission electron microscopy. Protein samples are negatively stained and dehydrated by the addition of heavy metals which provide contrast against the target protein. This technique is rather limited in terms of the resolution and information gathered from the sample, however, negative staining can provide a relatively quick and meaningful assessment of the sample's quality and purity, especially in a high-throughput format (Renaud *et al.*, 2018; Lyumkis, 2019).

Once a sufficiently optimised sample of the target protein has been acquired, it is then cryogenically cooled in liquid ethane or liquid ethane-propane to form a frozen-hydrated sample (Dubochet *et al.*, 1988; Tivol, Briegel and Jensen, 2008). The target protein is subsequently captured on an EM grid, with randomly distributed orientations, within the thinnest-possible layer of vitrified ice, able to withstand the vacuum of the transmission

electron microscope (Kim *et al.*, 2018). Importantly, this can be a highly variable process, involving factors such as buffer composition, grid substrate, humidity, ice thickness, protein stability and temperature, each of which can impact upon the quality of vitrified sample (Wu and Lander, 2020).

As the electron beam is fired at the sample-containing EM grid, thousands of electron micrographs can be created which result in hundreds of thousands, even millions, of single particle images. These images are then processed in an increasingly standardised computational workflow, beginning with motion correction and dose-weighting. This corrects any sample movement due to the electron beam itself, and filters image frames to account for accumulative radiation damage (Wu and Lander, 2020). Biological specimens within the corrected micrographs are then algorithmically identified and aligned by two-dimensions to initially remove any irrelevant or low-quality particles. The remainder are then subjected to three-dimensional alignment and classification, which can also separate any particle heterogeneity; this is an iterative process which ultimately improves the resolution of the final three-dimensional reconstruction (Penczek, 2010b). In the cases of particularly dynamic or flexible protein complexes, a more focussed refinement may be needed as multiple structural states make image reconstruction more challenging (Punjani and Fleet, 2021). Despite technological advancements and the development of image processing software, significant user input is still required to assess and advance each processing step, limiting the potential for fully automated cryo-EM image reconstruction (Zivanov *et al.*, 2018).

Where the maps and models of X-ray diffraction are iteratively refined within the same process, cryo-EM differs in that the maps and atomic models are generated independently, highlighting the need for robust structure validation, which is yet to be fully standardised (Wu and Lander, 2020). Firstly, cryo-EM map quality is assessed by the Fourier shell correlation (FSC) curve which analyses the agreement between two independently refined random subsets of the map data; not only does the FSC value indicate the map's nominal resolution, the shape of the curve can also be indicative of issues such as sample heterogeneity and duplicate particles (van Heel and Schatz, 2005; Penczek, 2010a). In addition, two-dimensional FSC curves can be expanded to a three-dimensional curve which describes the anisotropy of the reconstruction – a challenge also faced by X-ray diffraction, as mentioned (Naydenova and Russo, 2017). Furthermore, the atomic models generated by cryo-EM reconstruction must also undergo validation through several metrics, in a similar manner to those generated by X-ray diffraction. In summary, this ensures that the cryo-EM maps and atomic models accurately represent the data acquired from the protein sample, and the resulting image reconstruction is a true reflection of that data (Herzik, Fraser and Lander, 2019; Wu and Lander, 2020).

Overall, advancements in cryo-EM single particle analysis over the last decade have facilitated a plethora of discoveries with regards to the structural and mechanical properties of membrane proteins. As its use increases exponentially however, progress towards greater standardisation and automation must be made, to further enhance this technique's efficiency and user accessibility.

Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy has thus far played a key role in understanding the fundamental structural, chemical and dynamic properties of phospholipid bilayers, and their various components. However, in addition, NMR has also been central to the structural biology of membrane proteins, providing unique and complementary information to those garnered by X-ray crystallography and cryo-EM (Cross *et al.*, 2014; Opella and Marassi, 2017). A range of sample types are compatible with NMR, including detergent solubilised, detergent-free and native membrane, and NMR can even detect weak ligand binding events to correlate with conformational changes, for example (Yao *et al.*, 2015). As NMR is compatible with detergent-free samples, membrane proteins can be analysed in states relatively closer to their native microenvironment which contrasts with X-ray diffraction and cryo-EM in which at least some detergent or amphiphilic polymer are typically required; this, combined with their cryogenic temperatures, limits the dynamic plasticity of membrane proteins in order to stabilise a single conformation, which does not need to be the case with NMR. In fact, recent advances in NMR technology, methodology, and its combination with computational techniques means membrane proteins can be effectively probed at single atomic sites with ever increasing accuracy (Opella and Marassi, 2017).

Firstly, solution NMR has been used to determine the structure of membrane proteins within both detergent- and detergent-lipid micelles, including the outer mitochondrial voltage-dependent anion channel (VDAC-1) and the phototaxis receptor sensory rhodopsin II (pSRII), which forms a seven-transmembrane helix bundle in a similar manner to GPCRs (Hiller *et al.*, 2008; Gautier *et al.*, 2010; Nietlispach and Gautier, 2011). Importantly, each of these NMR-determined structures correlated with their X-ray diffraction counterparts, with no significant conformational differences between crystal and micelle samples (Opella and Marassi, 2017). In addition, the structures of several single-transmembrane helix membrane proteins, which typically do not form interhelical contacts strong enough for crystal formation for diffraction, have been determined by solution NMR which can overcome this limitation (Opella and Marassi, 2017). Furthermore, detergent solubilisation can easily disrupt ligand-receptor interactions, and thus solution NMR has been able to correlate the ligand binding of GPCRs

to conformational states, and additionally probe the native activity of several other membrane proteins (Chung *et al.*, 2012; Zhou and Cross, 2013).

Solution NMR is also compatible with membrane proteins contained within lipid bilayer systems including nanodiscs and SMALPs, eliminating the potential negative impacts of the use of detergents (Hagn *et al.*, 2013; Dörr *et al.*, 2014). This has proven particularly effective in determining the structures of several β-barrel bacterial outer membrane proteins, by measurement of distances between the inter-strand hydrogen bonding network; this contrasts with the intra-helical hydrogen bonding of α-helical proteins, which, though interesting, does not reveal their three-dimensional folding patterns (Hagn *et al.*, 2013; Sušac, Horst and Wüthrich, 2014; Dutta, Yao and Marassi, 2017). In addition, solution NMR in the absence of detergent ensures that protein structures can be studied in functional states. The outer membrane protein Ail shows great agreeability between NMR and crystal structures, however, its analysis by solution NMR with detergent eliminates its ligand binding capability; this was retained by instead performing detergent-free solution NMR in nanodiscs, suggesting the detergent molecules were interacting with key residues for ligand binding and/or functionality (Yamashita *et al.*, 2011; Ding *et al.*, 2015; Marassi *et al.*, 2015).

Solid-state NMR is also compatible with detergent-free samples, and membrane protein structures have been determined with both magic angle spinning (MAS) and orientated sample (OS) methods, which use site-specific resonance and restraint measurement, often via ^{13}C and ^{15}N, for example (Tang, Knox and Nevzorov, 2012; Baker *et al.*, 2015). MAS is able to convert isotropic chemical shift and spin exchange into inter-atomic distances and torsion angles, while OS uses anisotropic chemical shift and dipole coupling to give information on the dynamics and conformation of uniaxially ordered protein samples (Tian, Song and Cross, 1998; Park *et al.*, 2010). Importantly, solid-state NMR can also probe the mechanisms behind the functionality of membrane proteins in near-native samples, including the conductivity of ion channels such as KcsA, and the ligand binding and dynamics of GPCRs such as CXCR1 and rhodopsin (Bhate *et al.*, 2010; Smith, 2010; Park *et al.*, 2011; Tapaneeyakorn *et al.*, 2011). In addition, solid-state MAS NMR has proven useful for answering specific questions, using stable isotope-labelled amino acids as specific probes.

Overall, both solution and solid-state NMR have proven valuable tools in the analyses and determination of membrane protein structures, especially those with particularly challenging dynamics such as GPCRs. NMR is not only able to provide unique information *ab initio*, but also complements the rigid atomic models of crystallography in terms of ligand binding and conformational detail.

1.1.5 – Computational Techniques for Membrane Proteins

Despite the many developments in the expression, purification and technology available for membrane proteins over the last fifty years, the determination of their structures remains difficult due to their many inherent challenges, as discussed. Consequently, structure determination is a typically laborious task which can take many years to complete; the diacylglycerol kinase structure took thirteen years to elucidate by NMR, and a presenilin aspartate protease required 160,000 crystallisation attempts before success, for example (Van Horn *et al.*, 2009; Li *et al.*, 2013). As such, computational methods have been attempting to predict and model membrane protein structures to alleviate the challenges of experimental methods, in a complementary manner. This in itself poses significant challenges, though many advancements have also been made in the computational biology field, summarised effectively by Leman *et al.* 2015 and Almeida *et al* 2017.

Briefly, tools based upon sequence first aim to predict secondary structures of proteins, including transmembrane span, overall topology, and motifs involved in helical interactions, often aided by multiple sequence alignments and artificial intelligence. Subsequent to this, techniques to predict three-dimensional tertiary structures then include *de novo* prediction, homology modelling, fold recognition, and molecular dynamics simulations. Many computational tools including databases, software, and web servers exist for these purposes, and are discussed in the two mentioned reviews, but most are not relevant to this thesis (Leman, Ulmschneider and Gray, 2015; Almeida *et al.*, 2017).

The AlphaFold project is an artificial intelligence system able to predict a protein's three-dimensional structure from its amino acid sequence, created by a partnership between Google's DeepMind and the European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI) (Jumper *et al.*, 2021; Varadi *et al.*, 2022). AlphaFold's latest release contains over 200 million entries from 48 key organisms, and has been ranked as the best protein structure prediction tool by the Critical Assessment of Techniques for Protein Structure Prediction (CASP14) initiative, by a wide margin. While many other computational methods fall short in the atomic accuracy of their predictions, AlphaFold is thought to be accurate and comparable to experimentally resolved structures in the majority of cases; a median root mean square deviation (RMSD) value of 0.96 Å is observed between equivalent Cα atoms, compared to 2.83 Å from the next best method (Kufareva and Abagyan, 2012; Jumper *et al.*, 2021). The confidence of each residue is denoted by a pLDDT value between 0 and 100, corresponding to a superimposition-free difference of distance (Mariani *et al.*, 2013). A pLDDT score greater than 90 is considered high accuracy and suitable for the characterisation of binding sites, for example. Scores between 70 and 90 are generally

modelled well with a reliable backbone, between 50 and 70 is considered low confidence, and less than 50 pLDDT is a predictor of disordered regions or areas of complex formation (Tunyasuvunakool *et al.*, 2021). In addition, each AlphaFold structure has a Predicted Aligned Error (PAE) output which assesses confidence in the domain packing and overall topology of the protein, and is generated independently from the three-dimensional structure; a generally low PAE score indicates well defined relative positions and oriented domains.

Despite the seemingly resounding success of the AlphaFold project, there exist several limitations of its structure prediction capabilities, which remain to be active research areas. Firstly, this version of AlphaFold is not able to use multiple-chain sequences, meaning it cannot predict proteins in complex, and those residues which could be involved in complex formation generally exhibit low confidence scoring; this may be useful for identifying such regions, but cannot predict any likely conformation (Tunyasuvunakool *et al.*, 2021). In addition, AlphaFold has not been validated for use with mutations, in particular, the potential effects of destabilising mutations on protein structure and folding. Furthermore, AlphaFold is usually only able to produce one single conformation, even if the protein possesses multiple conformations, and AlphaFold also cannot predict positions of non-protein components such as co-factors or post-translational modifications which are often necessary for full functionality of the protein in question (Jumper *et al.*, 2021).

Overall, despite these limitations, AlphaFold is considered to be a very powerful tool in the prediction of protein structures from sequence, and has grown greatly in popularity when compared to current alternative methods. This topic is examined further in Chapter 5, in which GPCR structures predicted by the AlphaFold project have been compared to their experimentally resolved counterparts with the use of geometric morphometrics.

1.2 – The G Protein-Coupled Receptor (GPCR) Superfamily

1.2.1 – Classification and Taxonomy

Throughout the early 1900s, alongside discoveries of the cell membrane's nature, the discipline of pharmacology was tasked with explaining the physiological mechanisms which underpin the cellular effects of endogenous and exogenous compounds (Maehle, 2009). By observing the effects of curare and nicotine on skeletal muscle, the terms 'receptive substance' and the 'receptor concept' were proposed to explain the effects of these drugs (Langley, 1905; Maehle, 2004). These theories were confirmed and succinctly summarised by the Hill-Langmuir law of mass action which links drug concentration to receptor occupancy, forming the important foundation from which modern pharmacology was born (Poyner and Wheatley, 2009). Since then, there has been a multidisciplinary approach to further studying the receptor concept and analysis of the human genome highlights nearly eight thousand proteins involved in cell reception; one of these important groups are called the G protein-coupled receptors, or GPCRs (Kelly *et al.*, 2009).

The GPCR superfamily is the largest group of transmembrane proteins in the human genome, containing over 800 diverse receptors, though estimations suggest this could be closer to 1,300 (Kobilka, 2007; Tuteja, 2009). GPCRs are composed of seven stretches of 20 to 35 hydrophobic residues which equate to seven conserved transmembrane (TM) alpha helices, formed in an anti-clockwise pattern when visualised from the extracellular space (Schwartz *et al.*, 2006). This feature led to the term 'heptahelical, or 7TM receptors' where the transmembrane helix bundle is connected by three extra- and three intracellular loops, as illustrated in Figure 1.2.1. Additionally, this largely conserved structure has an N-terminal extracellular domain (ECD) which can vary in size, with the intracellular C-terminal end following helix 8. By recognising a wide variety of endogenous ligands, GPCRs are able to subsequently activate intracellular signalling pathways via heterotrimeric G proteins and are therefore crucial to a wide variety of physiological processes (Bockaert and Pin, 1999; Jacoby *et al.*, 2006).

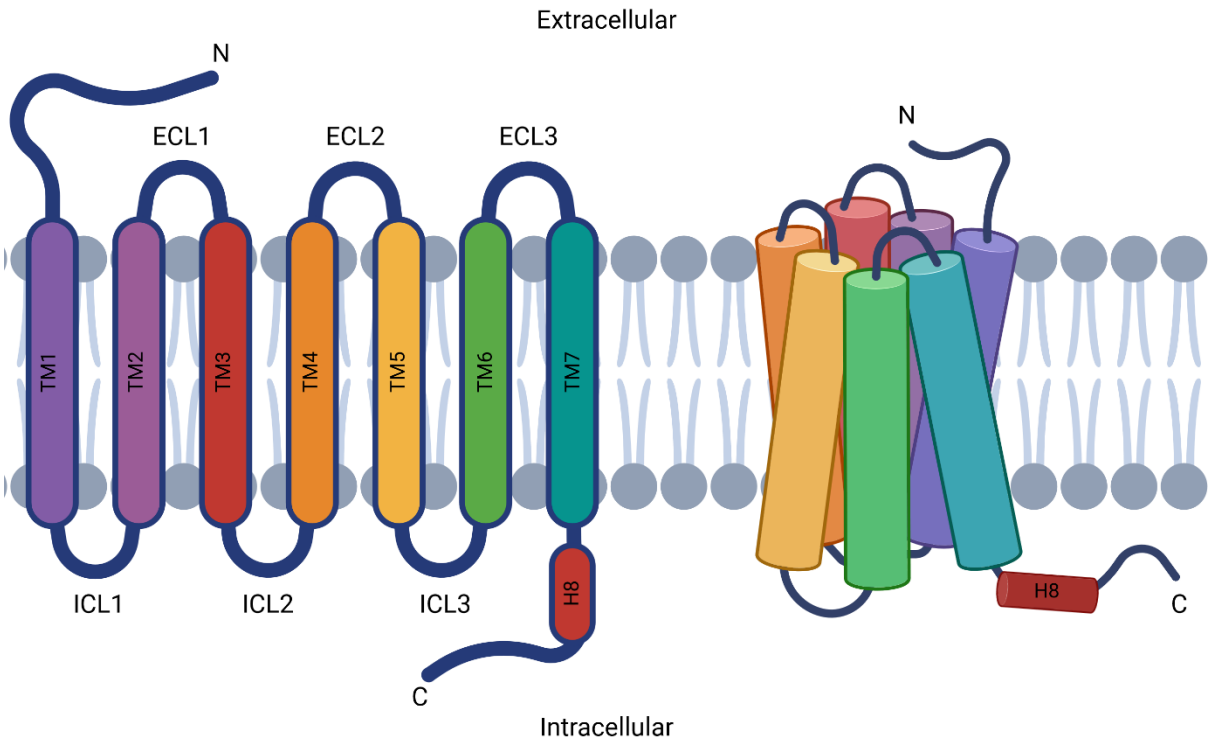

Figure 1.2.1. The Defining Structural Features of G Protein-Coupled Receptors. The two-dimensional illustration shows a GPCR with its extracellular N terminus (N), seven transmembrane helices (TM1-7, multicoloured) held together by three extracellular loops (ECL1-3) and three intracellular loops (ICL1-3), followed by helix 8 (H8) and the intracellular C terminus (C). The three-dimensional illustration highlights how the transmembrane helices form a bundle within the membrane, once the GPCR has been correctly folded. Created with BioRender.com.

As the GPCR superfamily is so large and diverse, a working model of their taxonomy was needed to understand how receptor features could be meaningfully classified and studied in more manageable group sizes (Kolakowski, 1994; Fredriksson and Schiöth, 2005). The most commonly used A to F system groups GPCRs into six families based upon transmembrane sequence homology, summarised in Table 1.2.1. While all of the metazoan GPCRs were classified into families A, B or C, more recently discovered mammalian GPCRs are not always as easily classified using the A to F system. Due to these limitations, the phylogenetic GRAFS taxonomy was created which suggested the Glutamate-like, Rhodopsin-like, Adhesion-like, Frizzled/Taste2-like and Secretin-like receptors would provide more useful or appropriate grouping (Fredriksson *et al.*, 2003). However, GPCR classification remains a point of contention within the field, many receptors are still considered orphans, and both taxonomy systems are used as often as each other. Orphan receptors are those whose cognate ligands are yet unknown, but have been classified as GPCRs based upon sequence and/or structure. For brevity, this thesis will consistently use the A to F system alongside naming the specific receptor sub-group, for example, the family A neurotensin receptors. In addition, the

Ballesteros-Weinstein numbering system will also be used; for example, P5.50 means proline found on transmembrane helix five, residue number relative to the most conserved residue (50) (Ballesteros and Weinstein, 1995; Isberg *et al.*, 2015).

Table 1.2.1. The A to F Classification Summary of GPCRs. Families A to C encompass the majority of the physiologically relevant receptors, while families D to F are often less well studied due to their limited application to mammalian physiology. Understandably, structural and pharmacological studies have largely prioritised mammalian receptors with application to human therapeutics.

Family	Description	Features	Ligand Examples
A	Rhodopsin-like	Lack substantial ECD; native ligands bind directly to TM	Adrenaline, neurotensin, opioids, photons, vasopressin
B	Secretin-like	Flexible, folded ECD with hormone binding motif	Calcitonin, glucagon, parathyroid hormone
C	Metabotropic glutamate receptors	Dynamic ECD with venus flytrap module	Calcium, GABA, glutamate
D	Fungal mating pheromone receptors	Activates MAPK signalling to form diploid zygotes	Cell membrane mating factors
E	cAMP receptors	Able to bind DNA	Cyclic AMP
F	Frizzled/Smoothened receptors	Mediates Wnt and Hedgehog signalling pathways	Cholesterol, Smoothened agonist (SAG), Wnt protein ligands

1.2.2 – The Role of GPCRs in Health and Disease

The GPCR superfamily contains at least 800 members, likely more, of which half are not considered a priority in terms of druggability due to their involvement in specific functions such as sight, smell and taste (Lagerström and Schiöth, 2008; Heng, Aubel and Fussenegger, 2013). However, the remaining half may be involved in the pathophysiology of human diseases and therefore have been subject to study for the eventual development of therapeutic treatment to improve patient health and outcomes. Indeed, GPCRs have been implicated in cardiovascular, cancer, endocrine, immune and neural diseases amongst others, and approximately 40% of all approved drugs mediate their effects through GPCRs (Thompson, Cole and Jose, 2008; Eglen and Reisine, 2009). These drugs are classified into four main categories: orthosteric which interact with the main ligand binding domain of the receptor, allosteric which bind to a site other than the orthosteric, drugs which modulate GRK and β-arrestin activity, and drugs which modulate receptor oligomerisation (Mason *et al.*, 2012; Heng, Aubel and Fussenegger, 2013). A brief review of diseases involving GPCRs and their drugs will now be given, to provide context as to why these receptors have been, and will always be of such importance.

Cancer

Cancer is one of the major challenges in human health, with around twenty million diagnoses, and ten million deaths per year globally; despite significant advances in treatment options and technology, prognoses remain poor overall with relatively high rates of mortality (Chhikara and Parang, 2023). While cancers can develop in many different forms and locations within the human body, they are underpinned by one common factor – uncontrolled cell growth and proliferation. Various GPCRs have been implicated in the development of cancers, not only being attributed to uncontrolled cell growth and proliferation, but also to angiogenesis, metastasis, cell differentiation, and aberrant apoptotic pathways (Heng, Aubel and Fussenegger, 2013).

Cancer cells often constitutively overexpress various potent mitogens for which their cognate receptors are GPCRs. The lysophosphatidic acid (LPA), prostaglandin, protease-activated (PAR), and sphingosine-1-phosphate (S1PR1) receptors are a few of such examples which mediate cancer cell growth, as well as various neuropeptide receptors (Pyne and Pyne, 2000; Umezu-Goto *et al.*, 2004; Wang and DuBois, 2006; Elste and Petersen, 2010). Ovarian cancer cells secrete LPA into the ascitic fluid alongside overexpression of the LPA receptor; one specific example of an autocrine-paracrine feedback loop which serves to amplify signalling towards uncontrolled ovarian cell growth and proliferation (Mills and Moolenaar, 2003). In other cases, GPCRs are known to interact with further growth-related pathways such as with the tyrosine kinase receptors; the epidermal growth factor (EGFR), insulin (InsR), and insulin-like growth factor-1 (IGF1R) receptors are all involved in the pathogenesis of various cancers (Kisfalvi *et al.*, 2007; Bhola and Grandis, 2008; Rozengurt, Sinnett-Smith and Kisfalvi, 2010). The cross-talk between EGFR and GPCRs including the acetylcholine, bradykinin, LPA, and prostaglandin E2 receptors have been implicated in squamous-cell carcinomas (Thomas *et al.*, 2006). Additionally, activation of the PAR and endothelin A receptors through EGF signals the proliferation of breast carcinoma, for example (Arora *et al.*, 2008; Fischgräbe *et al.*, 2010). Interestingly, the cross-talk between insulin and IGF1 receptors with G_q-coupled GPCRs has been disrupted by the anti-diabetic drug metformin in the context of preventing proliferation of pancreatic cancer cells (Kisfalvi *et al.*, 2007).

GPCRs which are involved in cancer cell growth can also exhibit other pleiotropic effects – the CXCR2, CXCR4, endothelin, and protease-activated receptors are also involved in angiogenesis via upregulating chemokines such as the vascular endothelial growth factor (VEGF), IL-8 and CCL5, for example (Strieter *et al.*, 2006; Martin, Galisteo and Gutkind, 2009; Wheeler-Jones, Farrar and Garonna, 2009; Yang *et al.*, 2010). These exert pro-angiogenic effects by acting directly upon the endothelium, or by promoting endothelial cell migration into

the tumour tissue itself; this process can also involve the amplifying effects of feedback loops such as the VEGF-mediated translocation of S1PR1 receptors to the endothelial cell membrane, which is itself a result of S1PR1 receptor activity (Pitson *et al.*, 2004). In addition to angiogenesis, chemokine networks and their associated GPCRs are also responsible for cancer metastasis; one of the prominent combinations is the chemokine SDF1 and its cognate GPCR CXCR4 which can metastasise many different cancer cell types into secondary locations such as the bone marrow, liver, lungs and lymph nodes (Gelmini *et al.*, 2008; de Nigris *et al.*, 2012). Furthermore, metastasis can also occur as a result of extracellular matrix degradation – the prostaglandin E2 receptor can cause this via upregulated cyclooxygenase-2 activity and secretion of metalloproteinase enzymes (Reich and Martin, 1996; Cao and Prescott, 2002).

The stem cell origin theory of cancer suggests that the abnormal differentiation of adult stem cells is the root cause of cancer; family F GPCRs which mediate Wnt/Hedgehog signalling pathways are thought to be closely involved in this, as they maintain the stem cell population for tissue regeneration and wound healing (Katoh and Katoh, 2007). Indeed, overactive Wnt signalling leads to increased nuclear β-catenin, which has been linked to cell differentiation and various cancer types (Polakis, 2000; Hoppler and Kavanagh, 2007). Additionally, the family F smoothened (SMO) receptor is a key mediator of Hedgehog signalling, and has been linked to transcription factors which promote oncogenesis; cyclopamine is a promising anti-cancer drug able to inhibit SMO receptors directly (Ayers and Thérond, 2010; Gould and Missailidis, 2011). Finally, many cancer cells are resistant to apoptosis, leaving them unable to be destroyed and cleared away without intervention. Approximately half of all cancers exhibit mutations which inactivate the p53-mediated apoptotic pathway; even without these mutations, abnormal GPCR activity can result in the same effect (Haupt and Haupt, 2006; Murph *et al.*, 2007). The LPA receptor which can be responsible for uncontrolled cancer cell growth, also reduces cellular and nuclear levels of p53, inhibiting apoptosis in lung carcinoma cells (Murph *et al.*, 2007). Indeed, cancer cells can arise from both constitutively active GPCR mutants, as well as aberrant expression which leads to autocrine-paracrine feedback loops, amplifying the cellular characteristics discussed.

Infectious Disease

GPCRs have also been implicated in the binding and entry of exogenous pathogens, resulting in GPCR roles in infectious disease (Suresh and Wanchu, 2006; Andrade *et al.*, 2012). The human immunodeficiency virus 1 (HIV-1) is one such example which, after recognition by the glycoprotein CD4, requires a secondary interaction to gain entry to T lymphocytes; this is often facilitated by a tyrosine-rich domain at the N terminus of chemokine GPCRs such as CCR5

and CXCR4 (Shimizu *et al.*, 2008; Pollakis and A Paxton, 2010). CCR5, being involved in the early stage of new HIV-1 infections, is one of the focussed targets of research for HIV treatment (Maeda *et al.*, 2012; Weichseldorfer *et al.*, 2022). Some pathogens are able to hijack host signalling pathways to enhance their infection, such as with the protozoan parasite *Trypanosoma cruzi*. This parasite secretes an enzyme to activate bradykinin, and can also induce endothelin receptor signalling, both of which promote it's infection by inflammatory oedema, causing Chagas disease (Andrade *et al.*, 2012; Coura and Borges-Pereira, 2012). Finally, some pathogens, such as herpes viruses, can also express their own GPCRs to facilitate their infection and evasion of host defences (Vischer, Vink and Smit, 2006). The Epstein-Barr virus, human cytomegalovirus, and Kaposi's sarcoma virus all encode GPCRs which share close sequence homology to human chemokine receptors, for example (Beisser *et al.*, 2008; Zuo *et al.*, 2009; Slinger *et al.*, 2011). These are able to influence viral replication, the survival of infected host cells, and evasion of host defences, through molecular mimicry of human GPCRs gained through co-evolution (Vischer, Vink and Smit, 2006).

Immunological Disorders

The human immune response is an incredibly complex system, but is commonly synonymous with inflammation. This involves the movement of cells, vascular permeability, and occurs in response to cellular stress, pathogens or other harmful stimuli. As expected, due to their wide variety of functions, GPCRs are studied as the potential targets of anti-inflammatory drugs to treat immune disorders (Sun and Ye, 2012). Chemokine, eicosanoid, and histamine receptors are all GPCRs involved in the inflammatory process, promoting the movement of leukocytes and pro-inflammatory cytokines to the site of inflammation (Olson and Ley, 2002; Khanapure *et al.*, 2007; O'Mahony, Akdis and Akdis, 2011). In addition, S1PR1 and S1PR2 receptors are expressed at the plasma membrane of mast cells, modulating migration and degranulation via their autocrine mode of action (Rivera, Proia and Olivera, 2008). It is during this degranulation process that antimicrobials and other inflammatory molecules such as histamine are released. Adenosine and ATP can also be found at sites of inflammation, and their cognate purinoceptors modulate both pro- and anti-inflammatory effects; the adenosine-2A receptor (A2AR) is one such widely studied GPCR which, in the context of inflammation, downregulates pro-inflammatory cytokines and inhibits T cell activation (Huang *et al.*, 1997; Haskó *et al.*, 2000; Ferrero, 2011). Finally, several other GPCRs have been linked to inflammation including several protease-activated receptors (PARs) which increase vascular permeability, the neurokinin receptors which promote vasodilation through the release of histamine and nitric oxide, and the anaphylatoxin receptors for C3a and C5a generated by the complement cascade (Campos and Calixto, 2000; Bunnett, 2006; Zhou, 2012). Indeed, GPCRs seem to

be an attractive target for the treatment of inflammatory diseases such as arthritis, inflammatory bowel disease and multiple sclerosis, for example (Arévalo-Martín *et al.*, 2008; Wright, Duncan and Sharkey, 2008; Tsou *et al.*, 2022).

Obesity

Obesity is a multi-faceted disease which has tripled globally since 1975, causing a major public health challenge, along with its related metabolic disorders (Boutari and Mantzoros, 2022). Obesity arises from a chronic imbalance between energy intake and expenditure, but often involves underlying physiological mechanisms with several key GPCRs. The peptide ghrelin is produced in the stomach and links the nervous systems via its cognate GPCR, the growth hormone secretagogue (GHS) receptor in the hypothalamus and pituitary gland (De Ambrogi, Volpe and Tamanini, 2003; Kojima and Kangawa, 2005). This 'hunger hormone' is secreted in response to fasting, reduces satiety in human participants and, importantly, administration of a ghrelin antagonist results in reduced body weight and fat mass in rodents (Laferrère, Hart and Bowers, 2006; Maletínská *et al.*, 2011). The melanin-concentrating hormone and melanocortin receptors also regulate energy homeostasis through food intake, along with the orexin/hypocretin system which are circadian specific examples (Saito and Nagasaki, 2008; Mountjoy, 2010; Teske, Billington and Kotz, 2010). In contrast, the neuropeptide B and Y system reduces food intake, boosting metabolic activity and energy expenditure over a long term administration (Mondal *et al.*, 2003; Aikawa *et al.*, 2008; Hondo, Ishii and Sakurai, 2008). Finally, the bile acid receptor, TGR5, is expressed in various tissues such as brown adipose tissue and skeletal muscle (Chen *et al.*, 2011). It's activation via cholesterol metabolism induces the cAMP-dependant production of T3 thyroxine which in turn increases energy expenditure and heat generation (Watanabe *et al.*, 2006). These GPCRs then, several of which were once considered orphans, are promising targets for the future treatment of obesity and its complex underlying mechanisms.

Type 2 Diabetes

Type 2 diabetes is commonly one of obesity's co-morbidities, found to affect approximately 6.3% of the global population, costing between three and nine times the average healthcare expenditure per capita and is currently the ninth leading cause of mortality worldwide (Lam and LeRoith, 2012; Khan *et al.*, 2020). It is characterised by insulin resistance, abnormal glucose homeostasis, and pancreatic dysfunction which can eventually lead to cardiovascular disease, renal failure and retinopathy, for example (Deshpande, Harris-Hayes and Schootman, 2008). The islet β-cells of the pancreas secrete insulin in response to the incretin gastrointestinal peptides, glucagon-like peptide 1 (GLP1) and glucose-dependant

insulinotropic polypeptide (GIP), upon food intake (Holst, Vilsbøll and Deacon, 2009). Activation of the GLP-1 receptor stimulates the secretion of insulin, but also inhibits that of glucagon; several synthetic analogues of GLP-1 agonists are used for the treatment of type 2 diabetes (Montanya, 2012). The GIP receptor also mediates glucose-dependant secretion of insulin, but additionally plays a role in glucagon secretion as well (Fujimoto *et al.*, 1978; Ehses *et al.*, 2001). The glucagon receptor itself, mainly expressed in the kidney and liver, often displays abnormal function in the context of type 2 diabetes, increasing blood levels of glucose from glycogen metabolism (Qureshi *et al.*, 2004). There has been some success in treating diabetes with glucagon receptor small molecule antagonists and monoclonal antibodies (Gu *et al.*, 2009; Mu *et al.*, 2011; Xiong *et al.*, 2012).

Likewise, free fatty acids (FFA), produced by colonic bacterial fermentation, are also involved in glucose homeostasis via their cognate GPR receptors; these are less well understood examples, but are thought to exert effects mainly upon islet β-cells, and stimulate secretion of both insulin and glucagon (Rayasam *et al.*, 2007; Feng *et al.*, 2012). Neurotransmitter systems also play an important role in type 2 diabetes, involving GPCRs such as the acetylcholine, gastrin-releasing peptide (GRP), pituitary adenylate cyclase-activating peptide (PACAP), and vasoactive intestinal peptide (VIP) receptors (Harmar *et al.*, 2012). Interestingly, the main transmitter within islet sensory innervation is the calcitonin gene-related peptide (CGRP) which inhibits glucose-dependent insulin secretion, providing a further avenue for possible therapeutics (Bretherton-Watt *et al.*, 1992). Finally, the GPR54, purinergic, and vasopressin receptors are also noteworthy, further demonstrating the complexity and wide variety of GPCRs in glucose homeostasis, insulin secretion, and type 2 diabetes (Oshikawa *et al.*, 2004; Hauge-Evans *et al.*, 2006).

Cardiovascular Disease

Another of obesity's co-morbidities are the group of cardiovascular diseases, including heart disease and stroke, which remain to be the leading cause of mortality worldwide (approximately twenty million deaths per year, globally) (Mensah, Roth and Fuster, 2019). While acute activation of GPCRs can often be cardio-protective, chronic activation of the approximately two hundred cardiac-specific GPCRs can lead to diseased states in the cardiomyocytes, endothelium, and fibroblasts, for example (Tang and Insel, 2004). Despite this large number however, relatively few have been thoroughly characterised in the context of cardiovascular disease and its treatment. The angiotensin 2 receptor is dysregulated in various states of heart disease, causing hypertrophy, fibrosis, and reduced cardiac function post-myocardial infarction; antagonists against this receptor, and angiotensin converting enzyme (ACE) inhibitors, not only reduce cardiac hypertrophy but can reverse tissue

remodelling and even delay heart failure post infarction as well (Paradis *et al.*, 2000; Cohn and Tognoni, 2001). The endothelin-1 receptor functions in a similar manner, and is targeted antagonistically to treat pulmonary arterial hypertension (Liu *et al.*, 2021). There are five adrenergic receptor sub-types, of which the β1 and β2 are expressed most abundantly in healthy cardiac tissue. During heart failure, β1AR is downregulated and β2AR loses efficacy, resulting in cardiac hypertrophy, decreased contractility, and myocyte apoptosis (Xiang and Kobilka, 2003; Tilley and Rockman, 2006).

To summarise, it is clear just how extensively GPCRs are involved in the dysregulation of cellular processes, and the pathophysiology of disease states. It Is impossible to capture the entirety of this topic, however, these examples highlight several points. Firstly, GPCRs are widely involved in all of the leading causes of morbidity and mortality worldwide, and are therefore regarded a high priority for research. Secondly, despite the sheer numbers of GPCRs expressed in diseased tissue, and despite being the target of 40% of all drugs, relatively few GPCRs have been well characterised in the context of these diseases, and therapeutic treatment options largely remain poor overall. The most recent lines of investigation look not only to the GPCRs themselves, but to their G proteins, GPCR receptor kinases (GRKs), and β-arrestin for possible therapeutic interventions (Jiang *et al.*, 2022).

1.2.3 – Ligand Binding and Models of Activation

From discovery of the wide variety of GPCRs to initial models of their resolved structures, focus was next placed on exactly how these receptors are able to facilitate signal transduction from the extracellular microenvironment to an intracellular cascade. Understanding the complex mechanisms behind ligand binding and activation of GPCRs was largely hindered by a lack of structural information under varying conditions (Tehan *et al.*, 2014). Generally, crystal structures fell into one of three categories of GPCR state: inactive when co-crystallised with antagonist or inverse agonist, active with agonist bound, and fully active with crystallisation of the agonist-receptor-G protein complex (Warne and Tate, 2013). Importantly, it is now understood that these three states are linked by intermediate states, and the process of activation is a dynamic transition between conformational states until a stable end point is reached (Weis and Kobilka, 2018).

GPCRs do exhibit a basal activity made feasible by thermal energy in their local microenvironment which can be enough to reach activation, even in the absence of ligand (the apo state); subsequent addition of agonist or antagonist can therefore positively or negatively modify basal activity (Tehan *et al.*, 2014). To probe this characteristic of GPCRs, single mutations have highlighted key residues and further mutants have been created which are constitutively active to shift this basal activity and the energy required to reach activation. Not

only are these key residues thought to be involved in the mechanisms of receptor activation, but so are highly conserved residues as well; the importance of the transmembrane domain was especially highlighted due to its key residues and conserved nature, indicating the transmembrane domain could form part of a common mechanism of GPCR activation (Gether and Kobilka, 1998). Another consideration relates to the wide variety of ligands that GPCRs recognise, and exactly how these interactions are facilitated (Strader *et al.*, 1994).

Initial models of activation were based upon the crystallisation of two key GPCRs in the active state – rhodopsin and the β2-adrenergic receptor (β2AR) (Standfuss *et al.*, 2011; Søren G. F. Rasmussen *et al.*, 2011). The retinal induced activation of rhodopsin displayed significant movement and outward rotation of transmembrane helix 6 (TM6) at the intracellular end, in the range of approximately 10 Å. It was hypothesised that this movement releases a tryptophan residue (W6.48) which resets a hydrophobic region, allowing hydrogen bonds to form and two tyrosine residues (Y5.58 and Y7.53) to occupy an intracellular cavity formed by the rotation of TM6 (Standfuss *et al.*, 2011). Meanwhile, agonist activated β2AR was observed to involve a 2 Å inward movement of TM5 around the proline kink (P5.50) which was thought to disrupt a network of interactions retaining β2AR in the inactive state. This intramolecular disruption was subsequently hypothesised to result in the rotation of TM6, and therefore its outward movement at the intracellular face (Søren G. F. Rasmussen *et al.*, 2011). In both cases, the mechanical rotation and outward movement of TM6 was highlighted as an important step in the stabilisation of active states in these receptors. Since then, the structural differences between the inactive and active states of several more GPCRs have been studied, including the muscarinic M2 and adenosine-2A (A2AR) receptors (Lebon *et al.*, 2011; Kruse *et al.*, 2013), to form a common mechanism of activation. This is thought to be generally applicable to family A GPCRs, but may remain at least somewhat consistent with the other families, despite their differences. Indeed, the ligand-binding mechanisms to induce activation are certainly different – more detail on family B receptors is covered in section 1.3.

This common mechanism of activation not only expands upon those proposed previously, but also remains consistent with data obtained from mutants probing key residues and conserved regions. Briefly, as summarised in Figure 1.2.3, the general mechanism of the activation of GPCRs involves the inward movement of TM5, the upwards movement of TM3, the mechanical rotation of TM6, and the inward movement of TM1 and TM7 (Tehan *et al.*, 2014). Each of these movements are thought to be mainly driven by TM3 and TM6 as these two helices are in contact with each other helix, aside from TM1. This common activation mechanism therefore, is thought to involve a ligand-mediated effect upon TM3 and TM6, by interaction with the major and/or minor binding pockets, or extracellular loops (Rosenkilde *et al.*, 2010). Furthermore, it was believed that the breaking of the ‘ionic lock’ between the highly

conserved R3.50 and an acidic residue in TM6 may be involved (Ballesteros *et al.*, 2001), however, the rearrangement of the hydrophobic core (L3.43, F6.44, ILVM6.40) is now thought to be more important (Tehan *et al.*, 2014). The hydrophobicity of the receptor core not only hinders activation, retaining the inactive state, but its rearrangement and hydration allows for the upward movement of TM3, and outward rotation of TM6 seen in all active structures. Precisely how these interactions influence each other and their exact mechanistic order is yet to be fully understood, however, they ultimately result in the opening of a cytoplasmic cleft in the GPCR, to facilitate binding of the G protein itself.

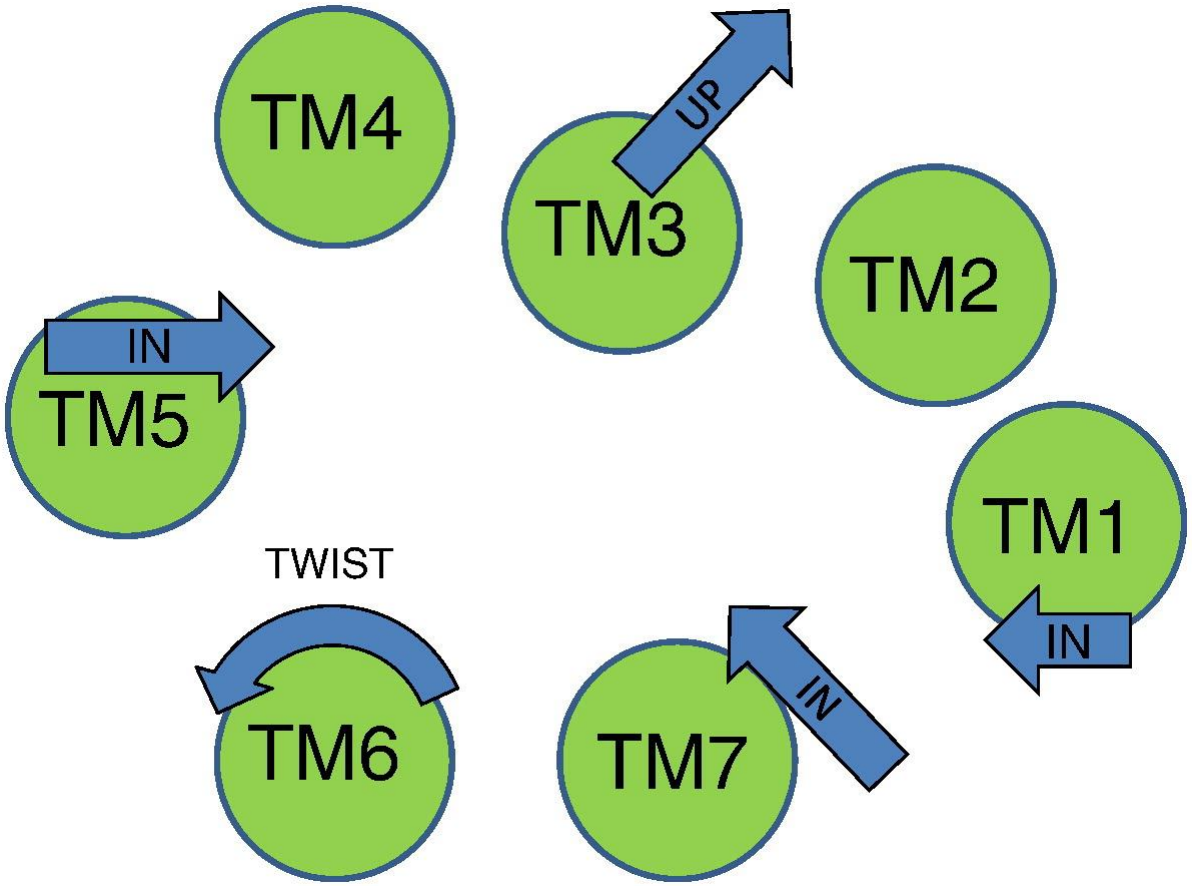

Figure 1.2.3. The Common Mechanism of GPCR Activation. TM5 moves inwards, TM3 moves up, TM6 undergoes an outwards mechanical rotation, and TMs 1 and 7 move inwards. Taken from Tehan *et al*., 2014. Not to scale.

While the importance of TM3, TM6, the ionic lock and the hydrophobic core have garnered considerable evidence, the ligand binding events which initiate activation seem to differ significantly between GPCR sub-families within family A. For example, the β2AR binding of agonist between residues N7.39, D3.32, S5.42 and S5.46 triggers the inward movement of TM5 and TM7. The movement of TM5 allows its proline kink P5.50 to sterically interact with I3.40 on TM3, moving it upwards along with L3.43. This, along with the rotation of F6.44 on TM6, breaks the hydrophobic core of β2AR, allowing the active state to be stabilised (Søren

G. F. Rasmussen *et al.*, 2011; Tehan *et al.*, 2014). Similarly, the photon-induced isomerisation of retinal from the 11-cis to trans form in rhodopsin causes the β-ionone ring to move upwards. This leaves a cavity which is subsequently filled by L3.40, P5.50 and W6.48; again, this triggers TM5 to move inwards at the P5.50 kink, L3.43 on TM3 to move upwards, and L6.44 on TM6 to rotate, breaking the hydrophobic core (Standfuss *et al.*, 2011; Tehan *et al.*, 2014). In the case of the muscarinic M2 receptor, binding of the 'super-agonist' iperoxo between D3.32 and N6.52 induces an inward movement of TM6 at the extracellular face, and therefore an outward movement at the intracellular face. This creates an additional hydrophobic region, drawing TM5 inwards just above P5.50, once again triggering the mechanism towards breaking the hydrophobic core (Haga *et al.*, 2012; Kruse *et al.*, 2013; Tehan *et al.*, 2014). Activation of A2AR differs in that the hydrogen bonding between residues T3.36, N6.55, H7.42 and S7.43 and agonist cause an inward movement of TM7, forcing Y7.49 over L2.46. This causes movement of TM2, creating space for L3.43 to move and, again, break the hydrophobic core (Jaakola *et al.*, 2008; Doré *et al.*, 2011; Lebon *et al.*, 2011; Xu *et al.*, 2011; Tehan *et al.*, 2014). Note that in this case, presence of a G protein is needed to achieve the stable opening of the cytoplasmic binding cleft upon receptor activation. Overall then, GPCR activation is predominantly achieved by the conformational changes of TM3 and TM6, and the rearrangement and hydration of the hydrophobic core, removing the hindrance towards a stable active state.

The concept of ligand binding, to receptor activation, to the binding of G protein gave rise to the 'ligand first' paradigm in which ligand binding must occur first to allow the receptor's cytoplasmic cleft to open for the G protein, relying greatly on random collision (Mafi, Kim and Goddard, 2022). However, an alternative 'pre-coupled' theory proposes that inactive G proteins may already be coupled to the inactive GPCR prior to ligand binding, upon which both GPCR and G protein are activated simultaneously. Long scale (approximately 20 μs) and meta-dynamic molecular simulations were recently used to demonstrate this theory with fifteen, well-studied, family A GPCRs; inactive G proteins, including G_s, G_i, G_o, and $G_{q/11}$, were shown to interact with inactive receptors, exothermically forming this pre-coupled complex (Barducci, Bussi and Parrinello, 2008; Mafi, Kim and Goddard, 2022). In concurrence with the 'ligand first' theory, the movements of TM3 and TM6, along with the ionic lock allow the G protein's α-5 helix to be partially inserted into the receptor at the conserved R3.50, without activation, and in the absence of activating ligand. The binding event itself was therefore shown to activate the pre-coupled GPCR-G protein complex, importantly opening the Gα subunit, which exposes GDP for exchange. Additionally, further arguments for the 'pre-coupled' theory include the observations that agonist alone is often not enough to stabilise the active state (Tehan *et al.*, 2014; Hilger *et al.*, 2020), and many GPCRs possess intrinsic basal

activity without ligand – in fact serotonin receptors 5-HT2A and 5-HT2C have high constitutive activity required for normal physiological function (Leurs *et al.*, 1998; Berg *et al.*, 2005).

In effect, there exists evidence for both the 'ligand first' and 'pre-coupled' theories, supporting the possibility that within a given cell population, some GPCRs may be pre-coupled and ready to signal with bias, leaving others to react to ligands more dynamically through the various G proteins and subsequent signalling pathways.

1.2.4 – GPCRs and their Signalling Pathways

It is now well known that GPCRs mediate signal transduction by the intracellular recruitment of G proteins, and subsequent propagation of downstream cascades via effector enzymes such as adenylate cyclase (Syrovatkina *et al.*, 2016). However, several other factors also further increase the complexity of this cellular process, including interactions with GPCR kinases (GRKs), β-arrestins, the concept of biased agonism towards certain pathways, receptor trafficking and endocytosis, and re/desensitisation before lysosomal destruction (Jiang *et al.*, 2022).

In essence, G proteins are heterotrimeric complexes composed of an α, β and γ subunit for which the human genome encodes eighteen, five and twelve subunit versions, respectively (Syrovatkina *et al.*, 2016). The βγ subunits interact strongly, and are considered as one functional unit which is anchored to the lipid membrane, whereas the α subunit dissociates from the βγ complex upon activation. Crucially, G proteins are classified based upon their Gα subunit type, which are further grouped into four main families, $G\alpha_s$, $G\alpha_i$, $G\alpha_{q/11}$ and $G\alpha_{12/13}$, each of which regulate their downstream targets differently (Gilman, 1987; Jiang *et al.*, 2022). A summary of G protein subunit characteristics is provided in Table 1.2.3. The Gα subunit is composed of three domains – an N-terminal helical domain, and the conserved α-helical and Ras-like GTPase domains (Lambright *et al.*, 1996; Chen and Manning, 2001). The N-terminus facilitates the interaction with the βγ complex, while the GTPase domain possesses a conformational plasticity – it is highly dynamic while bound to GDP, but rigid while bound to GTP, correlating with the need to expose GDP by conformational change during activation (Mazzoni, Malinski and Hamm, 1991; Goricanec *et al.*, 2016). Furthermore, the α-helical domain not only enhances Gα affinity for guanine nucleotides, but is also thought to increase the rate of GTP hydrolysis (Markby, Onrust and Bourne, 1993; Warner *et al.*, 1998).

Table 1.2.4. The Classification of G Protein Subunits. Adapted from Syrovkatina *et al.*, 2016.

Family	Members	Expression	Function
α_s	α_s	Ubiquitous	Stimulates adenylate cyclase to make cAMP from ATP
	α_{olf}	Olfactory neurons	
α_i	α_{i1}	Widely	Generally inhibits the production of cAMP from ATP Members may have different effector molecules (the target of G_t is cGMP phosphodiesterase, for example)
	α_{i2}	Widely	
	α_{i3}	Ubiquitous	
	α_{oA}	Neurons	
	α_{oB}	Neuroendocrine	
	α_{t1}	Retinal rods, taste cells	
	α_{t2}	Retinal cones	
	α_g	Brash and taste cells	
	α_z	Neurons, platelets	
α_q	α_q	Ubiquitous	Stimulates phospholipase C-β to make IP3 and DAG
	α_{11}	Ubiquitous	
	α_{14}	Kidney, liver, lung cells	
	α_{15}	Haematopoietic cells	
	α_{16}	Haematopoietic cells	
α_{12}	α_{12}	Ubiquitous	Rho GTPase signalling
	α_{13}	Ubiquitous	
βγ	β_1	Widely, retinal rods	Negatively regulates Gα Various diverse signalling effects
	β_2	Widely	
	β_3	Widely	
	β_4	Widely	
	β_5	Brain	
	γ_1	Retinal rods, brain	
	γ_2	Ubiquitous	
	γ_3	Brain	
	γ_4	Brain and other tissues	
	γ_5	Widely	
	γ_7	Widely	
	γ_8	Olfactory epithelium	
	γ_9	Olfactory epithelium	
	γ_{10}	Widely	
	γ_{11}	Widely	
	γ_{12}	Widely	
	γ_{13}	Brain, taste buds	

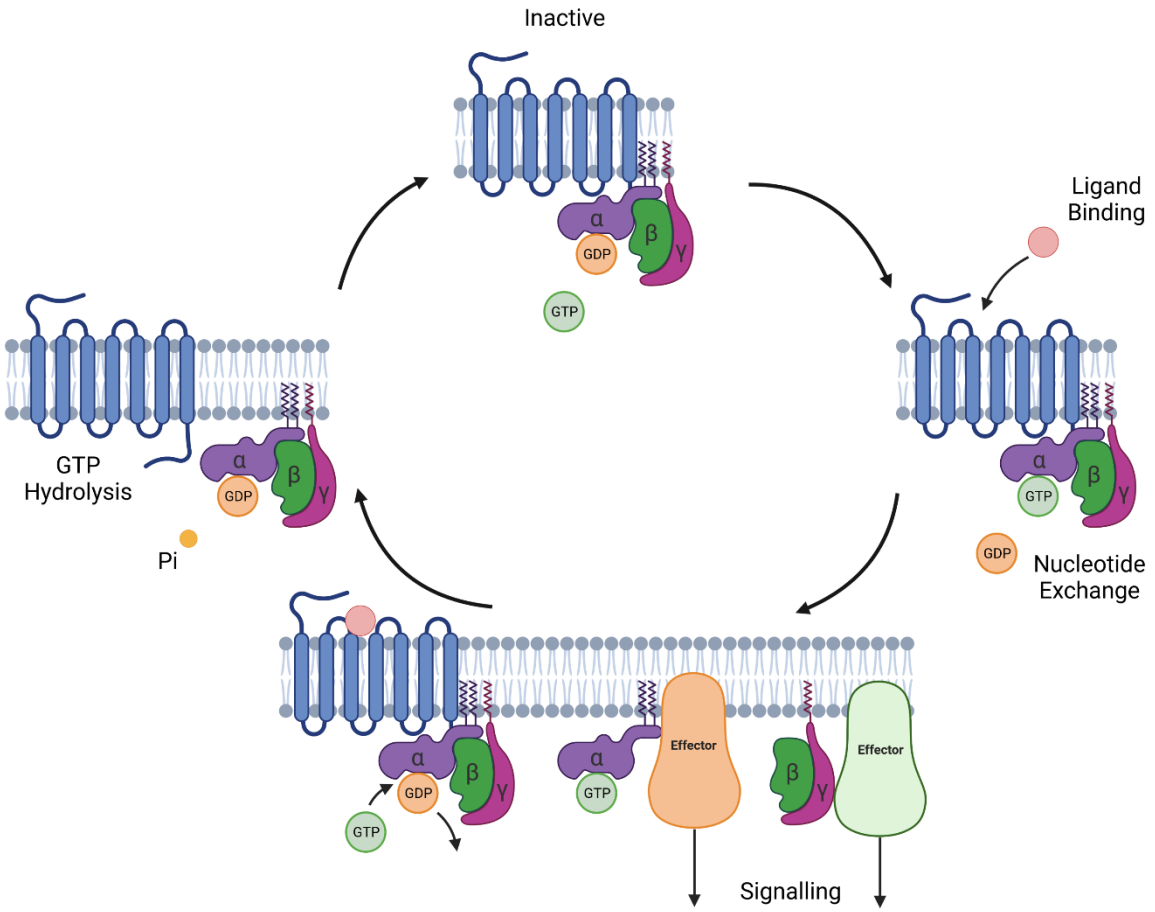

Figure 1.2.4.1. The Cycle of Heterotrimeric G Protein Activation. The ligand-induced activation of a GPCR facilitates GDP-GTP exchange within the G protein's Gα subunit, causing disassociation of the heterotrimer. Both the Gα-GTP and Gβγ subunits can then signal independently through various effector enzymes; greater signalling detail is shown in Figure 1.2.4.2. Finally, the Gα subunit's inherent hydrolytic activity converts its GTP back to GDP, releasing a free inorganic phosphate ion (Pi). This allows for reassociation of the heterotrimer, G protein turnover, and the activation cycle may begin again. Created with BioRender.com.

The mechanism of G protein activation, illustrated in Figure 1.2.4.1, first begins with ligand binding at the receptor, where the GPCR itself then acts as a guanine nucleotide exchange factor (GEF), making GDP available for exchange to GTP (Mahoney and Sunahara, 2016). As discussed in 1.2.2, this is made possible by the cytoplasmic cavity formed by the movement of receptor TM6, forming the location of interaction with the G protein α-5 helix. GDP-GTP exchange occurs due to the allosteric interference of the GDP binding site during conformational change, after which GTP associates with the Gα Ras-like domain due to GTP's naturally high cytoplasmic concentration (McKee *et al.*, 1999). It is this GTP binding which makes the heterotrimer unstable, causing dissociation between the now distinct Gα-GTP and Gβγ complexes (McCudden *et al.*, 2005).

Once separated from each other, each of the Gα and Gβγ complexes activate different downstream effector molecules, leading to different signalling profiles, especially with regards to the various Gα sub-families (Figure 1.2.4.2). $G\alpha_s$ and $G\alpha_i$ both regulate the effector enzyme adenylate cyclase which produces cyclic adenosine monophosphate (cAMP) from ATP; $G\alpha_s$ stimulates the enzyme which increases the intracellular concentration of cAMP, while $G\alpha_i$ reduces levels of cAMP as it is inhibitory (Gilman, 1987). cAMP is an important second messenger molecule which subsequently activates protein kinase A enzymes (PKA), affecting various cellular processes. $G\alpha_{q/11}$ activates the effector enzyme phospholipase C-β which produces inositol trisphosphate (IP_3) and diacylglycerol (DAG) from the hydrolysis of phosphatidylinositol-4,5-bisphosphate (PIP_2) (Simon, Strathmann and Gautam, 1991). DAG is membrane-bound and activates protein kinase C enzymes (PKC), while IP_3 is a soluble second messenger, promoting the release of Ca^{2+} from the lumen of the endoplasmic reticulum (Berridge, 2016). Finally, $G\alpha_{12/13}$ activate various targets including RhoGEF molecules which regulate small GTPases in the Rho family, also affecting various cellular processes via transcription factors (Chen *et al.*, 2005). Certainly, the Gβγ complex provides negative regulation of G protein signalling overall, as GDP-GTP exchange is prevented in the heterotrimeric complex (Ford *et al.*, 1998; Mahon *et al.*, 2006). However, as well as this, Gβγ also exerts effects on various downstream molecules, including GPCR kinases (GRKs) (Pitcher *et al.*, 1992), ion channels (Wickman *et al.*, 1994) and phospholipases (Exton, 1996), for example.

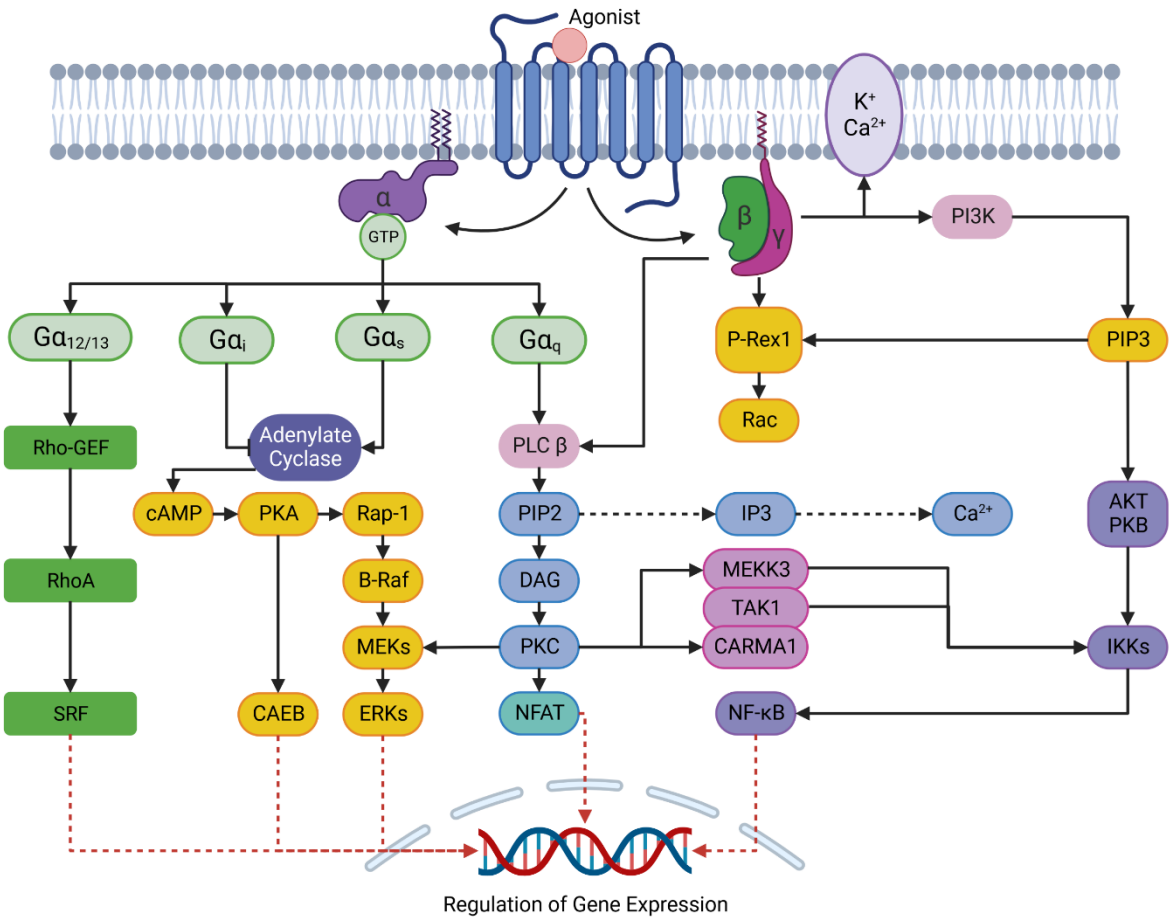

Figure 1.2.4.2. The Signalling Pathways of GPCRs. Upon GPCR, and subsequent G protein activation, the Gα-GTP subunit acts upon various signalling pathways, depending on its family. $G\alpha_s$ activates adenylate cyclase, producing cAMP which results in the activation of protein kinase A (PKA) towards CAEB and MEK/ERK cascades; conversely, $G\alpha_i$ inhibits this process. $G\alpha_{q/11}$ activates phospholipase C-β (PLC β) which activates protein kinase C (PKC) through IP3 and DAG. Finally, $G\alpha_{12/13}$ activates the RhoA GTPase, through guanine nucleotide exchange, towards transcription factors including SRF. In addition, the Gβγ subunit is also able to activate PLC β, various ion channels, and PI3K which propagates signalling through the AKT/PKB pathway towards NF-κB. Ultimately, each of these pathways affect the regulation of gene expression for many cellular and physiological processes. Created with BioRender.com.

Once signalling has occurred, the GTPase activity of Gα will convert the bound GTP back to GDP, terminating Gα's signalling capability, possibly with the aid of GTPase activating proteins (GAPs) which are classed as regulators of G protein signalling (RGS) (Druey *et al.*, 1996; Shi *et al.*, 2001). Simultaneously, Gβγ can reassociate with a Gα subunit, reforming the heterotrimer and terminating their signalling capability as well. Furthermore, the inhibition of GPCR and G protein signalling can also be the result of GPCR kinase (GRK) activity, and/or by the association of β-arrestin (Jiang *et al.*, 2022).

GPCR receptor kinases (GRKs) are a group of serine/threonine kinases which phosphorylate the third intracellular loop (ICL3) or C terminus of GPCRs to desensitise them, and ultimately downregulate and terminate agonist-induced signalling (Pfleger, Gresham and Koch, 2019). Each of the seven GRK isoforms are expressed in different tissue types, and unlike other kinases, do not require autophosphorylation for their complete activity (Premont *et al.*, 1999; Pearce, Komander and Alessi, 2010). While each isoform possesses unique N- and C-terminal characteristics, their overall structures are relatively conserved (Pitcher, Freedman and Lefkowitz, 1998). Each has a catalytic kinase domain within an RGS homology domain which can also interact with $G_{\alpha q}$ subunits, free $G_{\beta\gamma}$ subunits, and membrane lipids (Benovic *et al.*, 1989; Siderovski *et al.*, 1996; Sallese *et al.*, 2000; Homan and Tesmer, 2015). RGS stands for regulator of G protein signalling which typically enhance the intrinsic GTPase activity of Gα subunits (Stewart and Fisher, 2015). The GRK α-helical N-terminal domain contains an anti-parallel β-sheet which is thought to stabilise the kinase by hydrophobic membrane interactions. In contrast, the C-terminal domain, which is the primary binding site, is composed of several α-helices (McClendon *et al.*, 2014). It is also thought that interactions between the N and C termini of the GRK form a binding site for activated GPCRs; mutations at these locations result in the loss of full GRK phosphorylation activity, and similarly, mutation of GPCR phosphorylation sites leads to delayed desensitisation (Bouvier *et al.*, 1988; Palczewski *et al.*, 1993; Boguth *et al.*, 2010; Huang *et al.*, 2011).

As a consequence of GPCR C-terminal phosphorylation, the recruitment and association of β-arrestin can occur which promotes the clathrin-mediated internalisation of the GPCR signalling complex into the membrane of an endosome (Lefkowitz and Shenoy, 2005; Gurevich and Gurevich, 2014). β-arrestin achieves this by forming a salt bridge between a lysine/arginine motif on its N terminus and the phosphorylated sites on the GPCR C terminus; this causes structural changes within β-arrestin, activating this adaptor protein (Shukla *et al.*, 2013; Chen *et al.*, 2017; Ahn *et al.*, 2020). The recruitment of β-arrestin typically uncouples heterotrimeric G proteins from the receptor, due to its interaction with intracellular loop 3, causing complete termination of signalling following GRK-mediated desensitisation (Lohse *et al.*, 1992; Heitzler *et al.*, 2012). However, if β-arrestin interacts with the receptor C terminus only, the GPCR-G protein complex can remain intact, enabling the GPCR to continue signalling from the endosome itself, after internalisation; indeed, intracellular membranes such as endosomes and the Golgi network have been shown to be rich in G proteins (Ferguson *et al.*, 1996; Thomsen *et al.*, 2016; Nguyen *et al.*, 2019; Crilly and Puthenveedu, 2021). Interestingly, β-arrestin does not solely terminate GPCR signalling, but also has the potential to switch signalling profiles in a receptor-dependent or -independent manner, activating

pathways such as Src/EGFR and MEK/ERK which affect apoptosis, cell growth, and various other cellular processes (Luttrell *et al.*, 1999; Ahn *et al.*, 2020; Jiang *et al.*, 2022).

In essence, GPCRs are able to modulate a wide range of cellular processes in response to endogenous and exogenous ligands, and their role in cell signalling and homeostasis leads to their involvement in many human disease states. The continued study of the complex molecular mechanisms behind GPCRs and their downstream effectors, illustrated in Figure 1.2.4.2, are of utmost importance, as greater understanding will undoubtedly influence the design of novel drugs against dysregulation and pathophysiology.

1.3 – The Six GPCR Families

The differences between each of the six main GPCR families will now be discussed, with detail on their structural and conformational changes during activation, as well as describing a few key examples from each family. Overall, while two-thirds of human hormones and one-third of all drugs target GPCRs in the A, B1, C and F families, relatively little is known about the structural mechanisms of the non-family A receptors, in comparison to them.

1.3.1 – Family A

The family A GPCRs are by far the largest group, with 719 of these receptors in humans, and are able to bind a wide range of ligands including lipids, peptides, proteins, photons and small soluble molecules (Congreve *et al.*, 2020; Sutkeviciute and Vilardaga, 2020). Despite this wide range of agonism, family A GPCRs undergo a common mechanism of activation resulting in the movement of TM helices, including the characteristic outward movement of TM6, as described in section 1.2.3. This movement of TM helices links each of the conserved CWxP, DRY, Na^+ pocket, NPxxY and PIF motifs to form a transmembrane network of interactions between the ligand-binding and G protein-binding pockets, leading to G protein activation (Zhou *et al.*, 2019). As family A is so large, many of its members are still considered orphan with no known ligand, and some of these receptors even lack soluble ligand as is the case with the protease-activated receptors (PARs). Very interestingly, the orphan receptor GPR52 has been found to constitutively self-activate by folding ECL2 into its orthosteric binding site, offering a possible explanation for why so many orphan receptors exist, with no apparent ligand (Lin *et al.*, 2020). Overall, the family A GPCRs contain many well-characterised drug targets, and are split into the rhodopsin, and non-rhodopsin receptors.

Rhodopsin is a visual pigment of retinal photoreceptor cells, which converts photons into chemical signals via the G_t protein transducin, allowing organisms to detect and process light

(Zhou, Melcher and Xu, 2012). The bovine rhodopsin was the first GPCR structure to be resolved by Palczewski *et al.* in 2000, and has been used widely as a template for understanding the activation and signalling of GPCRs in general, not solely limited to rhodopsin or family A (Palczewski *et al.*, 2000). Since then, many more bovine rhodopsin structures, and four human rhodopsins, have been resolved in various conformational states and co-crystallised with various ligands, further strengthening understanding of the ligand-induced activation of GPCRs. Indeed, the characteristic movement of TM6 was first predicted four years earlier by Farrens *et al.* 1996, with the use of electron paramagnetic resonance (EPR) and site-directed spin-labelling of disulphide-linked cysteine residues (Farrens *et al.*, 1996). Subsequently, based upon the first rhodopsin structures, it was found that the highly conserved Pro267 in TM6 causes it to bend upon activation, tilting the helix away from the TM bundle, creating an intracellular G protein binding pocket approximately 14 Å in diameter (Scheerer *et al.*, 2008; Choe *et al.*, 2011; Standfuss *et al.*, 2011). Though these opsins created a paradigm for GPCRs in the early 2000s, the reliance on its structure has since diminished over time as many more varied GPCR structures have been resolved, highlighting differences which confine the opsins to their own unique sub-family within family A.

Following the resolution of the rhodopsin crystal structure, that of the β-adrenoceptor was obtained next, leading to the sub-family of the 42 aminergic GPCRs of family A. The aminergic receptors are further split into several more sub-families including the adrenergic, dopaminergic, histaminergic, muscarinic and serotoninergic receptors, with over 50 crystal structures resolved thus far (Strosberg, 1993; Missale *et al.*, 1998; Parsons and Ganellin, 2006; Nichols and Nichols, 2008; Kruse *et al.*, 2014). As several active, intermediate and inactive conformational states were resolved, the aminergic receptors played a significant role in the foundational investigations of GPCRs alongside rhodopsin, and translated successfully into therapeutic research; the aminergic GPCRs are the targets of approximately 25% of all current drugs (Lee, Basith and Choi, 2018). Notable examples include β-adrenoceptor antagonists to treat anxiety and migraine (β-blockers), histamine H1 receptor antagonists to treat allergic reactions (anti-histamines), the muscarinic acetylcholine receptors are targeted for neurodegenerative disorders such as Alzheimer's and Parkinson's disease, and finally the serotonin receptors are important targets for psychiatric disorders including depression and anxiety, for example. Overall, much of the molecular detail of GPCRs has been garnered from this particularly important sub-family of aminergic receptors.

Family A also contains the nucleotide-like, or purinergic receptor sub-family, including the P1 and P2 receptors which are preferentially activated by adenosine and various other nucleotides including ATP, di/triphosphates, purines and pyrimidines (Kaebisch *et al.*, 2015). The most well-studied purinergic receptors are the adenosine A1, A2A, A2B and A3 receptors,

which regulate sleep and arousal in the central nervous system, in response to the neuromodulator adenosine; while the A1 and A3 receptors couple predominately to $G\alpha_{i/o}$, both A2A and A2B are preferentially biased towards $G\alpha_s$ (Dunwiddie and Masino, 2001; Lee, Basith and Choi, 2018). The adenosine A2A receptor (A2AR) is by far the most well-characterised purinoceptor, with both active and inactive structures in complex with various ligands such as the antagonist caffeine and numerous inverse agonists, revealing structural insights into A2AR activation and modulation (Jaakola *et al.*, 2008; Doré *et al.*, 2011; Lee, Basith and Choi, 2018). Interestingly, it has been proposed that the active A2AR structures are not actually fully activated and should be considered as active-intermediates; co-crystallisation with a high-affinity ligand and G protein is thought to be required for complete activation with the characteristic intracellular movement of TM6 14 Å away from the TM bundle (Carpenter *et al.*, 2016).

While the ligands of many family A GPCRs are small soluble molecules, there are also numerous examples which are activated by peptides, including the angiotensin II, chemokine, endothelin, neurotensin, opioid, orexin, and protease-activated receptors (Rajagopalan and Rajarathnam, 2006; White *et al.*, 2012; Zhang *et al.*, 2012; Pasternak, 2014; H. Zhang *et al.*, 2015; Shihoya *et al.*, 2016; Yin *et al.*, 2016). Many of the ligand binding and activation mechanisms of these peptide binding family A GPCRs were poorly understood prior to the resolution of their crystal structures, which have highlighted some key differences to the rest of family A. Their activation mechanisms seem to correlate with those exemplified by rhodopsin, however, the peptide-binding receptors tend to have larger, more open binding pockets which do not penetrate as deeply towards the TM bundle, when compared to other family A members (Manglik *et al.*, 2012; Huang *et al.*, 2015). In addition, some of these peptide ligands have been observed to adopt a dual-binding mechanism between the receptor N terminus and ECLs, and it's TM bundle, in a similar manner to the peptide hormones of family B GPCRs (Lee, Basith and Choi, 2018). Interestingly, the more open binding pockets give a conformational plasticity which in turn allows receptors such as CXCR4 and CCR5 to accommodate various ligands of different classes, expanding their functional repertoire.

Finally, several members of family A are able to bind lipids as their cognate ligand, such as the lysophospholipid, free fatty-acid and cannabinoid receptors (Hanson *et al.*, 2012; Srivastava *et al.*, 2014; Chrencik *et al.*, 2015; van Jaarsveld, Houthuijzen and Voest, 2016). These receptors are thought to be of great therapeutic relevance due to their involvement in diabetes, cancer and multiple sclerosis, and have therefore received considerable attention, with several compounds reaching clinical trials. It is thought that these lipid-binding GPCRs recognise their ligands in a highly selective manner, as even minor modifications to these lipids often abolishes their activity completely (Houthuijzen *et al.*, 2014). An interesting

observation with these lipid-binding GPCRs highlights the spherical binding site of the lysophosphatidic acid 1 receptor versus the more linear and rigid binding pocket of the sphingosine 1-phosphate receptor, enabling the recognition of a more and less diverse range of ligands, respectively (Lee, Basith and Choi, 2018).

1.3.2 – Family B

The family B GPCRs are split into three groups: the B1 secretin-like receptors, the B2 adhesion receptors, and the little-known B3 methuselah-like sub-family; they are found in all vertebrates and some non-vertebrates such as *Caenorhabtidis elegans* and *Drosophila melanogaster*, but family B GPCRs are not found in fungi, plants or prokaryotes (Harmar, 2001; Schwartz and Frimurer, 2017). The B1 sub-family bind a series of peptide hormones including calcitonin, CGRP, CRF, glucagon, GLP, PTH and secretin, via their larger extracellular domain, coupling preferentially to $G\alpha_s$ which stimulates adenylate cyclase activity (Couvineau and Laburthe, 2012; Poyner and Hay, 2012). Meanwhile, the 33 adhesion receptors of sub-family B2 are characterised by a variety of ECD motifs needed for cell-cell adhesion and cell-matrix interaction; examples include epidermal growth factor-like, laminin and cadherin domains. In addition, most adhesion receptors also contain a conserved autoproteolysis domain, creating N- and C-terminal fragments often needed for adhesion receptor signalling (Rosa *et al.*, 2021). Finally, the B3 sub-family is named after the *Drosophila* gene *methuselah* which is involved in embryonic development, lifespan and stress responses in *Drosophila*, though these GPCRs have no structures resolved as of yet (Ja *et al.*, 2009; de Mendoza, Jones and Friedrich, 2016). Consequently, focus is placed upon the sub-family B1 receptors in this thesis, the calcitonin family of receptors in particular.

The calcitonin peptide family include calcitonin (CT), amylin (AMY), calcitonin gene-related peptide (CGRP), adrenomedullin (AM), and adrenomedullin 2/intermedin (AM2), which signal through two family B GPCRs – the calcitonin receptor (CTR) and calcitonin receptor-like receptor (CLR), as detailed in Table 1.3.2. In addition, a group of three accessory proteins called the receptor activity modifying proteins (RAMPs) are able to influence these receptors in various ways, and are covered in greater detail in section 1.4. In essence, these three RAMPs influence the calcitonin family of receptors via trafficking patterns, ligand selectivity and direct allostery, acting as pharmacological switches (Kotliar *et al.*, 2023). While this family of peptides only share limited homology by sequence, they are structurally similar, and are involved in a range of biological processes (Hay *et al.*, 2018). The cognate ligand for CTR is calcitonin, but switches over to amylin when CTR associates with a RAMP; calcitonin is involved in calcium homeostasis between blood plasma and bone, while amylin is a satiety hormone involved in nutrient uptake, two very different biological functions promoted simply

by RAMP association (Findlay and Sexton, 2004; Hay *et al.*, 2015). Conversely, RAMP association with CLR produces three receptors for the potent vasodilators CGRP, AM and AM2, with CLR alone having no currently known function – co-expression and association with a RAMP is therefore required for CLR functionality (Hinson, Kapas and Smith, 2000; Russell *et al.*, 2014). Again, more detail on the calcitonin receptors and the RAMPs is given in section 1.4.

Table 1.3.2. The Calcitonin Family of GPCRs and their Peptides. Each of the receptor-RAMP combinations alters ligand selectivity and extent of pharmacology, demonstrated by the pEC50 values for αCGRP-stimulated cAMP production, as an average summary of 46 publications (Hay *et al.*, 2018). CLR alone has no currently known function without a RAMP association. CTR = calcitonin receptor; CLR = calcitonin receptor-like receptor; CT = calcitonin; AMY1 = amylin 1; AMY2 = amylin 2; AMY3 = amylin 3; CGRP = calcitonin gene-related peptide; AM = adrenomedullin; AM2 = adrenomedullin 2/intermedin.

		Receptor Alone	RAMP1	RAMP2	RAMP3
CTR	Ligand Function pEC50	CT Ca^{2+} homeostasis 7.33	AMY1 Nutrient intake 9.28	AMY2 Nutrient intake 7.52	AMY3 Nutrient intake 8.10
CLR	Ligand Function pEC50	- - -	CGRP Vasodilation 9.61	AM Vasodilation 7.16	AM2 Vasodilation 7.18

As the calcitonin family of peptides and their receptors have been well-characterised, much of the binding mechanisms of family B GCPRs have also been elucidated from these receptors. The unique structural characteristic of family B GPCRs is a two-step binding process for their peptide hormone ligands, which contrasts with the majority of other GPCRs. This involves a rapid interaction between the peptide C terminus, and receptor extracellular domain, followed by a much slower and more complex interaction between the peptide N terminus and receptor transmembrane domain, leading to receptor and subsequent G protein activation (Liang, Belousoff, Zhao, *et al.*, 2020). While the C termini of the peptide hormones can adopt either an extended helical or unfolded conformation, the N termini penetrating into the receptor core seems to be a conserved activation mechanism for family B receptors. Indeed, several conserved motifs have been identified and attributed to this process, along with a larger displacement of TM6, approximately 20 Å, compared to the 6 to 14 Å typical of family A receptors; this is thought to be due to a significant kink in the middle of TM6, stabilised by the conserved NPGQ motif (Yin *et al.*, 2017; Zhao *et al.*, 2019; Ma *et al.*, 2020; Sutkeviciute and Vilardaga, 2020). Importantly, much of this information was garnered from structures in complex with agonists which are unbiased towards downstream signalling pathways, however, GLP-1R and PTH1R structures have revealed possible mechanisms for biased agonism via the mobility of the extracellular domain and interactions with the extracellular

loops (Zhang *et al.*, 2017; Lei *et al.*, 2018; Liang, Khoshouei, Glukhova, *et al.*, 2018; Sarkar *et al.*, 2019; Clark *et al.*, 2020; Zhao *et al.*, 2020). Suffice to say, the biased agonism of GPCRs remains to be a crucial subject of research endeavours.

1.3.3 – Family C

The family C GPCRs, also known as the metabotropic glutamate receptor-like family, consist of at least six sub-families including: the calcium-sensing receptor (CaSR), γ-aminobutyric acid type B receptors (GABA-B), metabotropic glutamate receptors (mGluR1-8), vomeronasal type 2 pheromone receptors (V2R), type 1 taste receptors (TAS1R), the promiscuous L-α-amino acid receptor (GPRC6A), and several orphan receptors as well. The family C GPCRs have a characteristically large extracellular domain housing a 'Venus flytrap module' which contains the orthosteric binding site, and a cysteine-rich domain, with the exception of GABA-B (Kunishima *et al.*, 2000; Geng *et al.*, 2013; Pin and Bettler, 2016). Uniquely, this family of GPCRs exist as obligate homo- or heterodimers; GABA-B1 and GABA-B2 receptors form heterodimers through interactions between their Venus flytrap modules and C-terminal coiled-coils, while the CaSR and TAS1R heterodimers, and homodimeric mGluR receptors form through their Venus flytraps and disulphide bonds between their cysteine-rich domains (Muto *et al.*, 2007). Furthermore, as the family C GPCRs bind a variety of ligands, they are involved in skeletal metabolism, calcium homeostasis, neurological synaptic transmission and the sense of taste, to name a few.

The structures of family C extracellular domains have revealed that agonist binding causes the Venus flytrap module to close which rearranges the interface between the two monomers, bringing them closer in proximity, considered to be a defining characteristic of family C GPCR activation (Kunishima *et al.*, 2000). In addition to this, full-length structures also show agonist-induced movement within the dimer TM helices, with TM6 forming an interface, and a rigid ECL2 relaying the conformational change from the ECD to the TM (Xue *et al.*, 2015; Koehl *et al.*, 2019). However, despite these helical rearrangements, agonist-induced mGluR structures do not display the characteristic outward movement of TM6 which may suggest that resolution of family C GPCRs in complex with G proteins is required to fully understand their conformational changes upon activation (Doré *et al.*, 2014; Christopher *et al.*, 2015, 2019).

1.3.4 – Family D

The family D GPCRs are fungal mating pheromone receptors, consisting of the Ste2 and Ste3 sub-families. These receptors are activated by the peptide mating pheromones, alpha-factor and a-factor, which activate the mitogen-activated protein kinase (MAPK) signalling cascade to induce fungal cell-cell fusion and the formation of a diploid zygote (Brown *et al.*, 2018;

Velazhahan *et al.*, 2021). While these GPCRs do indeed share the common feature of seven transmembrane helices, and share common GPCR signalling pathways, family D does not share any significant similarity of sequence with the vertebrate GPCR families. For brevity, and due to a lack of resolved structures, family D GPCRs are not considered in this thesis.

1.3.5 – Family E

The family E GPCRs consist of cAMP receptors found in 'slime mold' amoebas which live in soil, such as the *Dictyostelium discoideum* and *Polysphondylium pallidum*. These cAMP receptors are involved in the chemotactic movement of these cells, and their subsequent aggregation into a multicellular organism (Manahan *et al.*, 2004; Kawabe and Schaap, 2022). Similarly, the family E GPCRs also do not share significant sequence similarity with the vertebrate GPCR families, which, along with the lack of resolved structures means family E are also not considered in this thesis.

1.3.6 – Family F

The family F GPCRs are composed of the 10 frizzled isoforms (FZD), and the closely related smoothened receptors (SMO). The frizzled receptors are activated by the wingless/int-1 (WNT) lipoglycoproteins, and are preferentially biased towards $G\alpha_s$, but can signal through $G\alpha_i$ and $G\alpha_q$ as well. Meanwhile, the smoothened receptors transduce the hedgehog (HH) signalling pathway to induce the transcription of genes such as PTCH, Gli1 and Bcl-2, etc. Both the WNT and HH pathways are intrinsically involved in many developmental processes including embryogenesis, cell proliferation, stem cell renewal and tumorigenesis, mediated via the family F GPCRs (Kozielewicz, Turku and Schulte, 2020).

Structural insights into these receptors are mainly based upon the SMO receptors, with only three FZD structures resolved thus far. Their activation seems to be largely similar to those of family A, with the outward movement of TM6 disrupting a conserved R/K6.32 interaction with TM7 which retains the inactive state of family F GPCRs (Deshpande *et al.*, 2019; Qi *et al.*, 2019; Sutkeviciute and Vilardaga, 2020). In contrast to family A however, their coupling to $G\alpha_i$ differs; the α-5 helix of $G\alpha_i$ inserts into the core of SMO, parallel to the transmembrane bundle, and is tilted by 5 degrees when compared to family A coupling to $G\alpha_i$ (Wright *et al.*, 2019).

1.4 – The Receptor Activity Modifying Proteins (RAMPs)

1.4.1 – The Discovery of the RAMP Accessory Proteins

The gene encoding the family B calcitonin receptor (CTR) was first isolated and cloned from porcine renal epithelial cells, eventually leading to the discovery of a close homologue called the calcitonin receptor-like receptor (CLR), sharing 55% sequence homology between their human counterparts (Lin *et al.*, 1991; Flühmann *et al.*, 1995). As a result, the CLR receptor was ultimately proposed to be a potential cognate receptor for the calcitonin gene-related peptide (CGRP) (Aiyar *et al.*, 1996). While CTR was observed to bind salmon calcitonin with high affinity, stimulating adenylate cyclase activity, the CLR receptor's response to CGRP was unpredictable and seemed to be cell line dependent (Flühmann *et al.*, 1997). Several cell lines, including HEK293 and the *Xenopus* oocytes, were therefore hypothesised to contain an unknown factor required for CGRP-mediated signalling via the CLR receptor.

Consequently, cDNA was subsequently isolated from SK-N-MC neuroblastoma cells which led to the discovery of a 148 amino acid, single-transmembrane protein termed receptor activity modifying protein 1 (RAMP1) (McLatchie *et al.*, 1998). The CLR receptor was indeed proposed to be the cognate receptor for CGRP, with the RAMP1 accessory protein required for CLR's trafficking to the cell membrane; the apparent lack of signalling in certain cell lines was therefore attributed to a potential lack of RAMP1 expression. This was considered a revolutionary discovery, as it suggested that GPCRs did not solely function as independent molecules but could be included in multifaceted signalling complexes, a notion now widely accepted. In time, homology searches additionally identified two more human receptor activity modifying proteins, RAMP2 and RAMP3; in fact, RAMPs have been identified in 53 species thus far, most of which possess three distinct RAMP genes, some species of fish possess five (McLatchie *et al.*, 1998; Parameswaran and Spielman, 2006; Klein, Matson and Caron, 2016).

1.4.2 – What are RAMPs?

The receptor activity modifying proteins are bitopic, single transmembrane accessory proteins, unique to vertebrate organisms, and are expressed ubiquitously in all human tissue types (Kotliar *et al.*, 2023). While each of the three RAMPs share the same overall topology, with a structured extracellular domain, single transmembrane α-helix, and cytoplasmic tail consisting of just nine residues, the RAMPs share only 30% homology between their primary structures (Parameswaran and Spielman, 2006). The RAMP extracellular domain is composed of a bundle of three α-helices, approximately 90 to 100 residues in length; the ECD structure of RAMP1 and RAMP3 are more similar, with that of RAMP2 being 26 residues longer (Kotliar *et al.*, 2023). This further correlates with the higher sequence similarity shared between

RAMP1 and RAMP3, revealed by bioinformatic analysis; it is thought that RAMP1 and RAMP3 co-evolved with a subset of GPCRs distinct to that of RAMP2, offering possible explanation for this difference (Barbash *et al.*, 2017).

Another notable feature of RAMP extracellular domains are their post-translational modifications, which are thought to influence trafficking to the cell surface. RAMP2 and RAMP3 have one and four glycosylation sites, respectively, while RAMP1 has none. These differences are highlighted by the CLR-independent expression of RAMP2 and RAMP3 at the cell surface in COS-7 cells; mutations to introduce or remove these sites enabled or disabled trafficking activity, respectively (Flahaut, Rossier and Firsov, 2002; Husmann *et al.*, 2003). The ECD of RAMP1 instead contains a C-linked mannosylation motif at tryptophan 56 which regulates and/or enhances the stability of RAMP1 (Crine and Acharya, 2022; Mizuta *et al.*, 2023).

Despite their relatively short length, the RAMP C-terminal tails contain particularly important motifs. Firstly, RAMP1 possesses the QSKRT endoplasmic reticulum retention signal; interestingly, this is overridden with RAMP1's interaction with CLR, thereby promoting movement of the heterodimer to the cell surface (Steiner *et al.*, 2002). In addition, the serine-lysine (SK) motif within this signal is conserved across all three RAMPs, though its purpose has not yet been defined completely. Through mutation and truncation, the SK motif has been further attributed to cell surface expression, and the negative regulation of receptor internalisation as well (Kuwasako *et al.*, 2006). Uniquely, the C-terminal tail of RAMP3 contains a PDZ recognition motif which mediates the internalisation of CLR; PDZ domains are common scaffolding structures found in the signalling proteins of many organisms (Bomberger, Parameswaran, *et al.*, 2005; Bomberger, Spielman, *et al.*, 2005; Cottrell *et al.*, 2007; Klein, Matson and Caron, 2016). Overall, while the RAMP C-terminal tails possess important functions, they are not thought to be directly involved in GPCR-RAMP signalling, but may exert indirect effects via their interactions with intracellular components such as G proteins and their coupling events (Udawela *et al.*, 2006, 2008).

1.4.3 – The Effects of RAMPs on GPCR Behaviour

Thus far, up to 46 GPCRs are thought to interact with one or more RAMPs, exhibiting both long-lasting, stable complex formation as well as relatively transient interactions (Kotliar *et al.*, 2023). As discussed, the RAMP accessory proteins play the role of chaperone, promoting the trafficking of certain GPCRs to the cell surface; the initial example being the CLR receptor, which is poorly expressed at the cell surface, and has no known function, without the interaction of a RAMP (McLatchie *et al.*, 1998). Since 1998, RAMPs have been shown to promote the trafficking of several more GPCRs to the cell surface, including the family C

calcium-sensing receptor (CaSR), family B corticotropin-releasing hormone receptor 1 (CRH1R) and the family A G protein-coupled oestrogen receptor 1 (GPR30), to name a few (Bouschet and Henley, 2005; Bouschet, Martin and Henley, 2008; Lenhart *et al.*, 2013; Wootten *et al.*, 2013). Many GPCRs other than CLR can be expressed and may be functional without an interacting RAMP, leading to populations of both RAMP-free and RAMP-complexed GPCRs at the cell surface, as is the case with chemokine and glucagon (GCGR) receptors (Mackie *et al.*, 2019; McGlone *et al.*, 2021).

Furthermore, RAMPs also influence trafficking away from the cell surface, as with the internalisation and desensitisation of GPCRs following signalling. CLR is the most well-studied receptor in this regard; CLR and RAMP1, and CLR and RAMP2 (the CGRP and AM receptors, respectively) both internalise in a β-arrestin dependent manner, however, CLR and RAMP3 (the AM2 receptor) internalises due to RAMP3's PDZ motif interacting with the Na^+/H^+ exchanger regulatory factor (NHERF) or N-ethylmaleimide-sensitive factor (NSF) (Hilairet *et al.*, 2001; Bomberger, Parameswaran, *et al.*, 2005; Bomberger, Spielman, *et al.*, 2005). RAMP3's PDZ motif is also responsible for the rapid recycling of the AM2 and atypical chemokine receptor 3 (ACKR3) back to the cell surface, following internalisation (Mackie *et al.*, 2019). Interestingly, as GPCRs can continue to signal from internal membranes such as endosomes, the CGRP receptor demonstrated that this can also include RAMP-complexed receptors, with evidence suggesting that these complexes may be able to remain intact during the internalisation or recycling processes (Kuwasako *et al.*, 2000; Calebiro *et al.*, 2010; Yarwood *et al.*, 2017).

In addition to their chaperone and trafficking roles, the RAMPs can also influence the ligand selectivity and downstream signalling of several key GPCRs. Again, this is only well documented with regards to the calcitonin family of receptors thus far (McLatchie *et al.*, 1998; Husmann *et al.*, 2003; Morfis *et al.*, 2008). As discussed in section 1.3.2, the calcitonin receptor in complex with a RAMP becomes the receptor for amylin, switching the receptor's phenotype; the three amylin receptors (AMY1 to 3) bind amylin and calcitonin with high and low affinity, respectively, which is the opposite of when the calcitonin receptor is RAMP-free. Similarly, the CLR receptor in complex with RAMP1 becomes the CGRP receptor, but switches phenotype to bind adrenomedullin when complexed with RAMP2 or RAMP3. However, these receptor-RAMP complexes can also share ligand affinity between the phenotypes mentioned; the adrenomedullin 2 receptor can bind CGRP with moderate affinity, and the amylin 1 receptor binds CGRP with high affinity as well as amylin. This leads to a quite complex picture of ligand selectivity and affinity between these two receptors, three RAMPs and four peptide ligands, as discussed in section 1.3.2.

Aside from the calcitonin family, the extent of ligand selectivity via RAMPs is not as clear with other receptors. The effect of RAMP2 on the glucagon receptor is currently disputed, with evidence for both enhancing its activation, or not at all, and some receptors such as the vasoactive intestinal polypeptide receptor 1 (VIPR1) have never demonstrated RAMP-dependent ligand selectivity (Christopoulos *et al.*, 2003; Weston *et al.*, 2015; Cegla *et al.*, 2017; Shao *et al.*, 2022). The evidence thus far therefore suggests that the calcitonin family may be a special case in the context of ligand selectivity; a possible explanation for this may be due to an N-terminal α-helical motif maintained in the selective ligands prior to ligand binding, which is not found in the non-selective ligands until ligand binding has occurred (Liang, Belousoff, Zhao, *et al.*, 2020; Deganutti *et al.*, 2021). Therefore, this difference in the calcitonin family peptide secondary structure may be the reason why RAMP-dependent ligand selectivity has only been well documented with the calcitonin and CLR receptors, and not others.

Following ligand binding, RAMPs are also able to influence the downstream signalling profiles of GPCRs, often pleiotropically. The vasoactive intestinal polypeptide receptor 1 (VIPR1) is able to couple to multiple G proteins, stimulating both adenylate cyclase and phospholipase C activity. Phosphoinositide hydrolysis is enhanced by RAMP2 without any changes in cAMP accumulation, suggesting that RAMP2 could modulate the G protein coupling of the VIPR1 receptor (Christopoulos *et al.*, 2003). In addition, the calcitonin receptor in complex with RAMP1 or RAMP3 gives a 20 to 30 fold increase in amylin-stimulated cAMP production as opposed to receptor alone, as well as a 2 to 5 fold increase in intracellular Ca^{2+} (Morfis *et al.*, 2008). These data suggests that the AMY1 and AMY3 receptors preferentially couple to $G\alpha_s$ rather than $G\alpha_q$ compared to the calcitonin receptor alone, indicating that RAMPs do indeed influence the coupling of G proteins to GPCRs. Interestingly, RAMPs may also promote the uncoupling of G proteins as well. RAMP2 has been shown to enhance the glucagon-mediated activation of the glucagon receptor (GCGR), however, instead of increasing ligand affinity or the rate of trafficking to the cell surface, RAMP2 uncouples $G\alpha_i$ from GCGR, permitting $G\alpha_s$ signalling instead. This is not only RAMP-dependent but ligand-dependent as well; $G\alpha_i$ coupling to GCGR is unaffected by the glucagon-like peptide oxyntomodulin (Weston *et al.*, 2015). Finally, RAMPs may also enable the coupling of a wider range of G protein subunits to GPCRs; both the VIPR2 and CRF1R receptors show a RAMP-dependent enhancement of basal $G\alpha_{i/o/t/z}$ coupling when compared to receptor alone, promoting a unique signalling profile (Wootten *et al.*, 2013).

Overall, for such a small group of single-transmembrane accessory proteins, the RAMPs play quite a significant role in modulating the behaviour of select GPCRs through their trafficking patterns, glycosylation states and direct allostery, although many of these interactions have

not been fully elucidated as of yet. Ultimately, greater structural information of the RAMPs and their interactions with GPCRs would undoubtedly provide a platform to further understand and explain their varied biological roles.

1.4.4 – Structural and Dynamic Insights

Numerous structures of RAMPs have now been resolved, with and without a complexed GPCR, although many of these are only partial structures such as the RAMP extracellular domain or those which lack detail in certain areas such as the C-terminal tail. Despite these incomplete structures, some understanding of RAMP mechanisms of action has been garnered, and any resolved structures can also aid further analyses by computational methods. Overall, the RAMPs seem to only exert relatively subtle effects on the GPCR structure itself, making minimal contacts with ligands and mainly influencing the GPCR's conformation allosterically (Kotliar *et al.*, 2023).

The active human CLR and RAMP1, or CGRP receptor, structure was resolved to 3.3 Å by cryo-EM in complex with CGRP itself and $G\alpha_s$ in 2018, confirming findings based upon the previously resolved extracellular domains (Liang, Khoshouei, Deganutti, *et al.*, 2018). The TM helix of RAMP1 makes contact with TM helices 3, 4 and 5 of CLR, and the ECD of RAMP1 interacts with the ECD and ECL2 of CLR; unfortunately, the C-terminal tails of both were not resolved. As the contact between the RAMP1 ECD and CGRP is minimal, it was suggested that RAMP1 stabilises the ECD and ECL2 of CLR to promote ligand binding (Liang, Khoshouei, Deganutti, *et al.*, 2018). The full-length structures of CLR with RAMP2 and RAMP3 also displayed a similarly minimal contact with ligands, further strengthening the argument of RAMP allostery (Liang, Belousoff, Fletcher, *et al.*, 2020; Pioszak and Hay, 2020). Interestingly, the RAMPs also conveyed several other subtle conformational effects to CLR including the orientation of the ECD, the kink in CLR's TM6, the conformation of ICL2, and position of ECL3. Following further analysis of these data, it was suggested that the linker region between the RAMP ECD and TM actually confers its stabilising ability, and the dynamic C-terminal tail of RAMP3 makes transient contacts with the G protein itself (Kotliar *et al.*, 2023). In particular, the RAMP1 linker region interacts quite strongly with the CLR ECL2 when CGRP-bound, likely stabilising the active conformation (Josephs *et al.*, 2021).

As well as the CGRP and AM receptors, each of the three amylin receptors have also been resolved in complex with $G\alpha_s$ and several ligands, seemingly activated via distinct mechanisms which may involve RAMP allostery (Cao *et al.*, 2022). Despite similar overall interactions at TM helices 3, 4 and 5, and ECL2, the RAMPs again had little effect on the calcitonin receptor structure, only displaying subtle differences in the conformation of ICL2, and varying contact with $G\alpha_s$, although the linker region was found to be even more stabilising

with the calcitonin receptor than with CLR (Cao *et al.*, 2022). These results again highlight the subtle allosteric nature of the RAMPs, only forming minimal hydrophobic contact with the C-termini of these peptide ligands.

Overall, while the CLR receptor structure is now available, including those complexed with a RAMP, current data is lacking especially with regards to the unresolved C-terminal tails of both receptor and RAMP. In addition, structural information of interactions with β-arrestin and G proteins other than $G\alpha_s$ are also lacking. Furthermore, as the RAMPs are thought to interact with all 15 members of the family B GPCRs, it would be interesting to focus on these other combinations all well as the calcitonin family, especially as the structures of all 15 have now been resolved (Cong *et al.*, 2022). Similarly, any structural effects of RAMPs with family A and C GPCRs should also be of interest due to their differences with family B.

Several computational techniques, including molecular dynamics simulations, have been able to take advantage of these resolved RAMP structures in order to propose theories as to any potential RAMP effects on GPCR dynamics. The original cryo-EM structure of the CGRP receptor was used as a basis for several simulations (Liang, Khoshouei, Deganutti, *et al.*, 2018). RAMP1 was observed to stabilise the CLR ECD, which in turn stabilises the C-terminus of the CGRP peptide with solely transient interactions between CGRP and RAMP1; the same experiment without RAMP1 caused greater dynamics of the CLR ECD, and fewer conformational changes associated with signalling. In addition, the C-terminal tail of RAMP1, which was not resolved by cryo-EM, was modelled to also interact transiently with the CLR ICL2 and one of the α-helices of $G\alpha_s$. These simulation results not only correlate well with the experimentally resolved structural information, but also suggest that RAMP interactions do contribute to GPCR dynamics and stability; this is especially the case with the particularly flexible regions which were not resolved by cryo-EM (Kotliar *et al.*, 2023). In a similar manner, the interaction between RAMP1 and the C-terminus of amylin was also modelled with a full-length AMY1 receptor. These simulations did indeed show a similar stabilising effect from transient RAMP interactions, the ligand-binding residues of the calcitonin receptor were also RAMP1-dependent and showed a good overall level of agreement between the simulation and experimentally resolved AMY1 receptor (Bower *et al.*, 2018).

In addition to these experiments, molecular dynamics simulations have also shown that the N-terminal tail and ECD loops of the calcitonin receptor have greater flexibility when complexed with RAMP1, which is thought to allow for peptide binding due to this reduction in rigidity (Gingell *et al.*, 2016). These findings agree with the computational assessment of the two-step binding mechanism of CGRP, whereby loop 4 of the CLR ECD was shown to be of importance to peptide binding, which in turn may be another RAMP-dependent allosteric

mechanism via the state of loop 4 (Deganutti *et al.*, 2021). Furthermore, these studies also displayed minimal contact between RAMP1 and CGRP, in agreement with other computational experiments. Indeed, RAMPs have also been found to bias GPCRs towards certain G proteins via modelling and simulations, correlating with experimental results and further emphasising their importance to CGRP pharmacology and general receptor bias (Weston *et al.*, 2016). Modelling of mutants has also been employed to study the activation of the CGRP receptor complex; it was found that mutated residues at the extracellular face of the CLR TM bundle affected its signalling capability, which is made potent by TM helix packing, and the ECLs being of importance to ligand binding (Woolley *et al.*, 2017). Finally, the modelled CGRP receptor structure was compared to the experimentally resolved calcitonin receptor, and found that RAMP1 reorganised TMs 1, 6 and 7 of CLR, restricting the movement of TM6 and therefore ECL3, when compared to the RAMP-free calcitonin receptor (Woolley *et al.*, 2017).

Overall, modelling has therefore shown that the dynamics of CLR are indeed RAMP-dependent, with particular focus on ECL2, ECL3 and loop 4 of the ECD (Pham *et al.*, 2019; Kotliar *et al.*, 2023). While molecular dynamics have proven useful in probing GPCR-RAMP dynamics, its full potential has not yet been unlocked; homology modelling could detect further RAMP interactions, MD simulations could be used to design novel drugs based upon specific structural motifs, and eventually course-grain simulations will be able to give detail on the physical aspects of GPCR-RAMP complex formation, complementary to experimentally resolved structures.

1.4.5 – RAMPs as Potential Drug Targets

In summary, RAMPs have been found to interact with up to 46 GPCRs, modulating their behaviour through trafficking patterns, ligand selectivity, bias towards G protein coupling, and the dynamics of GPCR activation. As such, considering that GPCRs are widely implicated in human health and disease, RAMPs are also proposed to be associated with disease states and are therefore attractive and novel drug targets (Jacob, Wu and Wang, 2012).

CGRP and its receptor are known to be a key mechanism in the development of migraines, and thus several inhibitors have been developed to interfere with the CGRP-CLR-RAMP1 complex. Erenumab, eptinezumab, fremanzumab and galcanezumab are all therapeutic monoclonal antibodies which have been approved by the Food and Drug Administration (FDA), and are delivered either by subcutaneous injection or intravenous infusion (Reuter *et al.*, 2018; Tepper, 2018; Wattiez, Sowers and Russo, 2020). In addition to these, several small molecule inhibitors have also been approved including ubrogepant, rimegepant and atogepant (Edvinsson *et al.*, 2018; Scuteri *et al.*, 2022). Olcegepant and telcagepant are also CGRP

antagonists which bind to both CLR and RAMP1 directly, although the latter was discontinued due to its side effects, as was also the case with the orally delivered antagonist MK-3207 (Bell *et al.*, 2010; Hewitt *et al.*, 2011). However, several more therapeutics for the treatment of migraine are in development; the intranasal small molecule antagonist, zavegepant, is currently in clinical trials, while *in silico* drug repurposing has potentially identified two more CGRP receptor antagonists which are already FDA approved (Croop *et al.*, 2021; Aksoydan and Durdagi, 2022). CLR also forms receptors for adrenomedullin when complexed with RAMP2 or RAMP3, and has been implicated in the progression of tumour growth; adrenomedullin antagonists have been developed for their anti-cancer potential but remain in pre-clinical stages (Avgoustou *et al.*, 2020; Jailani *et al.*, 2022).

In terms of the calcitonin receptor, the amylin analogs pramlintide and cagrilintide have been effective in the treatment of diabetes mellitus and obesity, respectively, although the latter is not yet FDA approved (Ratner *et al.*, 2004; Ryan, Jobe and Martin, 2005). Cagrilintide is unique as it is a ‘dual amylin and calcitonin receptor agonist’ (DACRA), as is salmon calcitonin which was in clinical trials for the treatment of post-menopausal osteoporosis (Binkley *et al.*, 2012; Kruse *et al.*, 2021; Lau *et al.*, 2021). In addition, the DACRA KPB-088 was found to be promising in pre-clinical stage weight-loss trials, and the amylin peptide antagonist AC253 may be protective against the progression of Alzheimer’s disease (Soudy *et al.*, 2019; Larsen *et al.*, 2020). The calcitonin receptor therefore has several promising therapeutic options via both its agonism and antagonism.

While these are relatively clear examples of the involvement of RAMPs in disease states, with available therapeutic options, several more potential examples exist which have not yet been fully confirmed or addressed. The upregulation of CLR has been reported to hinder the treatment of acute myeloid leukaemia, a non-coding RNA within the RAMP2 gene may be involved in cancer-related angiogenesis, and AM2 is thought to promote the migration and invasion of endothelial cells, for example (Grandits and Wieser, 2021; Lai *et al.*, 2021; Nagar *et al.*, 2021). In addition, CLR is a potential biomarker for the prognosis of low-grade glioma, a malignant tumour of the central nervous system, and adrenomedullin has been a focal point in the treatment of COVID-19 due to its anti-inflammatory properties (Han *et al.*, 2021; Kita and Kitamura, 2022). More specifically in terms of RAMPs, select single nucleotide polymorphisms (SNPs) in RAMP2 correlate with the extent of headache post-concussion, and SNPs in RAMP3 may be involved in the age-related risk of fracture and body composition (Prakash *et al.*, 2019; La Fountaine *et al.*, 2022).

Overall, despite these associations, the exact role that RAMPs play in human or mammalian disease is yet to be fully determined. Given the success of therapeutically targeting the CLR-

RAMP1 with antibodies, robust drug screening assays are required to increase the success of additional therapeutics, including small molecule inhibitors, at clinical trial. Chapter 3 details the creation of a FRET-based interaction assay between the calcitonin family of receptors and the three RAMPs, as a platform for drug screening.

1.5 – Aims and Objectives

Currently, the druggability of GPCRs which interact with RAMPs remains poor, leaving many diseases without specific, effective treatment options. Similarly, drug screening methods which include both GPCRs and RAMPs also remain underutilised. As such, one of the two aims of this thesis is to develop a FRET-based interaction assay between the calcitonin family of GPCRs and the RAMPs, for its application to high-throughput drug screening. This will be achieved by *in silico* construct design, molecular cloning and subsequent expression in mammalian cells. Furthermore, each construct's functionality will be confirmed through their ability to induce cAMP production, before optimisation of the FRET assay itself. In addition to its potential involvement in drug screening, this interaction assay will also be used to gain further insights into the dynamics of RAMPs, and their ability to out-compete each other when interacting with a GPCR. The final objective of this project is to trial the use of *Drosophila melanogaster* to express the constructs for this FRET assay, as an alternative expression system to mammalian cells. This is a currently underutilised system, but has promising advantages over traditional approaches, as discussed.

As technology advances, the number of resolved GPCR structures increases each year (https://gpcrdb.org), and while structure determination is of great importance, structures alone do not translate to druggable GPCRs and treatable diseases. Further analyses and comparisons of experimentally resolved GPCR structures under varying conditions are needed to further build fundamental knowledge and understanding of them. As such, the second aim of this thesis is to adapt the mathematical techniques geometric morphometrics and principal component analysis for use with GPCR structures, and is introduced in Chapter 4. The main objective is to create and optimise a robust process compatible with resolved GPCR structures, and to establish its usefulness for structural analyses, via a series of case studies based on activation states and bound ligands. The subsequent objective of this project is its application to further understanding GPCR structures, including the potential impacts of thermostabilising mutations and the use of fusion proteins, an appraisal of *de novo* structures resolved by the AlphaFold project, and whether geometric morphometrics can help to elucidate the structural determinants of G protein coupling.

Chapter 2: General Methods

2.1 – Creation of Fluorescent Receptor and RAMP Constructs

2.1.1 – Materials

Unless specified, general reagents and consumables were obtained from ThermoFisher Scientific, U.K or Sigma Aldrich, U.K. Restriction enzymes, PCR reagents and molecular biology kits were from New England Biolabs, U.S. KOD polymerase was from Merck Millipore, Germany, and primers were obtained from Eurofins, Germany. The molecular weight DNA ladder used in all agarose gels was the Quick-Load® 1 kb DNA ladder from New England Biolabs, U.S. Agarose gels were visualised using a G:Box gel imager from Syngene, U.K.

2.1.2 – Complementary DNA Acquisition

All cDNA was kindly provided by the Prof. D. R. Poyner lab, including constructs for human calcitonin receptor-like receptor (CLR), calcitonin receptor (CTR) and receptor activity modifying proteins 1, 2 and 3 (RAMP1, RAMP2 and RAMP3), all of which were obtained within the mammalian expression vector pcDNA3.1$^-$ (Invitrogen, U.K). Both CLR and CTR receptors include an N-terminal T8 signal peptide and haemagglutinin (HA) epitope, as previously described (McLatchie *et al.*, 1998). Similarly, all three RAMPs have their putative signal peptide removed, and an N-terminal FLAG epitope inserted (Barwell, 2010). Note, these peptide insertions or deletions do not significantly affect their pharmacology (Wootten *et al.*, 2013).

2.1.3 – Construct Design

The provided receptor and RAMP cDNA sequences were used *in silico* to design the constructs needed to produce fluorescently tagged receptors and RAMPs. Each receptor would be cloned in-frame with yellow fluorescent protein (YFP) at the C-terminus, separated by a flexible serine/glycine linker sequence. Similarly, each RAMP would be cloned in-frame with cyan fluorescent protein (CFP) in the same manner. Note that the template sequences described were not further modified, only tagged with the fluorescent protein. Restriction sites were selected based on the absence of any internal restriction sites, and the ability to ligate into the desired vector's multiple cloning site. The final constructs created are shown below. The construction of CLR-YFP in both pcDNA3.1$^-$ and pUAST-attB was ultimately outsourced to DC Biosciences Ltd, Scotland.

gaattcATGGCCTTACCAGTGACCGCCTTGCTCCTGCCGCTAGCCTTGCTGCTCCACGCCGCCAGGCCGGATTAC
GCGTCTTACCCGTATGACGTCCCAGATTACGCATCGCTGGGAGGCCCTTCACTCGAGGGATCCGCAGAATTAGAA
GAGAGTCCTGAGGACTCAATTCAGTTGGGAGTTACTAGAAATAAAATCATGACAGCTCAATATGAATGTTACCAA
AAGATTATGCAAGACCCCATTCAACAAGCAGAAGGCGTTTACTGCAACAGAACCTGGGATGGATGGCTCTGCTGG
AACGATGTTGCAGCAGGAACTGAATCAATGCAGCTCTGCCCTGATTACTTTCAGGACTTTGATCCATCAGAAAAA
GTTACAAAGATCTGTGACCAAGATGGAAACTGGTTTAGACATCCAGCAAGCAACAGAACATGGACAAATTATACC
CAGTGTAATGTTAACACCCACGAGAAAGTGAAGACTGCACTAAATTTGTTTTACCTGACCATAATTGGACACGGA
TTGTCTATTGCATCACTGCTTATCTCGCTTGGCATATTCTTTTATTTCAAGAGCCTAAGTTGCCAAAGGATTACC
TTACACAAAAATCTGTTCTTCTCATTTGTTTGTAACTCTGTTGTAACAATCATTCACCTCACTGCAGTGGCCAAC
AACCAGGCCTTAGTAGCCACAAATCCTGTTAGTTGCAAAGTGTCCCAGTTCATTCATCTTTACCTGATGGGCTGT
AATTACTTTTGGATGCTCTGTGAAGGCATTTACCTACACACACTCATTGTGGTGGCCGTGTTTGCAGAGAAGCAA
CATTTAATGTGGTATTATTTTCTTGGCTGGGGATTTCCACTGATTCCTGCTTGTATACATGCCATTGCTAGAAGC
TTATATTACAATGACAATTGCTGGATCAGTTCTGATACCCATCTCCTCTACATTATCCATGGCCCAATTTGTGCT
GCTTTACTGGTGAATCTTTTTTTCTTGTTAAATATTGTACGCGTTCTCATCACCAAGTTAAAAGTTACACACCAA
GCGGAATCCAATCTGTACATGAAAGCTGTGAGAGCTACTCTTATCTTGGTGCCATTGCTTGGCATTGAATTTGTG
CTGATTCCATGGCGACCTGAAGGAAAGATTGCAGAGGAGGTATATGACTACATCATGCACATCCTTATGCACTTC
CAGGGTCTTTTGGTCTCTACCATTTTCTGCTTCTTTAATGGAGAGGTTCAAGCAATTCTGAGAAGAAACTGGAAT
CAATACAAAATCCAATTTGGAAACAGCTTTTCCAACTCAGAAGCTCTTCGTAGTGCGTCTTACACAGTGTCAACA
ATCAGTGATGGTCCAGGTTATAGTCATGACTGTCCTAGTGAACACTTAAATGGAAAAAGCATCCATGATATTGAA
AATGTTCTCTTAAAACCAGAAAATTTATATAATGGCGGCAGCGGCGGCAGCcccgggGTGAGCAAGGGCGAGGAG
CTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGC
GAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCC
TGGCCCACCCTCGTGACCACCTTCGGCTACGGCCTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCAC
GACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTAC
AAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAG
GAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAG
CAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCAC
TACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCGCC
CTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTC
GGCATGGACGAGCTGTACAAGTAAaagctt

Figure 2.1.3.1. cDNA sequence of human calcitonin receptor-like receptor tagged with yellow fluorescent protein (CLR-YFP). The start (ATG) and stop (TAA) codons are red. DNA encoding the cleavable signal peptide is blue, the HA epitope is orange and the flexible linker is purple. YFP is underlined. Restriction sites for *Eco*RI, *Xma*I, and *Hind*III are green. This construct was synthesised by DC Biosciences Ltd, Scotland.

ggtaccATGGCCTTACCAGTGACCGCCTTGCTCCTGCCGCTAGCCTTGCTGCTCCACGCCGCCAGGCCGGATTAC
GCGTCTTACCCGTATGACGTCCCAGATTACGCAGAATTTTCAAATCAAACCTATCCAACAATAGAGCCCAAGCCA
TTTCTTTACGTCGTAGGACGAAAGAAGATGATGGATGCACAGTACAAATGCTATGACCGAATGCAGCAGTTACCC
GCATACCAAGGAGAAGGTCCATATTGCAATCGCACCTGGGATGGATGGCTGTGCTGGGATGACACACCGGCTGGA
GTATTGTCCTATCAGTTCTGCCCAGATTATTTTCCGGATTTTGATCCATCAGAAAAGGTTACAAAATACTGTGAT
GAAAAAGGTGTTTGGTTTAAACATCCTGAAAACAATCGAACCTGGTCCAACTATACTATGTGCAATGCTTTCACT
CCTGAGAAACTGAAGAATGCATATGTTCTGTACTATTTGGCTATTGTGGGTCATTCTTTGTCAATTTTCACCCTA
GTGATTTCCCTGGGGATTTTCGTGTTTTTCAGAAAATTGACAACTATTTTTCCTTTGAATTGGAAATATAGGAAG
GCATTGAGCCTTGGCTGCCAAAGGGTAACCCTGCACAAGAACATGTTTCTTACTTACATTCTGAATTCTATGATT
ATCATCATCCACCTGGTTGAAGTAGTACCCAATGGAGAGCTCGTGCGAAGGGACCCGGTGAGCTGCAAGATTTTG
CATTTTTTCCACCAGTACATGATGGCCTGCAACTATTTCTGGATGCTCTGTGAAGGGATCTATCTTCATACACTC
ATTGTCGTGGCTGTGTTTACTGAGAAGCAACGCTTGCGGTGGTATTATCTCTTGGGCTGGGGGTTCCCGCTGGTG
CCAACCACTATCCATGCTATTACCAGGGCCGTGTACTTCAATGACAACTGCTGGCTGAGTGTGGAAACCCATTTG
CTTTACATAATCCATGGACCTGTCATGGCGGCACTTGTGGTCAATTTCTTCTTTTTGCTCAACATTGTCCGGGTG
CTTGTGACCAAAATGAGGGAAACCCATGAGGCGGAATCCCACATGTACCTGAAGGCTGTGAAGGCCACCATGATC
CTTGTGCCCCTGCTGGGAATCCAGTTTGTCGTCTTTCCCTGGAGACCTTCCAACAAGATGCTTGGGAAGATATAT
GATTACGTGATGCACTCTCTGATTCATTTCCAGGGCTTCTTTGTTGCGACCATCTACTGCTTCTGCAACAATGAG
GTCCAAACCACCGTGAAGCGCCAATGGGCCCAATTCAAAATTCAGTGGAACCAGCGTTGGGGGAGGCGCCCCTCC
AACCGCTCTGCTCGCGCTGCAGCCGCTGCTGCGGAGGCTGGCGACATCCCAATTTACATCTGCCATCAGGAGCCG
AGGAATGAACCAGCCAACAACCAAGGCGAGGAGAGTGCTGAGATCATCCCTTTGAATATCATAGAGCAAGAGTCA
TCTGCTGGCGGCAGCGGCGGCAGCggatccGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTG
GTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGC
AAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCTTCGGC
TACGGCCTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAA
GGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAG
GGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAG
CTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTC
AAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGAC
GGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGC
GATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAaag
ctt

Figure 2.1.3.2. cDNA sequence of human calcitonin receptor tagged with yellow fluorescent protein (CTR-YFP). The start (ATG) and stop (TAA) codons are red. DNA encoding the cleavable signal peptide is blue, the HA epitope is orange and the flexible linker is purple. YFP is underlined. Restriction sites for *Kpn*I, *Bam*HI, and *Hind*III are green. Sequencing data can be found in Appendix A1.

gaattcATGGACTACAAGGACGACGACGACAAGACTACTGCCTGCCAGGAGGCTAACTACGGTGCCCTCCTCCGG
GAGCTCTGCCTCACCCAGTTCCAGGTAGACATGGAGGCCGTCGGGGAGACGCTGTGGTGTGACTGGGGCAGGACC
ATCAGGAGCTACAGGGAGCTGGCCGACTGCACCTGGCACATGGCGGAGAAGCTGGGCTGCTTCTGGCCCAATGCA
GAGGTGGACAGGTTCTTCCTGGCAGTGCATGGCCGCTACTTCAGGAGCTGCCCCATCTCAGGCAGGGCCGTGCGG
GACCCGCCCGGCAGCATCCTCTACCCCTTCATCGTGGTCCCCATCACGGTGACCCTGCTGGTGACGGCACTGGTG
GTCTGGCAGAGCAAGCGCACTGAGGGCATTGTGGGCGGCAGCGGCGGCAGCggatccGTGAGCAAGGGCGAGGAG
CTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGC
GAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCC
TGGCCCACCCTCGTGACCACCCTGACCTGGGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAC
GACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTAC
AAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAG
GAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACATCAGCCACAACGTCTATATCACCGCCGACAAG
CAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCAC
TACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCC
CTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTC
GGCATGGACGAGCTGTACAAGTAAaagctt

Figure 2.1.3.3. cDNA sequence of human receptor activity modifying protein 1 tagged with cyan fluorescent protein (RAMP1-CFP). The start (ATG) and stop (TAA) codons are red. DNA encoding the FLAG epitope is orange and the flexible linker is purple. CFP is underlined. Restriction sites for *Eco*RI, *Bam*HI, and *Hind*III are green. Sequencing data can be found in Appendix A2.

gaattcATGGACTACAAGGACGACGACGACAAGACTCAGCCTCTTCCCACCACAGGCACACCAGGGTCAGAAGGG
GGGACGGTGAAGAACTATGAGACAGCTGTCCAATTTTGCTGGAATCATTATAAGGATCAAATGGATCCTATCGAA
AAGGATTGGTGCGACTGGGCCATGATTAGCAGGCCTTATAGCACCCTGCGAGATTGCCTGGAGCACTTTGCAGAG
TTGTTTGACCTGGGCTTCCCCAATCCCTTGGCAGAGAGGATCATCTTTGAGACTCACCAGATCCACTTTGCCAAC
TGCTCCCTGGTGCAGCCCACCTTCTCTGACCCCCCAGAGGATGTACTCCTGGCCATGATCATAGCCCCCATCTGC
CTCATCCCCTTCCTCATCACTCTTGTAGTATGGAGGAGTAAAGACAGTGAGGCCCAGGCCGGCGGCAGCGGCGGC
AGCggatccGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA
AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATC
TGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTGGGGCGTGCAGTGCTTCAGC
CGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACC
ATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGC
ATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACATCAGC
CACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAG
GACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGAC
AACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAG
TTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAaagctt

Figure 2.1.3.4. cDNA sequence of human receptor activity modifying protein 2 tagged with cyan fluorescent protein (RAMP2-CFP). The start (ATG) and stop (TAA) codons are red. DNA encoding the FLAG epitope is orange and the flexible linker is purple. CFP is underlined. Restriction sites for *Eco*RI, *Bam*HI, and *Hind*III are green. Sequencing data can be found in Appendix A3.

gaattcATGGACTACAAGGACGACGACGACAAGACTGGCGGCTGCAACGAGACAGGCATGTTGGAGAGGCTGCCC
CTGTGTGGGAAGGCTTTCGCAGACATGATGGGCAAGGTGGACGTCTGGAAGTGGTGCAACCTGTCCGAGTTCATC
GTGTACTATGAGAGTTTCACCAACTGCACCGAGATGGAGGCCAATGTCGTGGGCTGCTACTGGCCCAACCCCCTG
GCCCAGGGCTTCATCACCGGCATCCACAGGCAGTTCTTCTCCAACTGCACCGTGGACAGGGTCCACTTGGAGGAC
CCCCCAGACGAGGTTCTCATCCCGCTGATCGTTATACCCGTCGTTCTGACTGTCGCCATGGCTGGCCTGGTGGTG
TGGCGCAGCAAACGCACCGACACGCTGCTGGGCGGCAGCGGCGGCAGCggatccGTGAGCAAGGGCGAGGAGCTG
TTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAG
GGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGG
CCCACCCTCGTGACCACCCTGACCTGGGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGAC
TTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAG
ACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAG
GACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACATCAGCCACAACGTCTATATCACCGCCGACAAGCAG
AAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTAC
CAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTG
AGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGC
ATGGACGAGCTGTACAAGTAAaagctt

Figure 2.1.3.5. cDNA sequence of human receptor activity modifying protein 3 tagged with cyan fluorescent protein (RAMP3-CFP). The start (ATG) and stop (TAA) codons are red. DNA encoding the FLAG epitope is orange and the flexible linker is purple. CFP is underlined. Restriction sites for *Eco*RI, *Bam*HI, and *Hind*III are green. Sequencing data can be found in Appendix A4.

2.1.4 Primer Design

To achieve this cloning strategy, oligonucleotide PCR primers (Table 2.1.4) were designed to amplify each sequence and include restriction sites to allow complementary fragments to combine in the cloning vector pTZ19R and further sub-cloning into the expression vector pcDNA3.1^{-}. All primers were initially designed to be between 18 and 24 base pairs where possible, have a GC content of 50 to 65%, a melting point temperature (T_m) close to 58°C, and to finish the 3` ends on a G or C to aid with hydrogen bonding to the template. Antisense primers were reverse complemented using https://www.bioinformatics.org/sms/rev_comp.html.

Table 2.1.4. Sense and antisense oligonucleotide primers for mammalian expression. Restriction sites are green, start and stop codons are red. YFP has two sense primers, one for sub-cloning with CLR (Sense1), and the other for sub-cloning with CTR (Sense2).

Sequence	Direction	Oligonucleotide Primer	Restriction Site
CFP	Sense	ggggggatccGTGAGCAAGGGCGAGGAGCTGTTC	*Bam*HI
	Antisense	ccccaagcttTTACTTGTACAGCTCGTCCATGCCGAG	*Hin*dIII
CLR	Sense	ggggggaattcATGGCCTTACCAGTGACCGCC	*Eco*RI
	Antisense	ccccccccgggGCTGCCGCCGCTGCCGCCATTATATAAATTTTCTGG	*Xma*I
CTR	Sense	ggggggtaccATGGCCTTACCAGTGACCGCC	*Kpn*I
	Antisense	ccccggatccGCTGCCGCCGCTGCCGCCAGCAGATGACTCTTGCTC	*Bam*HI
RAMP1	Sense	ggggggaattcATGGACTACAAGGACGACGACGACAAGACTAC	*Eco*RI
	Antisense	ccccggatccGCTGCCGCCGCTGCCGCCCACAATGCCCTCAGTGCGCTTGCTC	*Bam*HI
RAMP2	Sense	ggggggaattcATGGACTACAAGGACGACGACGACAAGACTCA	*Eco*RI
	Antisense	ccccggatccGCTGCCGCCGCTGCCGCCCAGCAGCGTGTCGGTGCGTTTGC	*Bam*HI
RAMP3	Sense	ggggggaattcATGGACTACAAGGACGACGACGACAAGACTGG	*Eco*RI
	Antisense	ccccggatccGCTGCCGCCGCTGCCGCCCAGCAGCGTGTCGGTGCGTTTGC	*Bam*HI
YFP	Sense1	ggggcccgggGTGAGCAAGGGCGAGGAGCTGTTC	*Xma*I
	Sense2	ggggggatccGTGAGCAAGGGCGAGGAGCTGTTC	*Bam*HI
	Antisense	ccccaagcttTTACTTGTACAGCTCGTCCATGCCGAG	*Hin*dIII

2.1.5 Molecular Biology

Using the appropriate primers from Table 2.1.4, each cDNA sequence for CFP, CLR, CTR, RAMP1, RAMP2, RAMP3 and YFP was first amplified by polymerase chain reaction (optimised reaction conditions are shown in Table 2.1.5). The resulting PCR products were then purified with the Monarch PCR and DNA Cleanup Kit (NEB, U.S.) according to manufacturer's instructions. These PCR products and vector (pTZ19R) then underwent a double digestion with the appropriate restriction enzymes (Table 2.1.4), which were incubated at 37°C for 1 hour. Digested DNA fragments were then separated in a 1% agarose gel containing 0.5 µg/ml ethidium bromide by electrophoresis at 120 V, and extracted with the Monarch Gel Extraction Kit (NEB, U.S.) according to manufacturer's instructions. Target gene fragments were then combined with the complementary digested pTZ19R vector and ligated with T7 DNA ligase at 4°C for 20 minutes, followed by 20°C for either 2 hours, or overnight. Finally, 1 µl of the ligation reaction was transformed into chemically competent DH5α *E. coli* cells by resting on ice for 20 minutes, heat shock at 42°C for 30 seconds, a return to ice for 2 minutes, and recovery in 500 µl Luria-Bertani (LB) medium.

The transformed cells were allowed to recover for one hour at 37°C, 180 rpm in LB medium before plating onto 100 µg/ml ampicillin agar plates, incubated overnight at 37°C; several colonies were screened for the relevant DNA fragment by miniprep and restriction digest, and successful colonies were then confirmed by Sanger sequencing (Source Bioscience, U.K.). Using the same method, YFP was subsequently sub-cloned in-frame with the C terminal end of each receptor, and CFP was sub-cloned in-frame with the C terminal end of each RAMP, within pTZ19R (refer to Table 2.1.4 for the relevant enzymes). Ultimately, each complete sub-cloned construct was then digested with the appropriate restriction enzymes (Table 2.1.4) and ligated into empty pcDNA3.1^{-}, achieving the construct design of 2.1.3. All cloned constructs were then confirmed by diagnostic digest, further Sanger sequencing, and expanded to concentrations between 0.5 and 1 mg/ml with the Qiagen HiSpeed Plasmid Maxiprep Kit (Qiagen, U.K.) according to manufacturer's instructions.

Table 2.1.5. Optimised PCR reaction conditions. Template DNA, primer concentration and annealing/extension temperatures were optimised for each of the seven cDNA sequences used.

Component	**Volume**	**Final Concentration**
KOD Buffer (10x)	5 µl	1x
25 mM $MgSO_4$	3 µl	1.5 mM
2 mM dNTP Mix	5 µl	0.2 mM
10 µM Primers	1 µl each	0.2 µM
Template DNA	2 µl	1-20 ng
KOD Polymerase	0.5 µl	1.25 U
ddH_2O	32.5 µl	-
Step	**Temperature**	**Time**
Initial Denaturation	94°C	2 min
34 Cycles	94°C	30 sec
	45-65°C	1 min
	72°C	1 min/kb
Final Extension	72°C	7 min
Hold	4°C	∞

2.1.6 Cell Culture and Transfection

COS-7 mammalian cells (fibroblast-like kidney cells from the African green monkey) were cultured in Dulbecco's Modified Eagles Medium with L-glutamine, 10% (v/v) foetal bovine serum (FBS) and 5% (v/v) penicillin/streptomycin supplementation. Cells were grown in a humidified 95% air and 5% CO_2 incubator at 37°C. For transfection, cells were grown to approximately 80% confluency in 6, 24 or 96 well plates, and two transfection methods were used – by polyethylenimine (PEI) and Lipofectamine 2000 (Invitrogen, U.K.).

The PEI used was the branched polymer version from Sigma (408727), with an average molecular weight of approximately 25 kDa. This was mixed thoroughly with distilled water to a working concentration of 1mg/ml, before ensuring a pH of 7.4 with concentrated HCl and filter

sterilising at 0.22 µm before use. As shown in Table 2.1.6.1, DNA was mixed with serum free (SF) media, PEI added and vortexed immediately for 1 second, 15 times. Note, when co-transfecting with both a receptor and RAMP, 0.125 µg of each were used to the total of 0.25 µg DNA. This solution was incubated at room temperature for 10 minutes before supplementation with full media, mixing thoroughly. The growth media was then removed from the cell culture, and this transfection mixture added dropwise onto the cells. Media was topped up to the usual working volume after 2 hours, and these cells were then incubated at 37°C for 48 hours before use.

Table 2.1.6.1. Optimised PEI Transfection Mixtures. Note, several DNA:PEI ratios were tested, and 1:3 was broadly found to be best in terms of successful, efficient transfections.

	96-Well	24-Well	6-Well
Cell Density/Well	20k	50k	300k
SF Media (µl)	10	50	100
DNA (µg)	0.25	1	3
PEI (µg)	1	4	12
Full Media (µl)	10	200	500
Top-up Media (µl)	180	1000	2000

Lipofectamine 2000 transfections were performed according to manufacturer's protocols (ThermoFisher Scientific, U.K.). Briefly, as summarised in Table 2.1.6.2, DNA and Lipofectamine 2000 reagent were mixed to a 1:1 ratio in Opti-Mem medium, incubated for 5 minutes at room temperature and added to 80% confluent cells. These cells were also subsequently incubated for 48 hours at 37°C before use.

Table 2.1.6.2. Optimised Lipofectamine 2000 Transfection Mixtures. Transfections were performed according to manufacturer instructions.

	96-Well	24-Well	6-Well
Cell Density/Well	20k	50k	300k
Opti-Mem Medium (µl)	125	250	700
DNA (µg)	2.5	5	14
Lipofectamine 2000 (µl)	2.5	5	15

2.1.7 Assessing Construct Pharmacology by cAMP Assay

Before their use in further assays, each fluorescent construct's functionality was first confirmed by AlphaScreen cyclic AMP assay (Perkin Elmer, U.K.), to ensure the cloning strategy had not significantly affected the construct's ligand binding and signalling capabilities.

96-well plates were seeded with 20k COS-7 cells/well in complete DMEM to be co-transfected with a receptor and RAMP, as described in 2.1.6. 48 hours post transfection, growth media was removed from the cells and replaced with 90 µl of stimulation buffer (50 ml Hank's

Buffered Salt Solution (HBSS), 50 mg bovine serum albumin (BSA), 100 µl 500 mM 3-isobutyl-1-methylxanthine (IBMX), pH 7.4, prewarmed to 37°C). Cells were incubated in this stimulation buffer for 1 hour while ligands were prepared separately in a range between 1 pM and 100 nM. The internal control, forskolin, was prepared to 50 mM in 100% ethanol; forskolin increases intracellular cAMP levels by activating adenylate cyclase (Awad *et al.*, 1983). 10 µl of the appropriate ligand was added to each well and incubated at 37°C for 30 minutes. All media was removed, and cells washed once with 1x phosphate buffered saline (PBS) which was removed immediately. To stop cellular activity, 100 µl of ice-cold 100% ethanol was added to each well and allowed to evaporate overnight at room temperature.

The following day, the dried cell film was dissolved in 75 µl lysis buffer (38 ml dH_2O, 1.2 ml 10% Tween20, 0.2 ml 1 M HEPES, 40 mg BSA, pH 7.4), incubating at room temperature for 10 minutes on an orbital shaker. Next, the acceptor bead mixture (5 µl acceptor beads, 495 µl stimulation buffer) and donor bead mixture (5 µl donor beads, 1494.6 µl stimulation buffer, 0.375 µl biotinylated cAMP, incubated for 30 minutes before use) were made in reduced lighting, followed by the preparation of a cAMP standard curve between 10 pM and 1 µM. 10 µl of each standard and of each cell lysate were transferred into a 384-well white, opaque Optiplate, briefly centrifuged to draw contents to the bottom of the wells. In reduced lighting, 5 µl acceptor bead mix and 15 µl preincubated donor bead mix were added to each well, the plate was sealed, wrapped in foil, briefly centrifuged again, and incubated overnight at room temperature in reduced lighting. Finally, luminescence was measured the following day with a Mithras LB 940 Multimode Microplate Reader: excitation at 680 nm, emission at 520 to 620 nm.

2.1.8 cAMP Data Analysis

GraphPad Prism 9 (GraphPad Software Inc., San Diego, CA) was used to generate sigmoidal non-linear regression dose-response curves based on the raw cAMP data produced in 2.1.7. A standard curve was generated using the AlphaScreen competitive immunoassay (an example is shown in Figure 2.1.8) and used to interpolate raw cAMP data from the cell lysate samples, and determine their agonist-stimulated cAMP levels in pmol/mg. This type of inhibitory dose-response curve is based on a model with a four-parameter variable slope, including the top and bottom of the curve, logIC50 and HillSlope, according to Equation 2.1.8.

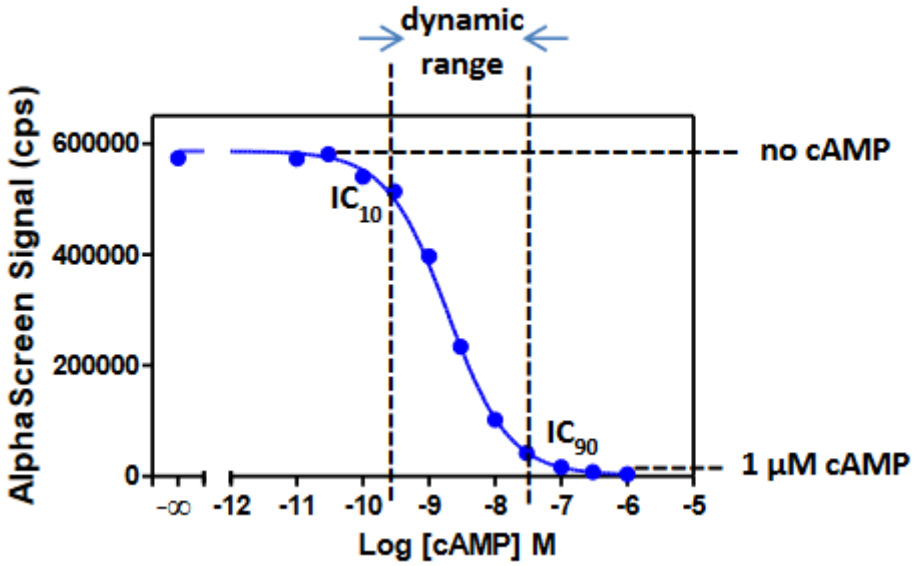

Figure 2.1.8. AlphaScreen cAMP Standard Curve Example. It is suggested that the most accurate and reliable range of the assay lies on its linear portion, between IC_{10} and IC_{90}. Taken from the AlphaScreen cAMP user manual (Perkin Elmer, U.K.).

$$Y = Bottom + \frac{(Top - Bottom)}{(1 + \ 10\text{^}((LogIC50 - X) * HillSlope))}$$

Equation 2.1.8. The Variable Slope Equation for Logarithms of Agonist Concentrations. Y is the response which follows a sigmoidal curve, and X is the logarithm of dose concentration. Bottom and top refer to the lowest and highest plateaus of the Y axis (response). For best fit analysis, HillSlope is assumed to be shared, meaning, dose treatment is assumed to shift IC50 but not basal response, maximal response, or the HillSlope variable itself.

2.2 Design and Development of FRET-Based Interaction Assay

2.2.1 Materials

Unless specified, general reagents and consumables were obtained from ThermoFisher Scientific, U.K. or Sigma Aldrich, U.K. FRET assays were performed in a black opaque 96-well format, using a Mithras LB 940 Multimode Multiplate Reader. Excitation filters 430/10 nm and 485/14 nm were used alongside emission filters 480/20 nm and 530/25 nm (Berthold Technologies, Germany).

2.2.2 Assay Procedure

The FRET assay method developed was based upon the gold-standard method created by Berthold Technologies themselves (Haase, Kamann and Brüx, 2011). 1 x 10^6 COS-7 cells per well were seeded into 6-well culture plates, and co-transfected with both a receptor-YFP and RAMP-CFP construct in a 1:1 ratio (to a total of 3 µg DNA per well). 48 hours after transfection, growth media was removed, cells washed in prewarmed PBS, and subsequently cells were detached into 300 µl PBS. 100 µl of cell suspension was transferred into a black opaque 96-well plate, in triplicate, and analysed in the LB 940 Mithras instrument. The fluorescence emission outputs for each of the donor, acceptor and FRET channels were measured sequentially for each well, using the round fluorescence scanning method. Data were collected in the MikroWin2000 software using the average of 10x10 scanned points in each well due to potential cell clustering. CFP fluorescence was measured at 480/20 nm after excitation at 430/10 nm; this is the donor channel, D. YFP fluorescence was measured at 530/25 nm after excitation at 485/14 nm; this is the acceptor channel, A. Finally, FRET signals were measured at 530/25 nm after excitation at 430/10 nm; this is the FRET channel, F. Note that any background or signal bleed through were measured for correction using non-transfected cells, and those transfected with solely donor or acceptor.

2.2.3 FRET Data Analysis

There are two main methods of processing FRET signal data; the first compares the ratio between donor-acceptor (I_{DA}) and donor (I_D) fluorescence intensity as an expression of FRET efficiency (E) as shown in Equation 2.2.3.1 (Sun *et al.*, 2011).

$$E = 1 - (\frac{I_{DA}}{I_D})$$

Equation 2.2.3.1. FRET Efficiency (*E*) Expressed As Donor Dequenching in the Presence of Acceptor. I_{DA} = FRET channel emission, cells expressing donor and acceptor. I_D = FRET channel emission, cells expressing donor only.

The second method described by Berthold Technologies (Haase, Kamann and Brüx, 2011) involves several stages of calculation, but benefits from taking background and signal bleed through into consideration (Equations 2.2.3.2 – 2.2.3.5). Firstly, signal bleed through for both donor and acceptor were first determined by assessing fluorescence intensity ratios (*d* and *a*) shown in Equations 2.2.3.2 and 2.2.3.3. These two factors *d* and *a* were then used to calculate the FRET signal corrected for bleed through (F_C) shown in Equation 2.2.3.4. F_c was then normalised for background correction, and to take differences in total expression levels into

account. This final process in Equation 2.2.3.5 gives the FRET efficiency, *E*, which also considers the quantum yields of each fluorophore.

$$d = F_D / D_D$$

Equation 2.2.3.2. Fluorescence Intensity Ratio for Donor Bleed Through (*d*). F_D = FRET channel emission, cells expressing donor only. D_D = Donor channel emission, cells expressing donor only.

$$a = F_A / A_A$$

Equation 2.2.3.3. Fluorescence Intensity Ratio for Acceptor Bleed Through (*a*). F_A = FRET channel emission, cells expressing acceptor only. A_A = Acceptor channel emission, cells expressing acceptor only.

$$F_C = F_{DA} - (D_{DA} * d) - (A_{DA} * a)$$

Equation 2.2.3.4. Bleed Through Corrected FRET Signal (F_c). F_{DA} = FRET channel emission, cells expressing both donor and acceptor. D_{DA} = Donor channel emission, cells expressing both donor and acceptor. A_{DA} = Acceptor channel emission, cells expressing both donor and acceptor.

$$E = 1 - [D_{DA} - (D_{DA} + F_c * \frac{Q_d}{Q_a})]$$

Equation 2.2.3.5. Normalised FRET Efficiency (*E*). D_{DA} = background corrected fluorescence intensity through the donor channel. Q_d and Q_a refer to the quantum yields of the donor and acceptor, respectively (Q_d = 0.4, Q_a = 0.61).

Ultimately, the two methods described both result in the FRET efficiency, *E*, which depends on the distance between donor and acceptor (*r*) due to the inverse 6th power law shown in Equation 2.2.3.6. Rearranging this enables the use of the experimentally derived FRET efficiency (*E*) and known Förster distance (R_0) of the fluorophore pair to find the distance (*r*) between them in nanometers (Sun *et al.*, 2011). The Förster distance is defined as the distance between donor and acceptor dipoles at 50% FRET efficiency, which is 4.751 nm for CFP and YFP (Patterson, Piston and Barisas, 2000; Bajar *et al.*, 2016).

$$E = \frac{1}{1 + (\frac{r}{R_0})^6}$$

$$r = R_0[\left(\frac{1}{E}\right) - 1]^{(\frac{1}{6})}$$

Equation 2.2.3.6. Donor-Acceptor Separation Distance (*r*). Rearranged equation to find *r*. *E* is FRET efficiency. *r* is distance between donor and acceptor. R_0 is the Förster distance of the fluorophore pair.

Both methods of determining FRET efficiency were performed and compared, with preference given to the more involved method as previously described (Periasamy *et al.*, 2008; Haase, Kamann and Brüx, 2011; Youvan *et al.*, 1997). All data was first processed in Microsoft Excel before transfer into GraphPad Prism 9 (GraphPad Software Inc., San Diego, CA).

2.3 Generating a Transgenic *Drosophila* Expression System

2.3.1 Materials

Unless specified, general reagents and consumables were obtained from ThermoFisher Scientific, U.K. or Sigma Aldrich, U.K. Enzymes were obtained from New England Biolabs, U.S. Primers were from Eurofins, Germany. Host *Drosophila* were obtained and cultured in the Dr C. Slack lab at 25°C, 60% humidity with a standard food pellet of 1.5% agar, 10% brewer's yeast and 5% sucrose. For increased scaling, continuous cultures were maintained and tipped to greater capacity. Flies were anesthetised with a CO_2 injection needle and observed under a 10x objective stereomicroscope. The molecular weight DNA ladder used in all agarose gels was the Quick-Load® 1 kb DNA ladder from New England Biolabs, U.S. Agarose gels were visualised using a G:Box gel imager from Syngene, U.K.

2.3.2 Cloning Strategy

In order to express a fluorescently tagged CGRP receptor in *Drosophila melanogaster* photoreceptor cells (PRC), the CLR-YFP and RAMP1-CFP constructs produced in section 2.1 were individually sub-cloned into the *Drosophila* expression vector pUAST-attB. This required amplification by PCR (following protocols outlined in 2.1), changing restriction sites for insertion into the pUAST-attB multiple cloning site; the *Hind*III site in CLR-YFP was changed to *Kpn*I, and *HindIII* in RAMP1-CFP was changed to *Eag*I. The primers used are listed in Table 2.3.2. Successful clones were again confirmed by diagnostic digest, sequencing and expanded to concentrations between 0.5 and 1 mg/ml before use.

Table 2.3.2 Sense and antisense oligonucleotide primers for *Drosophila* expression. Restriction sites are green, start and stop codons are red.

Sequence	Direction	Oligonucleotide Primer	Restriction Site
CLR-YFP	Sense	ggggGAATTCATGGCCTTACCAGTGACCGCC	*Eco*RI
	Antisense	ccccGGTACCTTACTTGTACAGCTCGTCCATGCCGAG	*Kpn*I
RAMP1-CFP	Sense	ggggGAATTCATGGACTACAAGGACGACGACGACAAGACTAC	*Eco*RI
	Antisense	ccccCGGCCGTTACTTGTACAGCTCGTCCATGCCGAG	*Eag*I

2.3.3 Generating Transgenic *Drosophila*

The two constructs cloned into the pUAST-attB vector were used for P-element mediated transformation of *Drosophila* embryos, where the plasmid is permanently inserted into the *Drosophila* genome by ΦC31 integrase. 500 ng/µl DNA samples were sent to the Fly Facility, Department of Genetics, University of Cambridge for microinjection and integration. RAMP1-CFP was inserted into the stock 13-20: vas-int; attp40 line of embryos (chromosome 2, 25C6) while CLR-YFP was inserted into the stock 13-18: nos-int; attP2 line of embryos (chromosome 3, 68A4). These embryos were received, cultured, and first crossed with white-eyed female virgins of a host line ($w^{-/-}$), to transfer the transgene from the germline to the somatic cell line. This reintroduces the red eye phenotype to easily track movement of the transgene through subsequent crosses. The resulting adult males with red eyes were again crossed with white-eyed female virgins a second time, to select for successful crosses and to boost stocks.

Next, adult males containing the transgene were crossed with the Curly derivative of Oster balancer line ($Cy^{+/-}$) which introduces an irradiated X chromosome to create a stable transgenic stock, meaning transgenic *Drosophila* would now exhibit the red eye, curly wing phenotype. Finally, these adult males were crossed with the GMR-GAL4 driver line for PRC-specific expression using the glass-binding enhancer element (GMR). At this final stage, transgenic *Drosophila* expressing the target construct will exhibit the red eye, straight wing phenotype, and are ready for use in experiments.

2.3.4 Confirming Transgenesis by Fluorescence

To determine whether the transgene was indeed incorporated successfully, PRC expression of the fluorescent target constructs was assessed by a rapid fluorescence assay, prior to more involved techniques such as SDS-PAGE and Western blotting. 20 *Drosophila* heads were manually dissected under a 10x objective and homogenised into 50 µl lysis buffer consisting of 25 mM Tris-Cl, 1% SDS. Samples were heated at 60°C for five minutes, centrifuged at 300 rcf and the supernatant harvested. Fluorescence of the supernatant was assessed against buffer only in a Mithras LB 940 Multimode Microplate Reader; excitation at 405 nm, emission at 485 nm. In addition, a single head of the RAMP1-CFP/GMR-GAL4 strain was visualised with a Leica TCS SP8 confocal microscope at 10x, 40x and 60x magnification, set to detect CFP fluorescence. A wide-field baseline was performed first, with checks for background auto-fluorescence. Note, only *Drosophila* expressing the RAMP1-CFP construct could be used to gather preliminary data.

2.4 The Application of Geometric Morphometrics to GPCR Structure

2.4.1 Software and Websites Used

Note that all of these are freely available online.

AlphaFold Protein Structure Database - https://alphafold.com/

GPCR Database (*GPCRdb*) - https://gpcrdb.org/structure/

Membrane Proteins of Known Structure (*mpstruc*) - https://blanco.biomol.uci.edu/mpstruc/

MorphoJ v1.07 - https://morphometrics.uk/MorphoJ_page.html

PAST Data Analyser v4.03 - https://past.en.lo4d.com/windows

Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB) - https://rcsb.org/

Swiss-PdbViewer (DeepView) v4.10 - http://expasy.org/spdbv/

2.4.2 Identification and Retrieval of Structural Models

In order to identify the GPCR structures available for inclusion into the sample selection, the *GPCRdb* (https://gpcrdb.org/structure/ (accessed 22/10/2021) (Kooistra *et al.*, 2021)) was first used to obtain a database of all currently known GPCR structures, which were subsequently cross-referenced to the *mpstruc* database (https://blanco.biomol.uci.edu/mpstruc/ (accessed 22/10/2021) (Shimizu *et al.*, 2018)) to ensure a comprehensive list was obtained. These databases were first downloaded, along with a plethora of information related to each structure's characteristics including resolution, activation state, bound ligands, mutations, etc. Next, the .pdb file of each selected structure was downloaded from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (https://rcsb.org/ (accessed 22/10/2021)), by use of their corresponding PDB codes. At this point, no GPCR structures were filtered or excluded from the sample selection in order to begin with a large, varied and unbiased selection; this of course included resolved structures representing as many of the sub-families within the A-F classifications of GPCRs as possible, if available at the point of sample selection. Additionally, selected GPCR structures predicted by the AlphaFold project (Jumper *et al.*, 2021) were also downloaded as .pdb files.

2.4.3 Visualisation and Manipulation of GPCR Models

Initially, the Swiss-PdbViewer software v4.10 (Guex and Peitsch, 1997) was used to visualise and manipulate each structure prior to data collection. Firstly, the entire receptor chain was

displayed as a ribbon format, and each transmembrane helix was assigned a colour for ease of identification. Note that any model manipulation was merely superficial and would not introduce any differences or bias into the data collection process. Next, the amino acid residues located at the very end of each transmembrane helix were identified manually, and the name and sequence number recorded. As GPCRs have seven transmembrane helices, this results in fourteen selected residues for each structure: seven at the extracellular face, and seven at the intracellular face.

2.4.4 Collection of Landmark Coordinate Data

The geometric morphometric analyses were performed on both the extra- and intracellular ends of GPCR transmembrane domains, to encompass the overall morphology of structures resolved under various conditions. Figure 2.4.4 shows a simplified schematic of a GPCR, and how the selection of residues marked by the fourteen stars, encompasses the morphology of both ends of the transmembrane bundle. Each .pdb file was opened as text in Microsoft Excel and the X, Y and Z coordinates located and recorded for each of the fourteen residues, as identified for each structure in 2.4.3. Specifically, these were the coordinates for the alpha-carbon (Cα) for each residue, which were then divided into two groups (extra- and intracellular) and finally exported as .txt files. This manual process of coordinate collection was repeated exactly for every structure analysed due to the differences between and quality of the .pdb file annotation; future semi-automation of this data collection could feasibly be achieved with the use of Python script programming. Finally, the XYZ coordinate data was exported as text files, compatible with their import into MorphoJ.

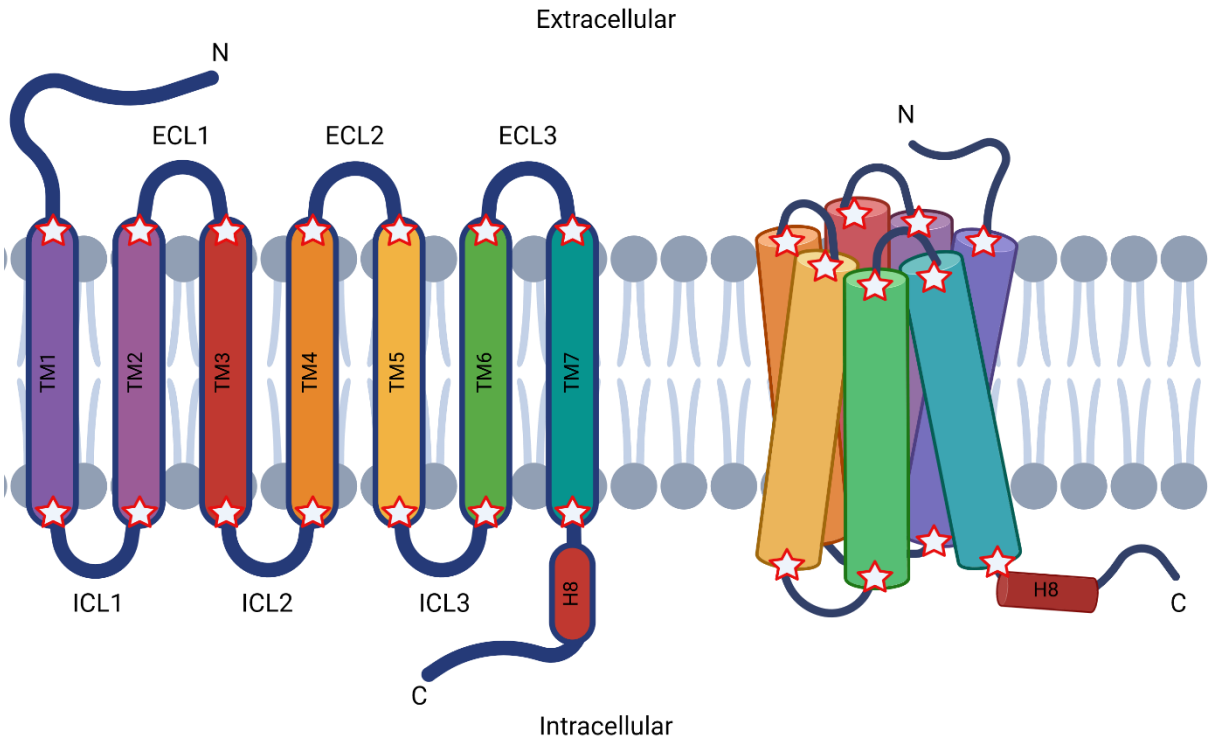

Figure 2.4.4. Simplified Landmark Selection Schematic. Both ends of each of the seven transmembrane helices were selected as discrete landmarks for morphometric analyses, depicted by the fourteen stars. This selection enabled an approximate encompassing of receptor morphology at both the extra- and intracellular face of the transmembrane bundle. Created with BioRender.com.

2.4.5 Geometric Morphometric Analysis

After importing the XYZ coordinate data collected from the defined landmarks into MorphoJ (Klingenberg, 2011), all analyses first began with a generalised procrustean superimposition; note that the extra- and intracellular data were processed separately each time. The figures shown in section 2.4.5 are representative examples of each step in the geometric morphometric analytical process, using coordinate data from the intracellular face of the beta-2 adrenergic receptor sub-family structures. As further discussed in section 4.2, this is an orthogonal transformation in which the shapes formed by the coordinate data are aligned and rescaled into standardised Procrustese shape coordinates. By doing so, the sum of squares between landmarks is minimised, and the cumulative effects of size and orientation are eliminated to ensure that the transformed shape coordinates are purely reflective of shape morphology. For clarity, the ‘shapes formed by coordinate data’ represent the positions of the approximate ends of each transmembrane helix, for either the extra- or intracellular face, and therefore capture the morphology of the resolved structures in the dataset. Figure 2.4.5.1 shows an example comparison of the three axes of the Procrustese shape coordinates as a

result of this orthogonal transformation, which were subsequently used to produce the covariance matrix needed for principal component analysis.

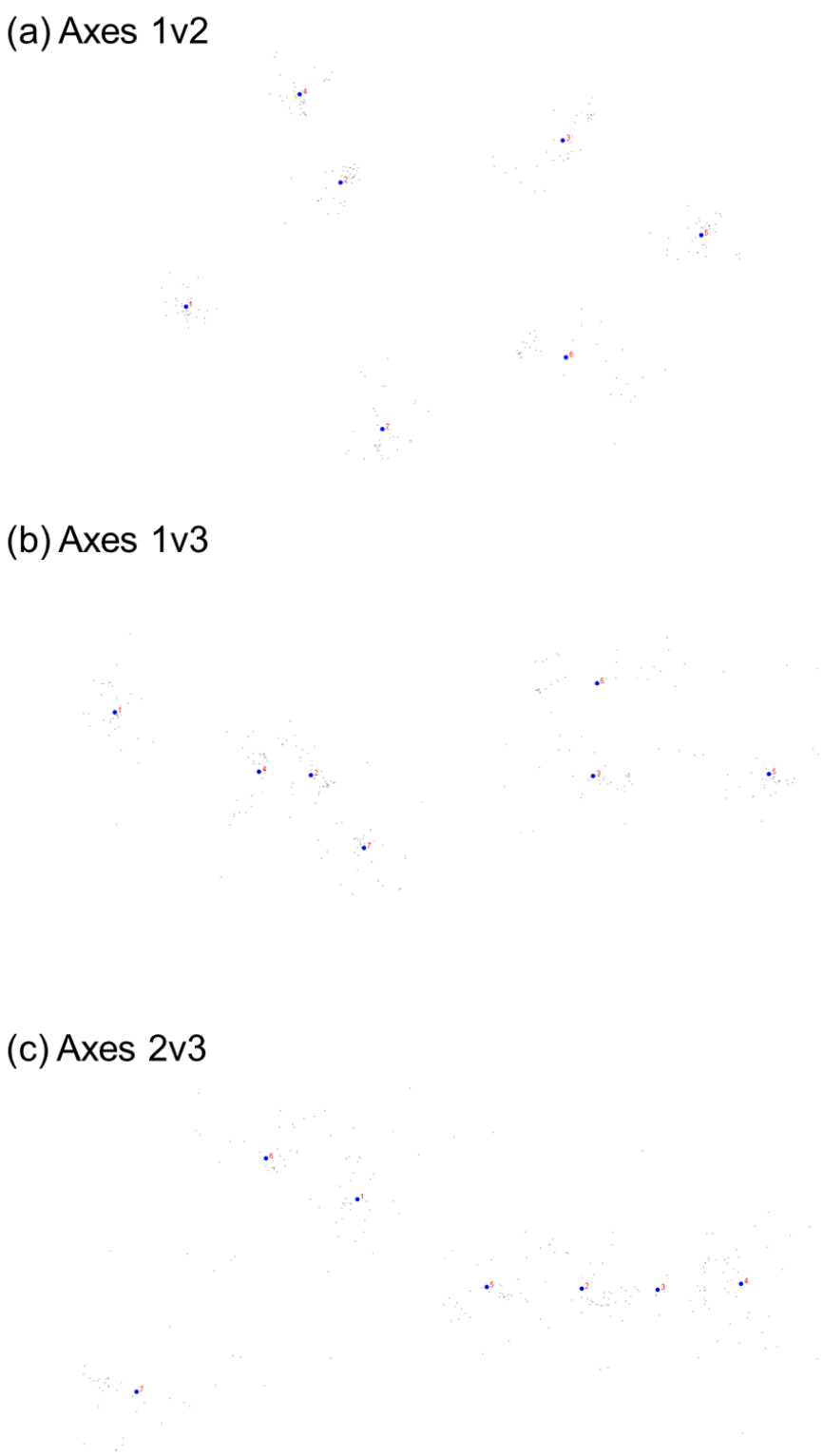

Figure 2.4.5.1. An Example of Procrustese Shape Coordinates. The XYZ coordinate data for the intracellular face of the beta-2 adrenergic receptor structures after procrustean superimposition. This is an orthogonal transformation which aligns and rescales raw coordinates, to eliminate the cumulative effects of size and orientation, ensuring the transformed Procrustese shape coordinates are reflective of pure shape morphology. Procrustean superimposition is required to generate covariance matrices which are in turn needed for principal component analysis. (a) is X vs Y, (b) is X vs Z, and (c) is Y vs Z.

Similarly to procrustean superimposition, principal component analysis is also an orthogonal transformation multivariate technique which seeks to determine the shortest distances between points. The use of a covariance matrix therefore enables this analysis to be two-dimensional, and thus landmark coordinates can be compared relative to one another. In addition, the covariance matrix generated from the Procrustese shape coordinates also provides the basis for eigenvector and eigenvalue calculations within the principal component analysis itself. As discussed, an eigenvector is the direction of the line representing the greatest variation in the shape coordinates, and its eigenvalue therefore, is the variance of the data in that same direction. As principal component analysis reduces the dimensionality of the coordinates without losing the original information, the principal components themselves are therefore linear combinations of the original variables, and capture the greatest variation within the data. As each principal component is orthogonal to each other principal component, they each capture a different aspect of the coordinate data; they are also ranked in order of greatest variation, beginning with PC1, then PC2 and so on. Figure 2.4.5.2 shows example principal components of the intracellular beta-2 adrenergic receptor dataset, with PC1 accounting for 42.071% of the variance and PC2 accounting for 14.602%, for example.

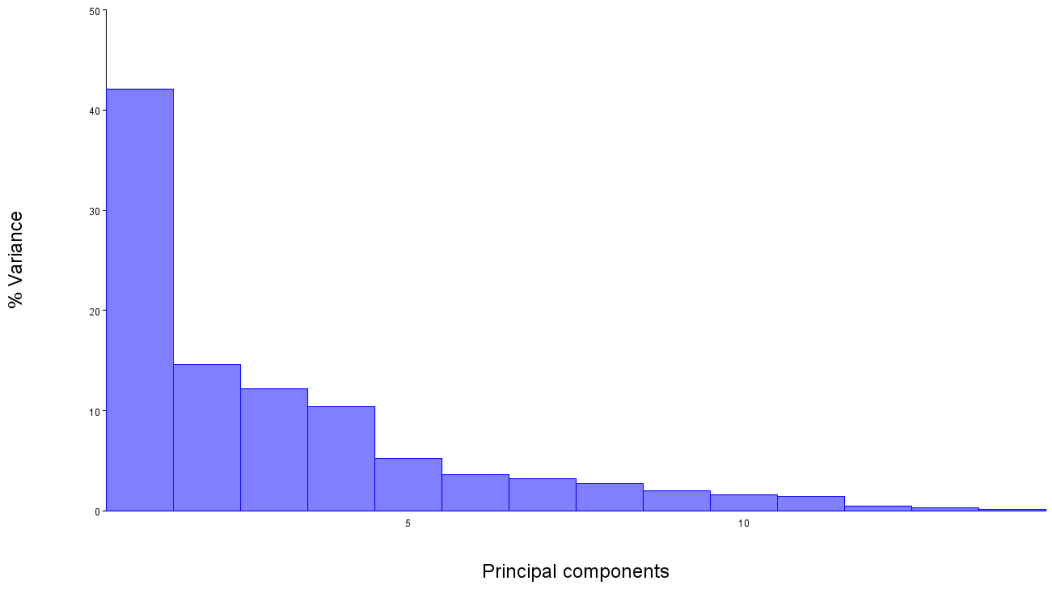

Figure 2.4.5.2. An Example of Variance Explained by Principal Components. The covariance matrix for the intracellular beta-2 adrenergic receptor structures was used for principal component analysis, producing principal components based on eigenvector and eigenvalue calculations. PC1 comprises the greatest variation in the data, with PC2 the next greatest, etc. PC1=42.071%, PC2=14.602%, PC3=12.142%, PC4=10.423%, PC5=5.202%, PC6=3.611%, PC7=3.211%, PC8=2.688%, PC9=2.039%, PC10=1.609%, PC11=1.432%, PC12=0.473%, PC13=0.333%, and PC14=0.164%.

As PC1 and PC2 represent the two greatest aspects of variation in the dataset, they are often focussed upon, at least initially. Generally, if the cumulative variance of PC1 and PC2 is greater than 50%, their comparison is often considered sufficient; if less than 50%, PC3 might also then be explored. This can be quite a restrictive approach to the interpretation of principal component data, and it may well be that subtle but interesting causes of variation are represented by relatively lowly ranked components. In any case, given that PC1 and PC2 do account for the greatest variance in the dataset, their initial interpretation was facilitated by Scree plots; Figure 2.4.5.3 shows example Scree plots for the intracellular beta-2 adrenergic receptor structures PC1 (a) and PC2 (b). Scree plots show a central point for each landmark called a centroid, and is an average position of all structures in the dataset. The line emerging from the centroid represents both the direction and proportional magnitude of the change for that particular principal component. Crucially, once these two principal components are compared directly, as shown in Figure 2.4.5.4, the further along the principal component axis a particular structure is, the further down the Scree plot line it is, moving away from the centroid. For example, in Figure 2.4.5.4, 3SN6 is the most positively shifted structure for PC1, and 7BZ2 is most positively shifted for PC2.

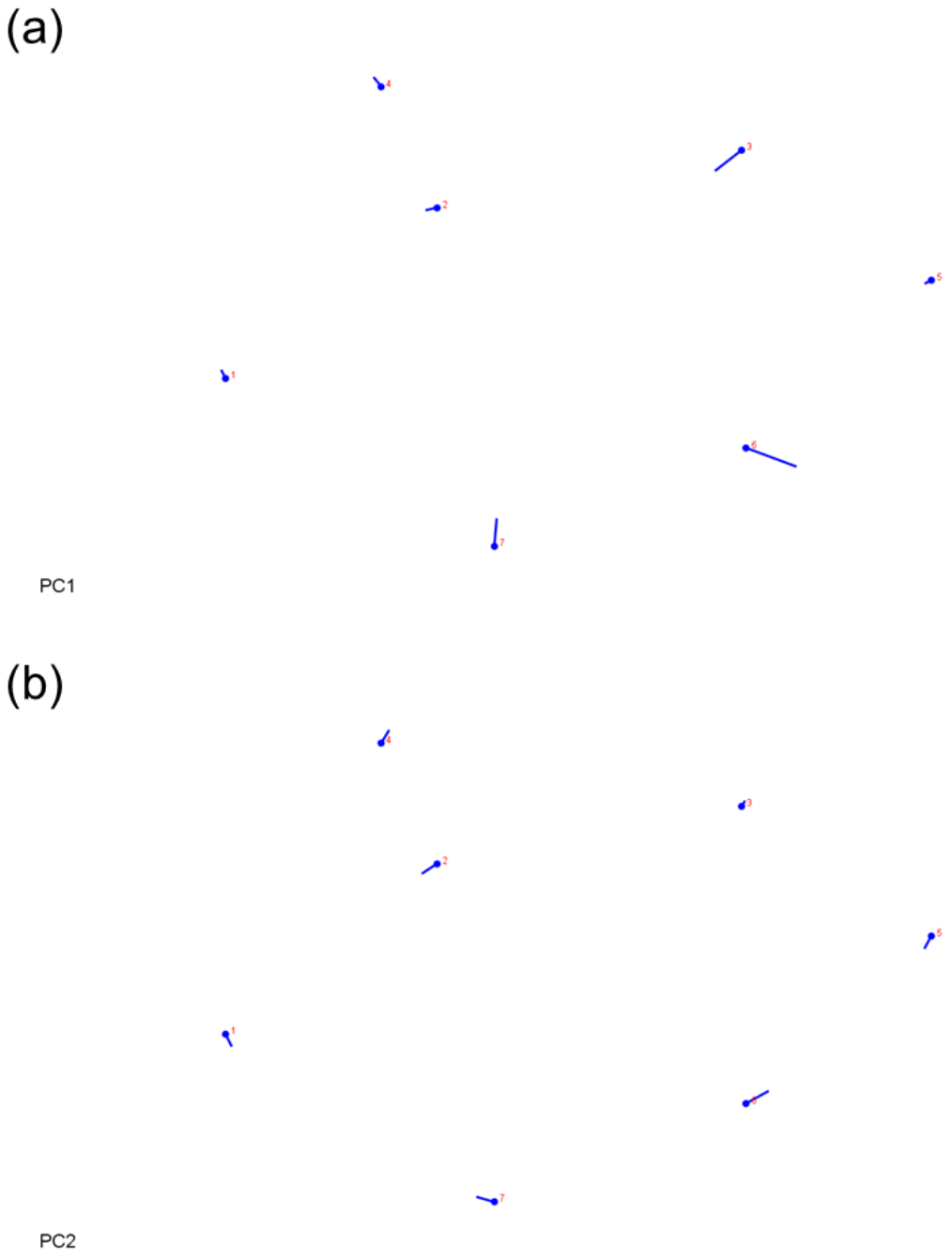

Figure 2.4.5.3. An Example of Principal Component Scree Plots. Each point represents the average centroid of each landmark, with the 'stick' moving in the direction of the principal component with magnitude.

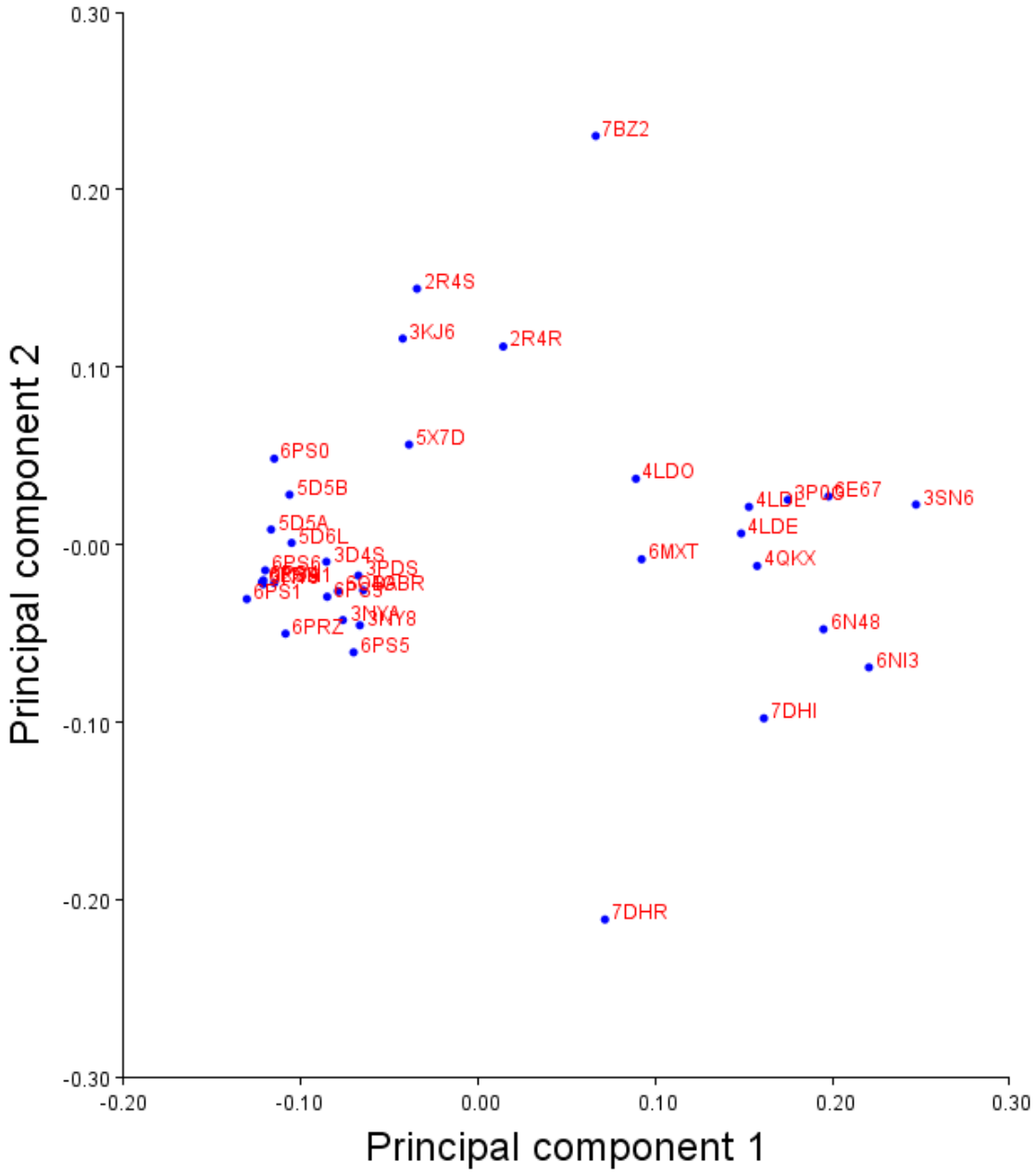

Figure 2.4.5.4. An Example of PC Score Comparison. The morphospace generated to compare principal component 1 and 2. Each receptor is represented by a blue circle, and are labelled by their PDB codes.

It is then from this final comparison of principle components that each structure can be labelled with a particular characteristic such as active, inactive, thermostabilised, etc., and patterns, groupings or outliers may be observed within the principal component data.

2.4.6 Statistical Analysis

In order to observe these potential groupings, and test their statistical significance, the principal component scores from the principal component analysis were exported as text files, and subsequently imported into PAST v4.03 data analyser (Hammer, Harper and Ryan, 2001). The PC scores were then subjected to an analysis of similarity, or ANOSIM, and a one-way permutational multivariate analysis of variance, or PERMANOVA, both with 9999 permutations and pairwise comparisons using Euclidean distances. One-way permutational multivariate analysis of variance (PERMANOVA) tests the distance between the centroids of each data group, while one-way analysis of similarity (ANOSIM) tests group dissimilarity. Both multivariate tests involve 9999 permutations as standard, and pairwise comparisons were performed using Euclidean distances. PERMANOVA gives p and F values, testing for differences between groups by distance; a larger F value indicates a more pronounced group separation. ANOSIM gives p and R values, using the mean rank of distances between and within groups; an R value of 1 indicates complete dissimilarity.

Chapter 3: Establishing A FRET-Based Interaction Assay for the Calcitonin Family of GPCRs and the RAMPs

3.1 – Introduction

As discussed in sections 1.3.2 and 1.4, the calcitonin family of GPCRs are modulated by a group of three accessory proteins called the receptor activity modifying proteins, or RAMPs, which lead to distinct receptor phenotypes and pharmacological profiles. As such, the capabilities of both the calcitonin receptor (CTR), and the calcitonin receptor-like receptor (CLR) expand to include binding affinity for calcitonin, amylin, and the potent vasodilators CGRP and adrenomedullin. As this was such a unique mechanism of action for GPCRs, the interactions between these receptors and RAMPs have been a significant focal point over the last thirty years, particularly in terms of pharmacological characterisation and the elucidation of their structural details. Currently, as highlighted by Kotliar *et al.* 2023, there is now a need to develop drug screening assays which include both receptors and RAMPs, in order to discover effective therapeutics which have a greater chance of success at clinical trial (Kotliar *et al.*, 2023).

As opposed to therapeutically targeting the orthosteric sites of GPCRs, it is hypothesised that the receptor-RAMP interface itself may prove to be a druggable allosteric site. Pragmatic drug screening approaches might therefore take drug-induced changes in receptor-RAMP interactions into account, in addition to the well-characterised pharmacology of these receptors. Consequently, this chapter details the creation of an interaction assay between the calcitonin family of GPCRs and the RAMPs, which is able to detect changes in distance between these membrane proteins, and is applicable to high-throughput drug screening approaches. This assay would be first implemented in a mammalian cell expression system, however, due to the variable and often unpredictable nature of transient transfections, a stable alternative would be attempted within a *Drosophila* expression system; the interaction assay itself, is one based upon fluorescence and FRET.

Förster resonance energy transfer, or FRET, is a mechanism which describes the non-radiative transfer of energy between the chromophores of donor and acceptor fluorescent molecules, and forms the basis of this GPCR-RAMP interaction assay (Figure 3.1) (Forster, 1946; Hussain, 2009). An excited donor chromophore is able to transfer energy to, and excite, an acceptor chromophore via dipole-dipole coupling, for which the efficiency of this energy transfer can be quantified, and is inversely proportional to the sixth power of the distance between donor and acceptor. This makes FRET very sensitive to small changes in distance,

and the quantification of FRET efficiencies can be used to determine if donor and acceptor molecules are within a certain distance of each other, typically to the scale of <10 nm. FRET is therefore widely used to study the interactions between macromolecules within cells, and is most often achieved via the fusion of compatible GFP variants onto target macromolecules, the most common of which are the CFP (donor, Ex430 nm, Em480 nm) and YFP (acceptor, Ex485 nm, Em530 nm) FRET pair (Bajar *et al.*, 2016).

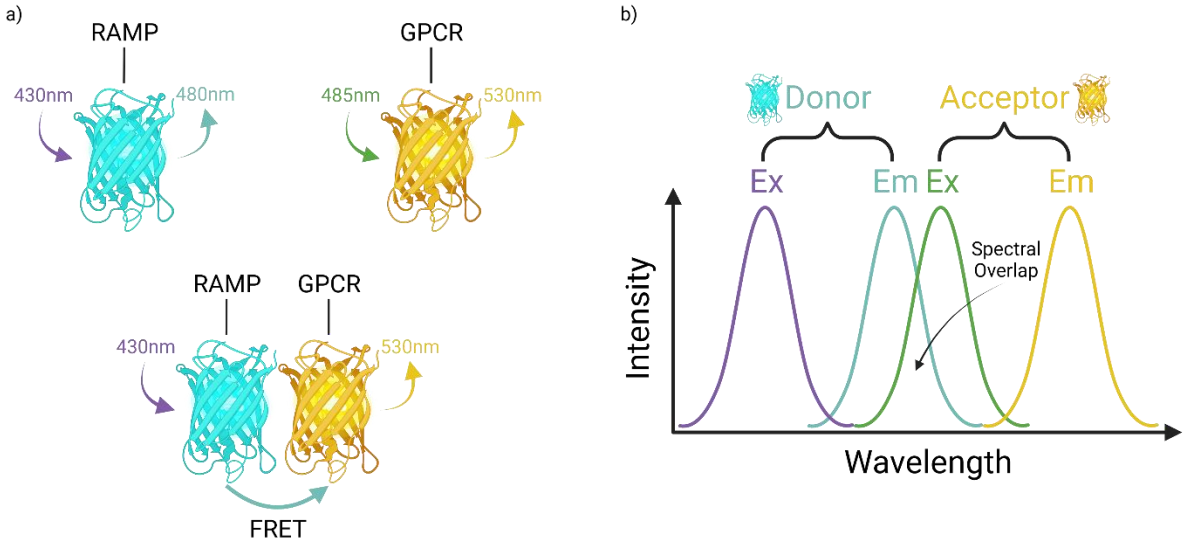

Figure 3.1. The CFP-YFP FRET Pair. The spectra of CFP (donor) emission (Em, 480 nm) overlaps with the spectra of YFP (acceptor) excitation (Ex, 485 nm), enabling non-radiative energy transfer by FRET, over distances less than 10 nm. By tagging each receptor and RAMP with YFP and CFP respectively, the quantification of FRET efficiency enables detection of the GPCR-RAMP complex formation, and determination of the distance between them. Created with BioRender.com.

The aim of this body of work was therefore to tag each of the calcitonin, and calcitonin receptor-like receptors with YFP, and each of the three RAMPs with CFP, creating this interaction assay's FRET pair by molecular biology techniques. These tagged constructs would then be transfected into COS-7 mammalian cells with an optimised protocol, followed by a confirmation of their functionality via cAMP production, before the establishment and optimisation of the FRET-based GPCR-RAMP interaction assay itself. In addition, a preliminary proof of concept experiment was carried out to determine whether *Drosophila melanogaster* could be used as an alternative expression system for this FRET assay, detailed in section 3.5. Please note that this body of work was affected by COVID-19 restrictions.

3.2 – Creation of Fluorescently Tagged Receptors and RAMPs

3.2.1 – Cloning Strategy for Mammalian Expression

In order to achieve the initial objective of producing fluorescently tagged receptors and RAMPs, DNA constructs were first designed *in silico* according to the plan detailed in Methods section 2.1. The overall aim of the cloning strategy was to insert a fluorescent protein (either CFP or YFP) at the C terminus of each protein to be studied (either a RAMP or GPCR), separated by a flexible linker sequence. Both CTR and CLR would be tagged with YFP, whereas each of the three RAMPs would be tagged with CFP. Firstly, sense and anti-sense oligonucleotide primers were designed against each receptor, RAMP, and fluorescent protein, listed in section 2.1.4. Each construct would first have its DNA components created by PCR, adding appropriate restriction sites, allowing for subsequent assembly within the small cloning vector pTZ19R, as illustrated in Figure 3.2.1.1. Using the RAMP1-CFP construct as an example, the RAMP1 fragment of DNA would be first created by PCR and ligated into pTZ19R following the digestion of both with the appropriate restriction enzymes. Next, the C terminus of RAMP1 within pTZ19R would be digested by restriction enzyme, allowing the PCR-created CFP fragment to be ligated into the plasmid at the RAMP1 C terminus. This process would then be repeated for each of the five constructs to be assembled. Finally, each tagged construct would then be digested out of pTZ19R, and ligated into a new vector, pcDNA3.1^{-}, as illustrated in Figure 3.2.1.2. This would enable the expression of the fluorescently tagged constructs in mammalian cell culture, fully achieving the aim of the cloning strategy, as evidenced by sections 3.2.2 to 3.2.4.

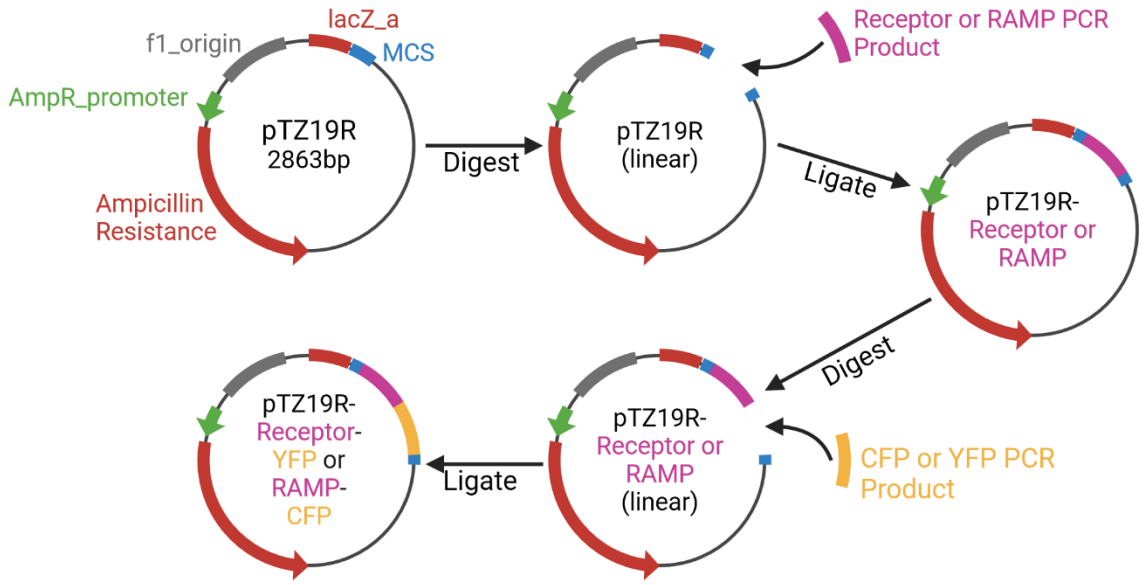

Figure 3.2.1.1. Cloning Strategy to Create Fluorescently Tagged Receptors and RAMPs. Each receptor would be tagged with YFP, and each RAMP would be tagged with CFP. All of the necessary DNA fragments were created and amplified by PCR, with a two-step assembly in the small cloning vector pTZ19R by restriction digest and subsequent ligation. Refer to section 2.1.3 for construct design and restriction enzyme names. Created with BioRender.com.

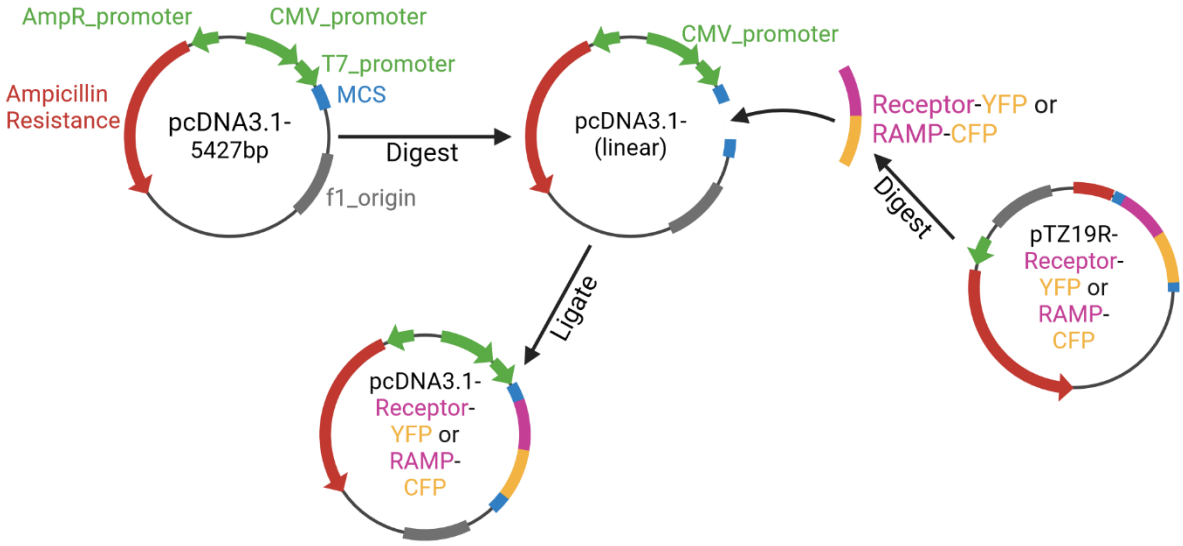

Figure 3.2.1.2. Sub-Cloning Strategy to Insert Fluorescently Tagged Receptors and RAMPs into the Mammalian Expression Vector, pcDNA3.1^{-}. Each of the receptor-YFP and RAMP-CFP constructs would be restriction digested out of the cloning vector pTZ19R, and subsequently ligated into the mammalian expression vector pcDNA3.1^{-}. Refer to section 2.1.3 for construct design and restriction enzyme names. Created with BioRender.com.

3.2.2 – Creation of Construct Fragments by PCR

Initially, the conditions for polymerase chain reaction (PCR) were extensively optimised for each construct fragment in terms of template DNA and oligonucleotide primer concentration, annealing temperatures and extension time, as outlined in Methods section 2.1.5. Ultimately, each of the individually optimised PCR conditions produced a strong, pure PCR product, each of which was ideal for subsequent cloning attempts. As shown in Figure 3.2.2.1, each of RAMP1 (lanes 1 and 2), RAMP2 (lanes 3 and 4) and RAMP3 (lanes 5 and 6) gave a band of approximately 0.4 kb, and each of CFP (lanes 7 and 8) and YFP (lanes 9 and 10) gave a band of approximately 0.8 kb. In addition, as shown in Figure 3.2.2.2, each of CTR (lanes 1 and 2) and CLR (lanes 3 and 4) gave a band of approximately 1.5 kb. Lanes 5 to 8 of this figure are explained in Figure 3.2.3.2. The PCR product for RAMP1 was expected to be 432 bp, RAMP2 was 459 bp, RAMP3 was 429 bp, CFP and YFP were 729 bp, CTR was 1530 bp, and CLR was 1482 bp. Each of these PCR products, demarcated by the red boxes, had therefore been amplified to the fragment sizes expected, based upon the initial cDNA sequences, and could be carried forward into cloning attempts.

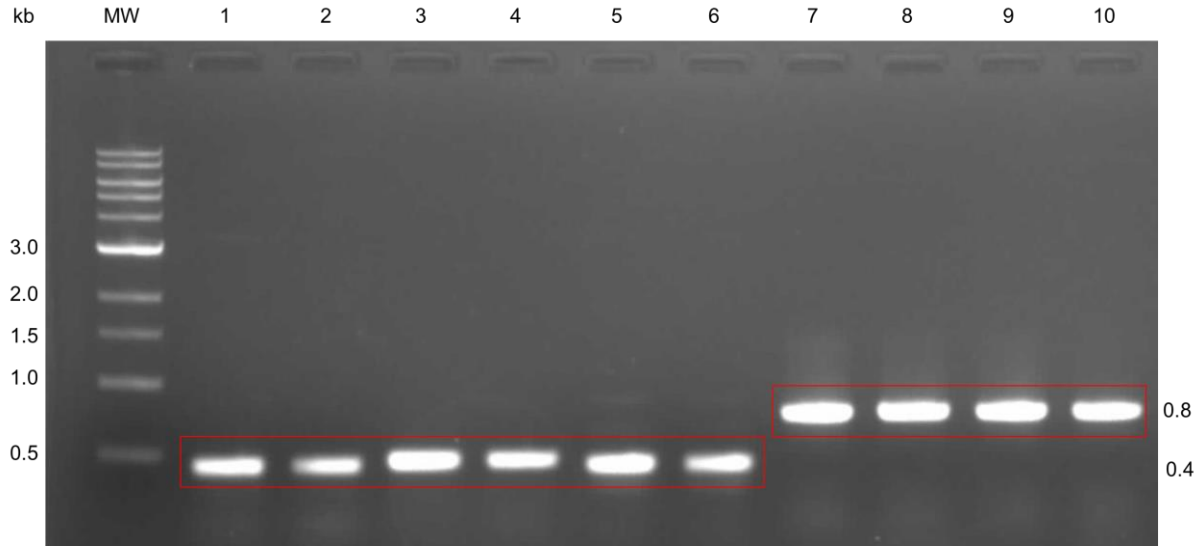

Figure 3.2.2.1. Receptor Activity Modifying Protein (RAMP) and Fluorescent Protein (CFP/YFP) PCR Products. When amplified using KOD polymerase and oligonucleotide primers listed in section 2.1.4, RAMP1 (lanes 1 and 2), RAMP2 (lanes 3 and 4) and RAMP3 (lanes 5 and 6) each resulted in an approximate band at 0.4 kb. In addition, CFP (lanes 7 and 8) and YFP (lanes 9 and 10) each resulted in an approximate band at 0.8 kb. MW is the NEB 1 kb DNA ladder.

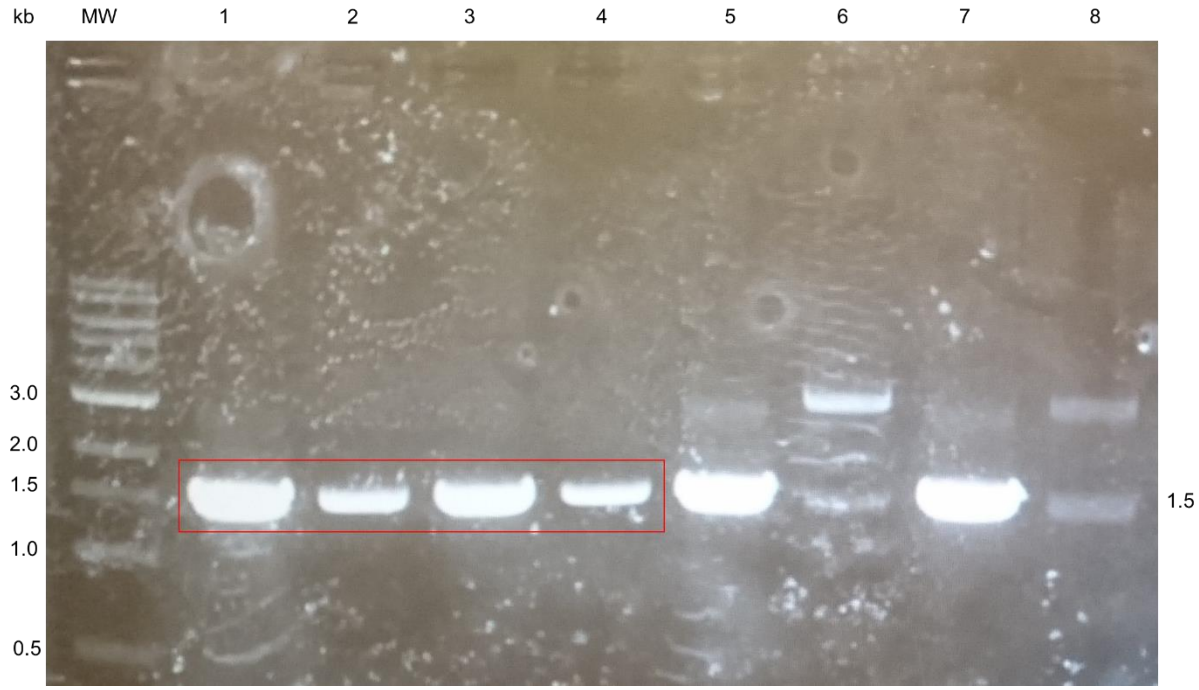

Figure 3.2.2.2. Calcitonin Receptor (CTR) and Calcitonin Receptor-Like Receptor (CLR) PCR Products. When amplified using KOD polymerase and oligonucleotide primers listed in section 2.1.4, CTR (lanes 1 and 2) and CLR (lanes 3 and 4) each resulted in approximate bands at 1.5 kb. MW is the NEB 1 kb DNA ladder. The red box demarcates the lanes relevant to this experiment; lanes 5 to 8 are not relevant to this figure, but are explained in Figure 3.2.3.2.

3.2.3 – Construct Assembly In pTZ19R

Prior to cloning attempts, each of the PCR products created in section 3.2.2 were first purified through a PCR cleanup kit, digested with the appropriate restriction enzymes detailed in section 2.1.4, run on an agarose gel, and purified a second time with a gel extraction kit. In tandem, the cloning vector pTZ19R was also digested with the corresponding restriction enzymes for each fragment, run on an agarose gel, and purified with the gel extraction kit. Each of the purified PCR products and linearised plasmid were then combined into several ligation reactions with T7 DNA ligase, and subsequently transformed into DH5α *E. coli* cells. These were then grown on agar plates at 37°C, and colonies were selected to be screened for successful clones by diagnostic digest and Sanger sequencing.

As shown in Figure 3.2.3.1, each of the three RAMPs, and both CFP and YFP were successfully digested out of the cloning vector pTZ19R, resulting in two bands being present within each lane. RAMP1 (lanes 1-3), RAMP2 (lane 4) and RAMP3 (lanes 7-9) each gave a band of approximately 0.4 kb, CFP (lanes 10-12) and YFP (lanes 13-15) each gave a band of approximately 0.8 kb, as demarcated by the red boxes, and finally, the linearised pTZ19R plasmid gave a band of approximately 2.9 kb with each digest (lanes 1-15). In addition, Figure

3.2.3.2 shows the diagnostic digest of both CTR (lane 6) and CLR (lane 8), both giving a band of approximately 1.5 kb when digested from pTZ19R (lanes 6 and 8), as demarcated by the red boxes, with pTZ19R giving a band of 2.9 kb each time. Lanes 1 to 4 are explained in Figure 3.2.2.2. Note that the upper band in lanes 1, 2, 3, 5 and 6 presented as a double-band which is explained by an incomplete digestion of the plasmid. Overall, these results mean each construct fragment had been successfully cloned into the pTZ19R plasmid, the sequences of which were then confirmed to be as expected by Sanger sequencing.

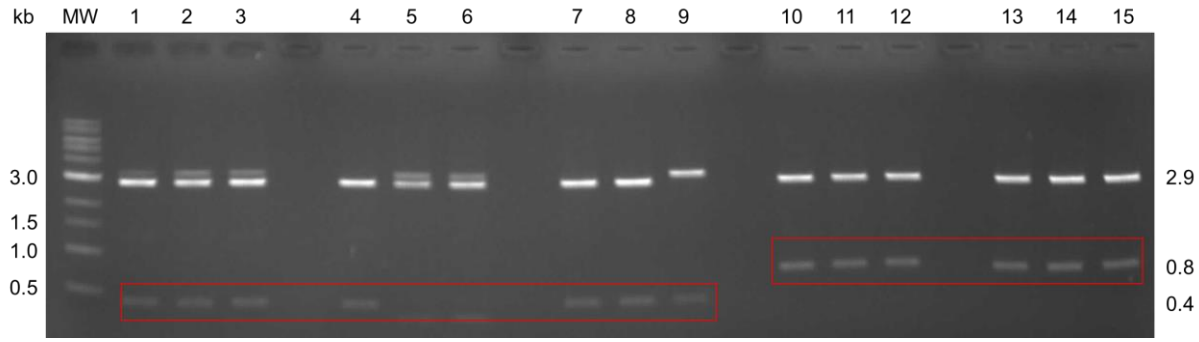

Figure 3.2.3.1. Restriction Digestion of Cloning Vector pTZ19R Containing RAMPs1-3, CFP and YFP. Following their ligation into pTZ19R with T7 DNA ligase, RAMP1 (lanes 1-3), RAMP2 (lane 4) and RAMP3 (lanes 7-9) each result in an approximate band at 0.4 kb when digested with the appropriate restriction enzymes listed in section 2.1.4. In addition, CFP (lanes 10-12) and YFP (lanes 13-15) each result in an approximate band at 0.8 kb, while the linearised host vector pTZ19R results in a band at 2.9 kb. MW is the NEB 1 kb DNA ladder. The red box demarcates the lanes relevant to this experiment.

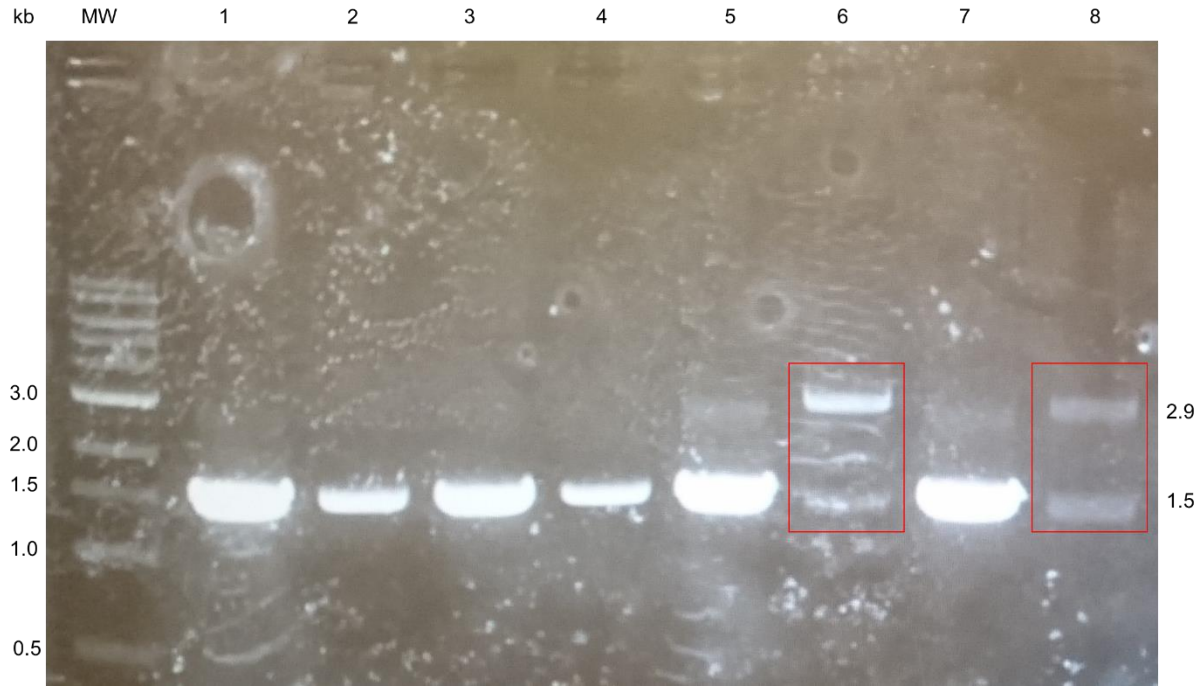

Figure 3.2.3.2. Restriction Digestion of Cloning Vector pTZ19R Containing CTR and CLR. Following their ligation into pTZ19R with T7 DNA ligase, CTR (lane 6) and CLR (lane 8) each result in an approximate band at 1.5 kb when digested with the appropriate restriction enzymes listed in section 2.1.4. In addition, the linearised host vector pTZ19R results in an approximate band at 2.9 kb. MW is the NEB 1 kb DNA ladder. The red box demarcates the lanes relevant to this experiment; lanes 1 to 5 and lane 7 are not relevant to this figure, and lanes 1 to 4 are explained in Figure 3.2.2.2.

Following this initial insertion of each construct fragment into pTZ19R, the second step of assembly first involved the restriction digest of each receptor and RAMP at their C termini in order to ligate either YFP or CFP, respectively, to form the tagged constructs in pTZ19R. Following restriction digestion, each sample was again run on an agarose gel and purified with a gel extraction kit prior to ligation. These were also transformed into DH5α *E. coli* cells, and colonies screened for successful clones. As shown in Figure 3.2.3.3, each of RAMP1-CFP (lane 2), RAMP2-CFP (lanes 4-6), and RAMP3-CFP (lanes 7-9) gave a band of approximately 1.2 kb when digested from pTZ19R (lanes 1-9) which gave bands of approximately 2.9 kb each time. Lanes 1 and 3 were unsuccessful cloning attempts. In addition, CTR-YFP (lane 2) gave a band of approximately 2.3 kb when digested from pTZ19R (lane 2) which gave a band of approximately 2.9 kb, as shown in Figure 3.2.3.4. Lane 1 is explain in Figure 3.2.4.2, and lanes 3 to 6 are irrelevant to this experiment. These results mean that the fluorescently tagged constructs of the three RAMPs and CTR had been successfully assembled within pTZ19R, and were again confirmed to be of the expected sequence via Sanger sequencing. Unfortunately, after several unsuccessful attempts to assemble the CLR-YFP construct, the decision was ultimately made to outsource its synthesis, so as not to further waste time. It was

unclear as to exactly why this construct could not be cloned; it was proposed that the *Xma*I restriction site or enzyme may have been responsible, as it was unique to this construct. In any case, the four successful clones were progressed into the final step of the cloning strategy.

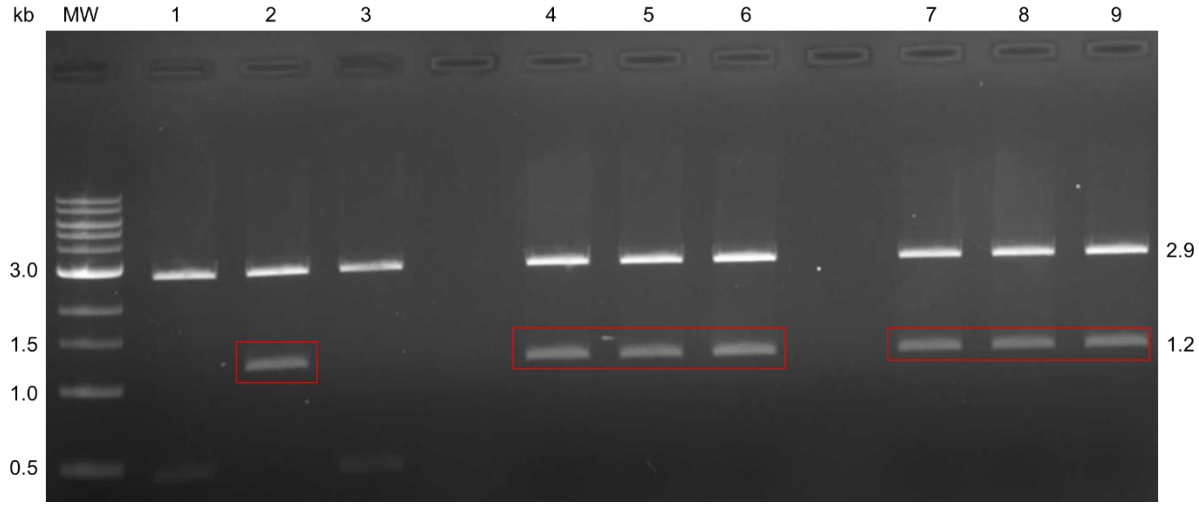

Figure 3.2.3.3. Restriction Digestion of RAMP-CFP Constructs Assembled in pTZ19R. Following their assembly in pTZ19R, RAMP1-CFP (lane 2), RAMP2-CFP (lanes 4-6) and RAMP3-CFP (lanes 7-9) each result in an approximate band at 1.2 kb when digested with the appropriate restriction enzymes listed in section 2.1.4. The linearised host vector pTZ19R results in an approximate band at 2.9 kb. MW is the NEB 1 kb DNA ladder. The red box demarcates the lanes relevant to this experiment; lanes 1 and 3 were unsuccessful cloning attempts.

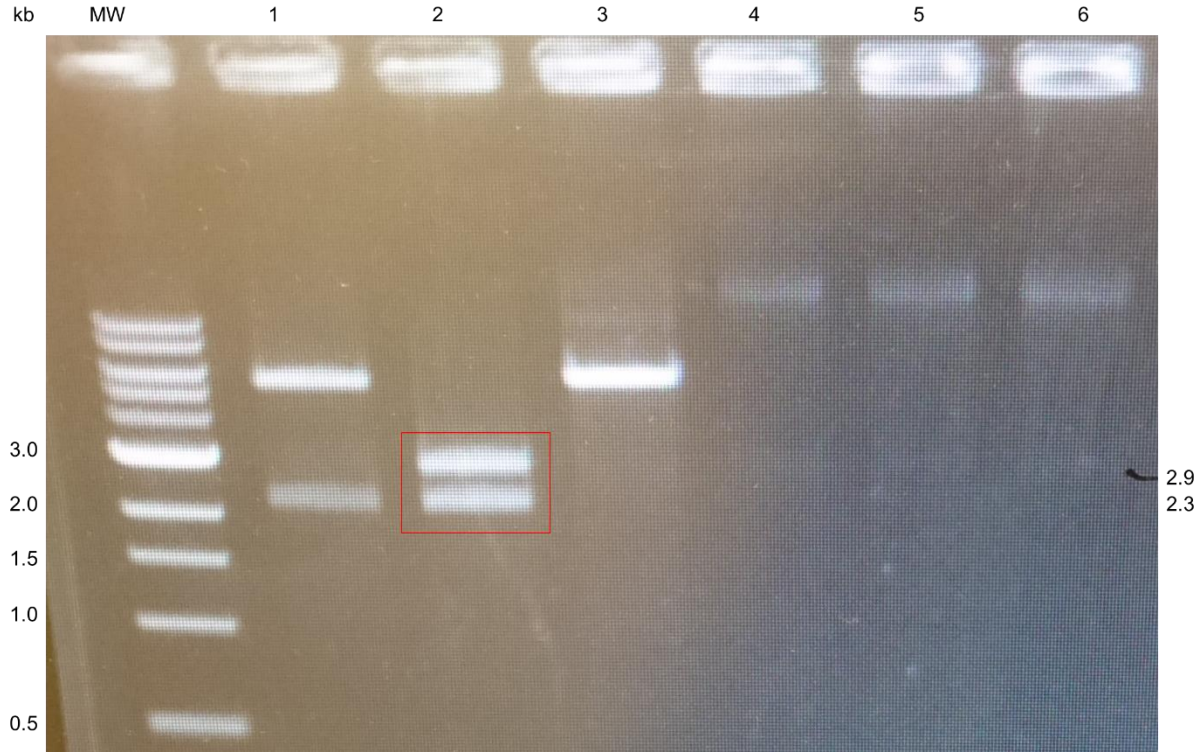

Figure 3.2.3.4. Restriction Digestion of the CTR-YFP Construct Assembled in pTZ19R. Following their assembly in pTZ19R, CTR-YFP (lane 2) results in an approximate band at 2.3 kb when digested with the appropriate restriction enzymes listed in section 2.1.4. The linearised host vector pTZ19R results in an approximate band at 2.9 kb. MW is the NEB 1 kb DNA ladder. The red box demarcates the lanes relevant to this experiment; lane 1 is explained in Figure 3.2.4.2, and lanes 3 to 6 are irrelevant.

3.2.4 – Subcloning Into pcDNA3.1$^-$

While the small cloning vector pTZ19R is suited to construct assembly and plasmid amplification within *E. coli* cells, a separate vector is needed for construct expression in mammalian cells; pcDNA3.1$^-$ was selected for this task. Simply, each of the three RAMP-CFP and CTR-YFP constructs were digested out of pTZ19R, and ligated into the complementary restriction sites in the similarly digested pcDNA3.1$^-$ plasmid. Once again, these ligation reactions were transformed into DH5α *E. coli* cells, and colonies screened for successful sub-clones. As shown in Figure 3.2.4.1, RAMP1-CFP (lanes 1 and 2), RAMP2-CFP (lanes 3 and 4) and RAMP3-CFP (lanes 5 and 6) each gave a band of 1.2 kb when digested from pcDNA3.1$^-$ (lanes 1-6) which gave bands of 5.5 kb each time. Note that the upper bands presented as double-banded which is explained by incomplete restriction digestion. In addition, as shown in Figure 3.2.4.2, CTR-YFP (lane 1) gave a band of 2.3 kb when digested from the 5.5 kb pcDNA3.1$^-$ plasmid (lane 1). Lane 2 is explained by Figure 3.2.3.4, and lanes 3 to 6 are irrelevant to this experiment. Again, it should be noted that the CLR-YFP construct

synthesis had been outsourced. Overall, these results mean the RAMP-CFP and CTR-YFP constructs had been successfully ligated into pcDNA3.1^{-}, confirmed by Sanger sequencing. The cloning strategy of section 3.2.1 was therefore achieved, and the constructs were thus ready for transfection and expression within mammalian cell cultures.

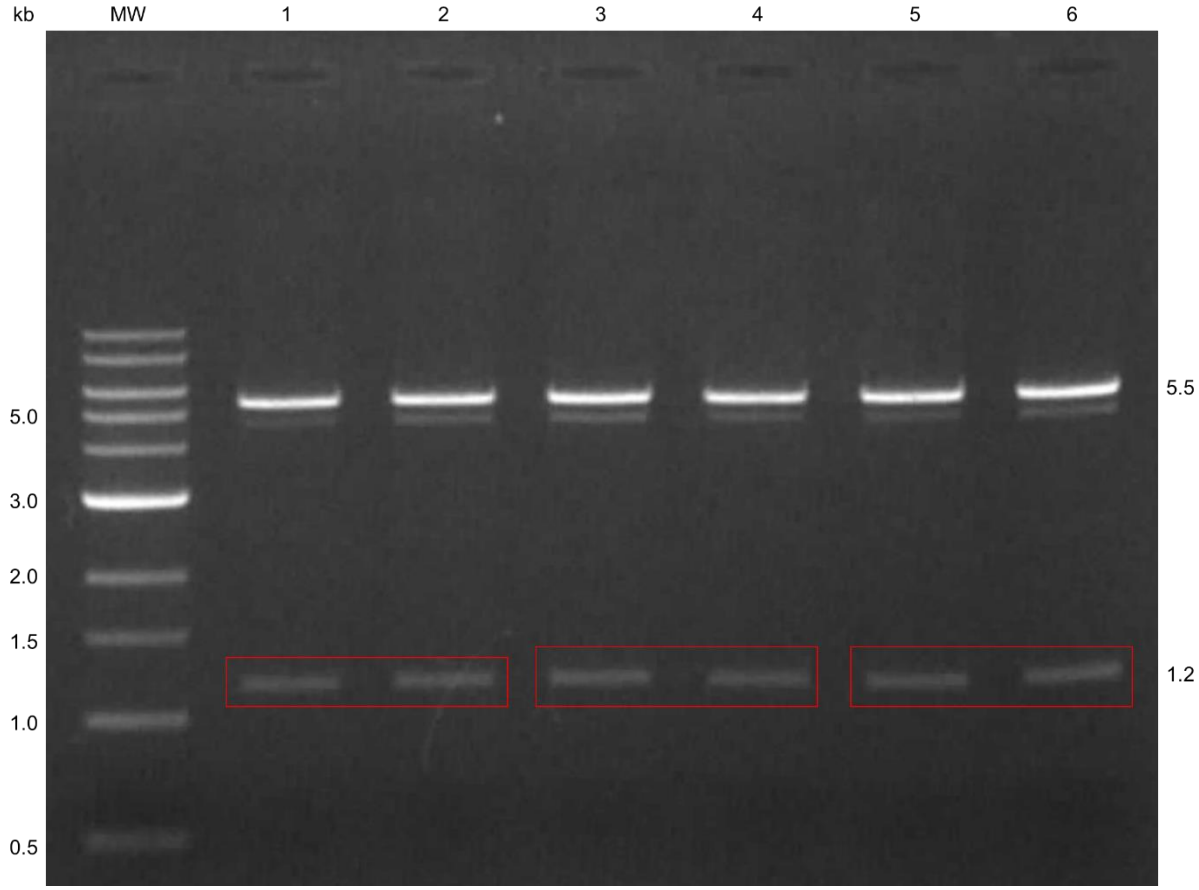

Figure 3.2.4.1. Restriction Digestion of RAMP-CFP Constructs Sub-Cloned into pcDNA3.1^{-}. Following sub-cloning into pcDNA3.1^{-}, RAMP1-CFP (lanes 1 and 2), RAMP2-CFP (lanes 3 and 4) and RAMP3-CFP (lanes 5 and 6) each result in an approximate band at 1.2 kb when digested with the appropriate restriction enzymes listed in 2.1.4. The linearised host vector pcDNA3.1^{-} results in an approximate band at 5.5 kb. MW is the NEB 1 kb DNA ladder.

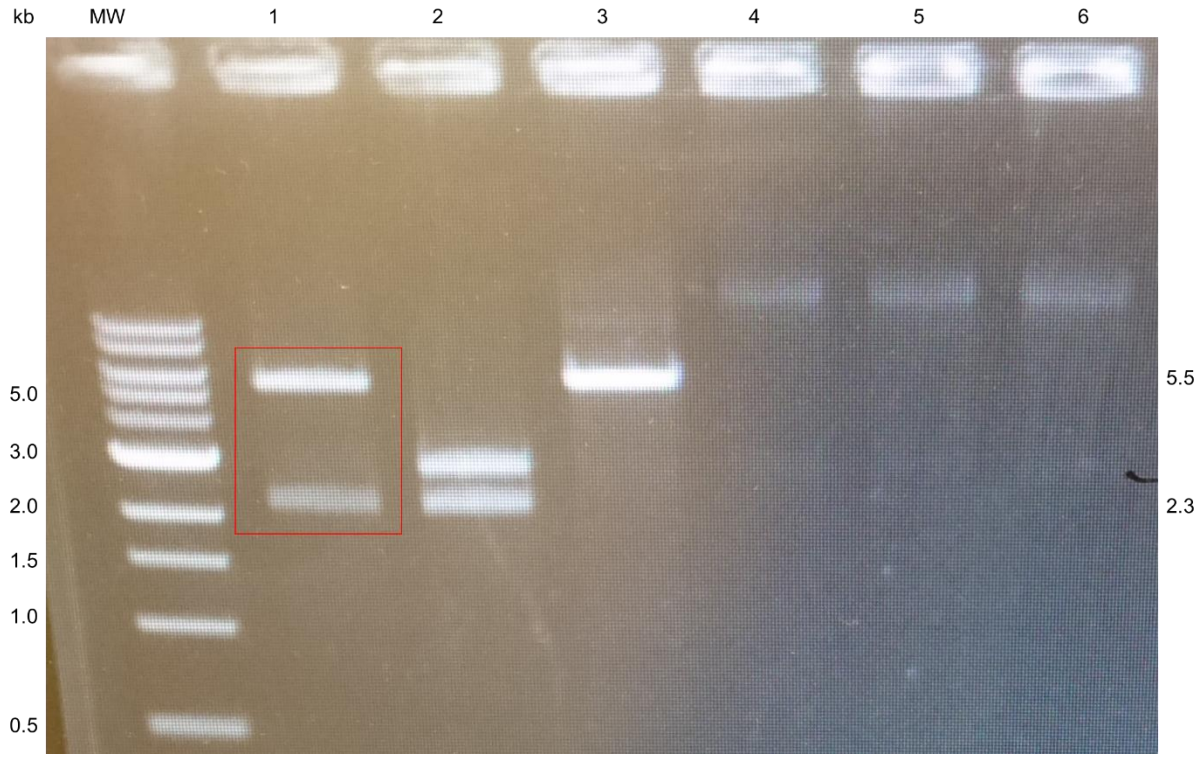

Figure 3.2.4.2. Restriction Digestion of the CTR-YFP Construct Sub-Cloned into pcDNA3.1$^-$. Following sub-cloning into pcDNA3.1$^-$, CTR-YFP (lane 1) results in an approximate band at 2.3 kb when digested with the appropriate restriction enzymes listed in section 2.1.4. The linearised host vector pcDNA3.1$^-$ results in an approximate band at 5.5 kb. MW is the NEB 1 kb DNA ladder. The red box demarcates the lanes relevant to this experiment; lane 2 is explained in Figure 3.2.3.4, and lanes 3 to 6 are irrelevant.

3.3 – Confirmation of Construct Expression and Functionality in Mammalian Cells

3.3.1 – Optimisation of Transfection Protocols

Given that experiments involving the expression or functionality of membrane proteins rely on efficient and effective transfections, a screening was next performed to identify the optimal transfection reagent and experimental conditions for the expression of the created constructs in mammalian cells. Receptor-YFP constructs were transfected alone, as well as conditions including a RAMP-CFP construct, as outlined in Methods section 2.1.6. COS-7 cells were seeded in complete DMEM medium at 20,000 cells per well in a 96-well format and transiently transfected with 0.25 µg total DNA, 24 hours after seeding; when co-transfecting both receptor and RAMP, this equates to 0.125 µg of each. Two established transfection methods were trialled, using either polyethylenimine (PEI) or the Lipofectamine 2000 reagent. Transfected cells were incubated at 37°C for both 24 and 48 hours post-transfection before measurement

at 480-530 nm in a fluorescent microplate reader. All values have been corrected to the background signal produced by untransfected cells, and are presented as relative light units (RLU), as shown in Figure 3.3.1.

Firstly, contrary to some established protocols, a 24 hour transfection incubation (data not shown) simply did not work, failing to produce fluorescence significantly higher than untransfected cells. However, a transfection incubation of 48 hours was found to be broadly successful (Figure 3.3.1), with the expression of the tagged constructs fluorescing significantly higher than untransfected cells. As shown in Figure 3.3.1a, several DNA:PEI ratios were screened, with 1:3 performing the best overall, giving an RLU range of 90 to 184, with an average of 133. Despite these successful transfections however, PEI was found to be a less expensive but more unreliable reagent and thus Lipofectamine 2000 was screened as an alternative option. As shown in Figure 3.3.1b, Lipofectamine 2000 gave lower RLU measurements overall, a range of 33 to 101, with an average of 70, but was found to be much more reliable across separate transfections and experiments. As such, the decision was made to solely use Lipofectamine 2000 with 48 hour incubations before cell harvesting, with this reagent being utilised in all further experiments involving the transfection of COS-7 cells.

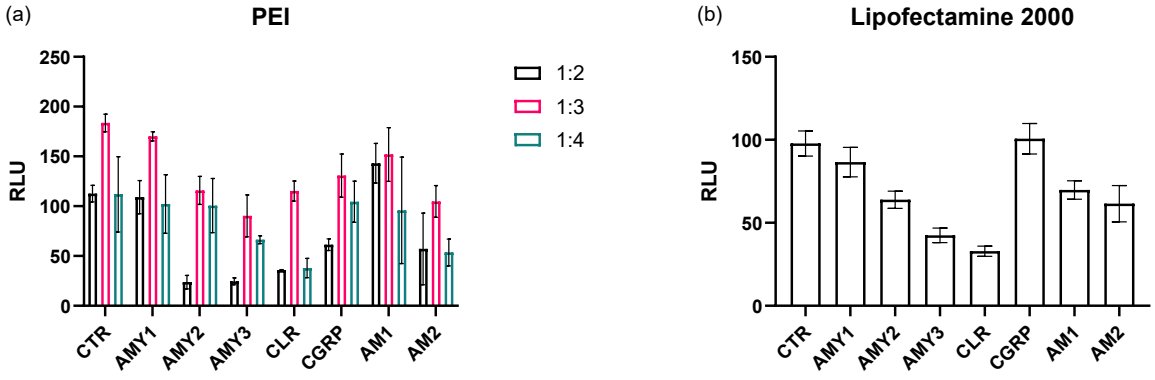

Figure 3.3.1. A Screening of Transfection Reagents. COS-7 cells were transiently transfected for 48 hours with fluorescently tagged constructs in a 96-well format, with 20,000 cells and 0.25 µg total DNA per well. 1 mg/ml PEI was trialled in a range of DNA:PEI ratios (a) as well as Lipofectamine 2000 which used a 1:1 ratio according to manufacturer instructions (b). Refer to Table 1.3.2 for details on these receptor-RAMP abbreviations; CTR is calcitonin receptor alone, AMY1 is CTR and RAMP1, AMY2 is CTR and RAMP2, AMY3 is CTR and RAMP3, CLR is calcitonin receptor-like receptor alone, CGRP is CLR and RAMP1, AM1 is CLR and RAMP2, and AM2 is CLR and RAMP3. Data is presented as relative light units (RLU), with overall fluorescence measured at ex 480 nm, em 530 nm from at least three independent experiments. Error bars represent the standard error of the mean (SEM).

3.3.2 – Production of cAMP in Response to CGRP

Consequently, after the target constructs had been created successfully and their transfections confirmed to be effective in COS-7 cells, it was next necessary to determine whether the modifications to the stock constructs had affected their ability to function pharmacologically, as receptor-RAMP complexes should. Given that the calcitonin family of receptors can be stimulated to produce cAMP with several calcitonin-related peptide ligands, both the untagged and fluorescently-tagged constructs were transfected into COS-7 cells and their cAMP production quantified after stimulation with the calcitonin gene-related peptide (CGRP). Note that the untagged constructs were the equivalent receptors and RAMPs within pcDNA3.1$^-$ without CFP or YFP, utilised as non-fluorescent or unmodified controls. CGRP itself was chosen as all of the selected receptor-RAMP combinations (CTR, AMY1-3, CGRP, AM1 and AM2 receptors) respond to this peptide ligand, albeit to varying extents. Cells were transfected and stimulated as outlined in Methods section 2.17 and subsequently assayed for cAMP production with the AlphaScreen kit manufactured by Perkin Elmer, Inc.

Firstly, a series of cAMP standards were prepared in duplicate, between 10 pM and 1 μM as shown in Figure 3.3.2.1a; as the immunoassay itself is competitive, an increase in cAMP results in a reduction in the luminescent AlphaScreen assay signal. Non-linear regression was performed on this log(agonist) vs response curve which determined its top and bottom values to be 10586 and 1010, its Hill slope to be -0.7298, its logIC50 to be -9.53, and an R^2 value of 0.9965. Now importantly, the AlphaScreen signals detected from the transfected cell lysates would not fit onto this standard curve, thus it was therefore interpolated and transformed to produce a simulated standard curve as shown in Figure 3.3.2.1b. This process expands the standard curve, but importantly maintains the parameters of the original curve's non-linear regression, such that the top and bottom values became 111362 and 4815, but the Hill slope and logIC50 values remained to be -0.7298 and -9.53, with an R^2 value of 1. Essentially, as opposed to trying to fit assay data onto a standard curve, this process expands the standard curve to fit the assay data obtained; the curve's regression remains the same 'shape', but its range is extended, and can therefore estimate the amount of cAMP produced by cell samples.

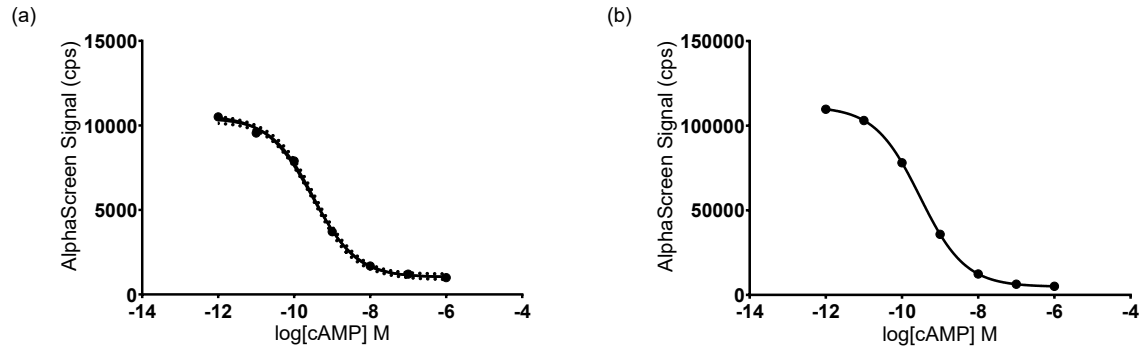

Figure 3.3.2.1. AlphaScreen cAMP Assay Standard Curves. (a) Standards of cAMP were prepared in duplicate between 10 pM and 1 μM and their AlphaScreen signals determined by competitive immunoassay, giving a logIC50 value of -9.53. This standard curve was then interpolated to simulate a second curve (b) appropriate for estimating the amount of cAMP produced by cell samples. Importantly, this extends the range of the standards to fit the AlphaScreen data obtained, but maintains the parameters of the original regression. Data plotted as four-parameter non-linear regression, as an average of n=3.

Next, the AlphaScreen signal values for cells transfected with the untagged and fluorescently-tagged constructs were similarly processed. These values were interpolated onto the simulated standard curve to give an estimation of the sample's cAMP production, transformed to the inversed log format (X=10^X), and represented as the amount of cAMP produced in pmol/mg compared to the concentration of CGRP that sample was stimulated by (log[CGRP] M). The dose-response curves of each receptor-RAMP combination are shown in Figures 3.3.2.2 and 3.3.2.3, with comparison between the untagged (black) and tagged (pink) versions of the constructs, the parameters of which are summarised in Table 3.3.2. In addition, the differences between the untagged and tagged construct dose-responses were each tested for significance by a two-tailed, unpaired t-test of which none were found to be significantly different (*p* values listed in Table 3.3.2). Overall, these results therefore mean that fluorescently tagging the receptor and RAMP constructs with either YFP or CFP did not significantly alter their ability to increase levels of cAMP in response to CGRP.

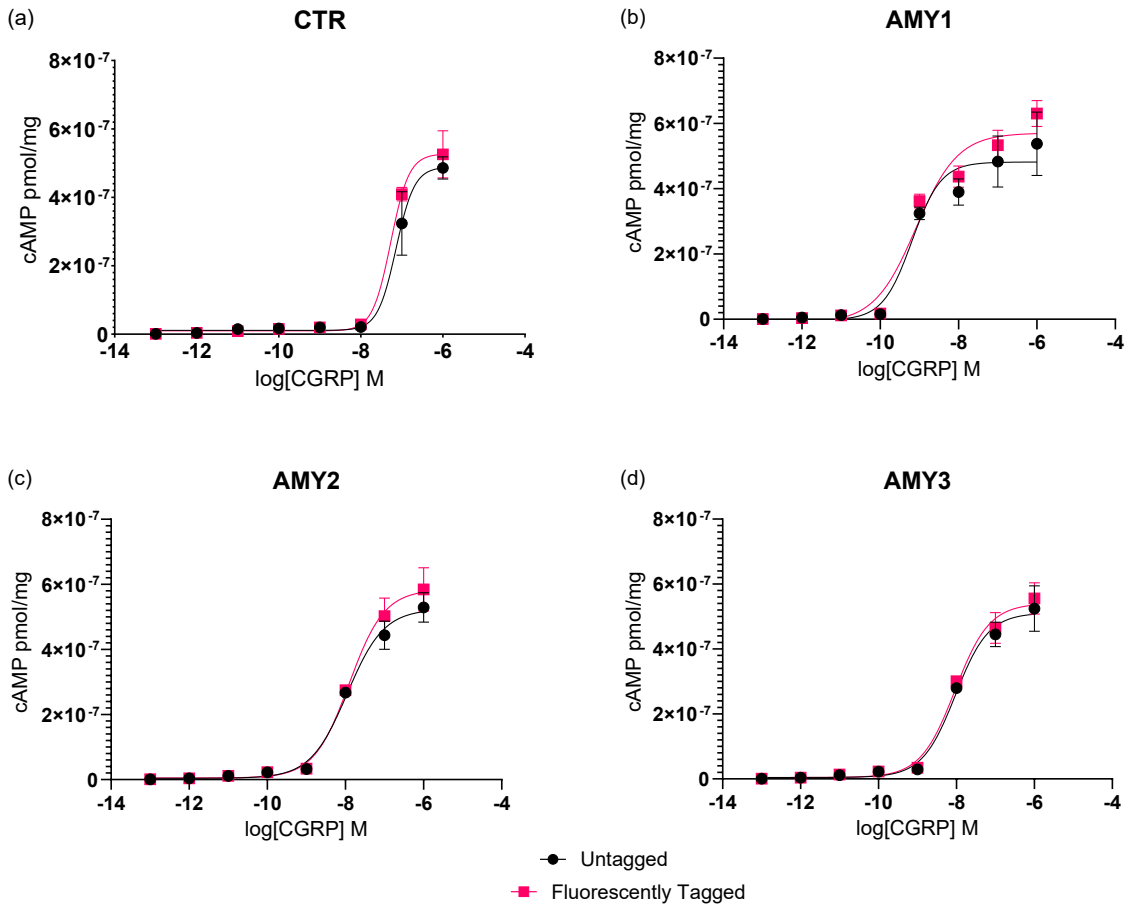

Figure 3.3.2.2. CGRP-Stimulated Production of cAMP in Cells Expressing the Calcitonin or Amylin Receptors. COS-7 cells were transfected with untagged (black) and fluorescently-tagged (pink) constructs corresponding to the calcitonin (CTR) and amylin (AMY1 to 3) receptors; refer to Table 1.3.2 for detail on these receptor-RAMP phenotypes. After stimulation with CGRP agonist, the cellular production of cAMP was determined by AlphaScreen immunoassay, interpolating from the cAMP standards run in-parallel (Figure 3.3.2.1). The parameters of each curve are summarised in Table 3.3.2, along with the results of any significant differences between untagged and tagged curves. Data is an average of n=3, and error bars represent the standard error of the mean.

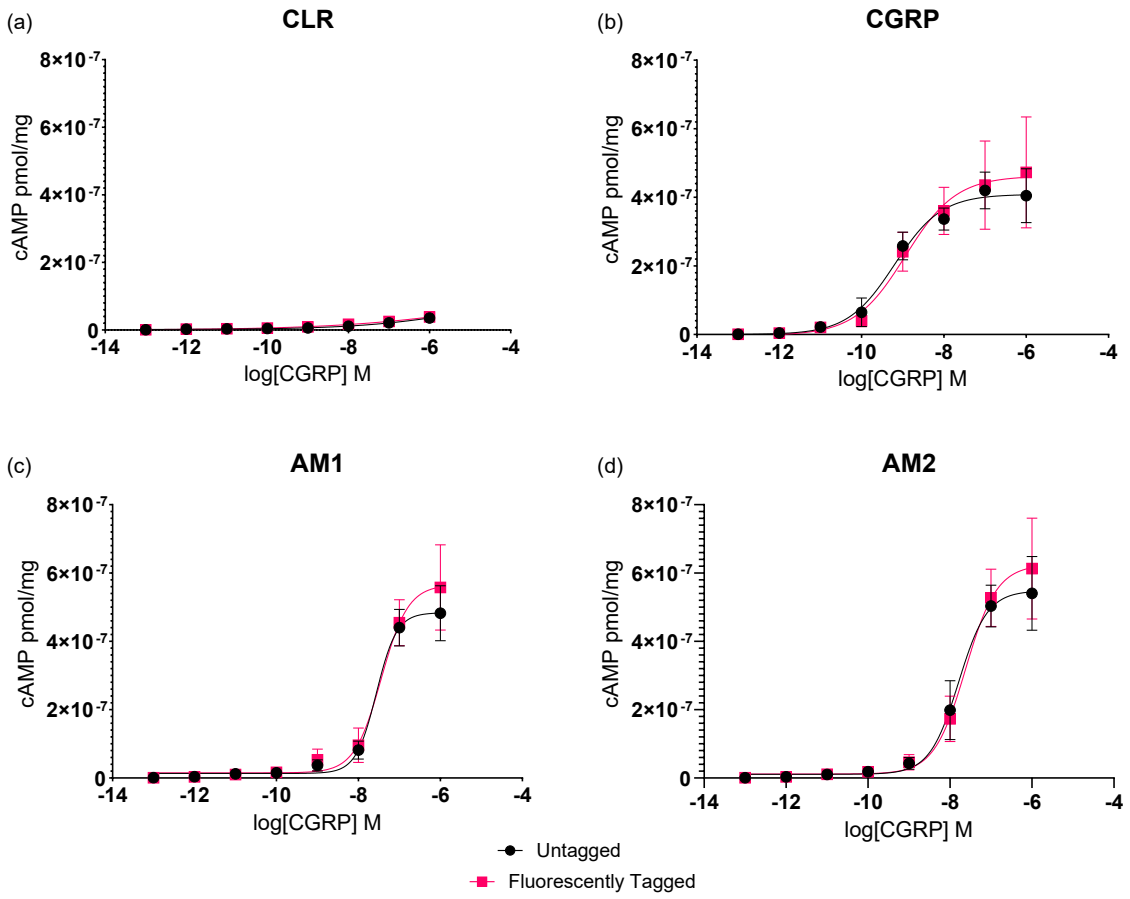

Figure 3.3.2.3. CGRP-Stimulated Production of cAMP in Cells Expressing the CGRP or Adrenomedullin Receptors. COS-7 cells were transfected with untagged (black) and fluorescently-tagged (pink) constructs corresponding to the calcitonin receptor-like receptor (CLR), CGRP and adrenomedullin (AM1 and AM2) receptors; refer to Table 1.3.2 for detail on these receptor-RAMP phenotypes. After stimulation with CGRP agonist, the cellular production of cAMP was determined by AlphaScreen immunoassay, interpolating from the cAMP standards run in-parallel (Figure 3.3.2.1). The parameters of each curve are summarised in Table 3.3.2, along with the results of any significant differences between untagged and tagged curves. Data is an average of n=3, and error bars represent the standard error of the mean.

Table 3.3.2. A Summary of Dose-Response Parameters from AlphaScreen cAMP Immunoassays. The parameters for the top, bottom, Hill slope, logEC50 and R^2 values are summarised for each curve in Figures 3.3.2.2 and 3.3.2.3, with comparison between the untagged and fluorescently-tagged versions of each receptor-RAMP combination. logEC50 values for CLR and CLR-YFP were not determined (ND) due to their lack of signalling in the absence of a RAMP. In addition, the results of two-tailed, unpaired t-tests are shown along with *p* values, indicating there were no significant differences between cells transfected with either untagged or tagged receptor-RAMP constructs, in terms of their cAMP production. Data is as an average of n=3 with both biological and technical replicates.

Transfection		Top	Bottom	Hill Slope	logEC50	R^2 Value	Significance
CTR	CTR	4.888e-007	1.110e-008	1.941	-7.143	0.9257	None *p* = 0.8778
CTR-YFP		5.273e-007	9.609e-009	1.950	-7.272	0.9676	
CTR+RAMP1	AMY1	4.816e-007	-2.601e-009	1.169	-9.192	0.8932	None *p* = 0.8253
CTR-YFP+ RAMP1-CFP		5.705e-007	-1.153e-008	0.8145	-9.087	0.9485	
CTR+RAMP2	AMY2	5.239e-007	3.876e-009	0.9238	-7.958	0.9725	None *p* = 0.8960
CTR-YFP+ RAMP2-CFP		5.839e-007	4.907e-009	0.9631	-7.906	0.9620	
CTR+RAMP3	AMY3	5.111e-007	4.395e-009	1.022	-8.028	0.9574	None *p* = 0.9308
CTR-YFP+ RAMP3-CFP		5.404e-007	4.359e-009	0.9841	-8.038	0.9689	
CLR	CLR	1.023e-007	1.461e-009	0.3296	ND	0.7924	None *p* = 0.7258
CLR-YFP		1.252e-007	-3.609e-011	0.2356	ND	0.7657	
CLR+RAMP1	CGRP	4.087e-007	-1.696e-010	0.7662	-9.198	0.8926	None *p* = 0.9363
CLR-YFP+ RAMP1-CFP		4.622e-007	-2.864e-009	0.7060	-8.933	0.7373	
CLR+RAMP2	AM1	4.836e-007	1.324e-008	1.744	-7.563	0.9315	None *p* = 0.8912
CLR-YFP+ RAMP2-CFP		5.654e-007	1.496e-008	1.325	-7.445	0.8779	
CLR+RAMP3	AM2	5.481e-007	1.080e-008	1.234	-7.793	0.8877	None *p* = 0.9384
CLR-YFP+ RAMP3-CFP		6.243e-007	1.152e-008	1.132	-7.623	0.8671	

Finally, after proving that the receptor and RAMP constructs were able to increase cAMP in COS-7 cells after stimulation with CGRP, and that fluorescently tagging them with either YFP or CFP did not cause any significant difference to that increase in cAMP, how do these results compare to those reported in the literature? Figure 3.3.2.4 shows a comparison of the pEC50 values determined in this body of work (untagged = black and tagged = pink) against those detailed in an update to the pharmacology of the calcitonin/CGRP family of peptides (grey) (Hay *et al.*, 2018). These grey data are a collection of pEC50 values reported in the literature,

originating from several research groups, with perhaps slightly different constructs, derived from a variety of cell lines including HEK293 and COS-7. To summarise, the pEC50 values for the various receptor-RAMP combinations in this body of work lie well within the reported ranges for the calcitonin family receptors stimulated specifically with CGRP.

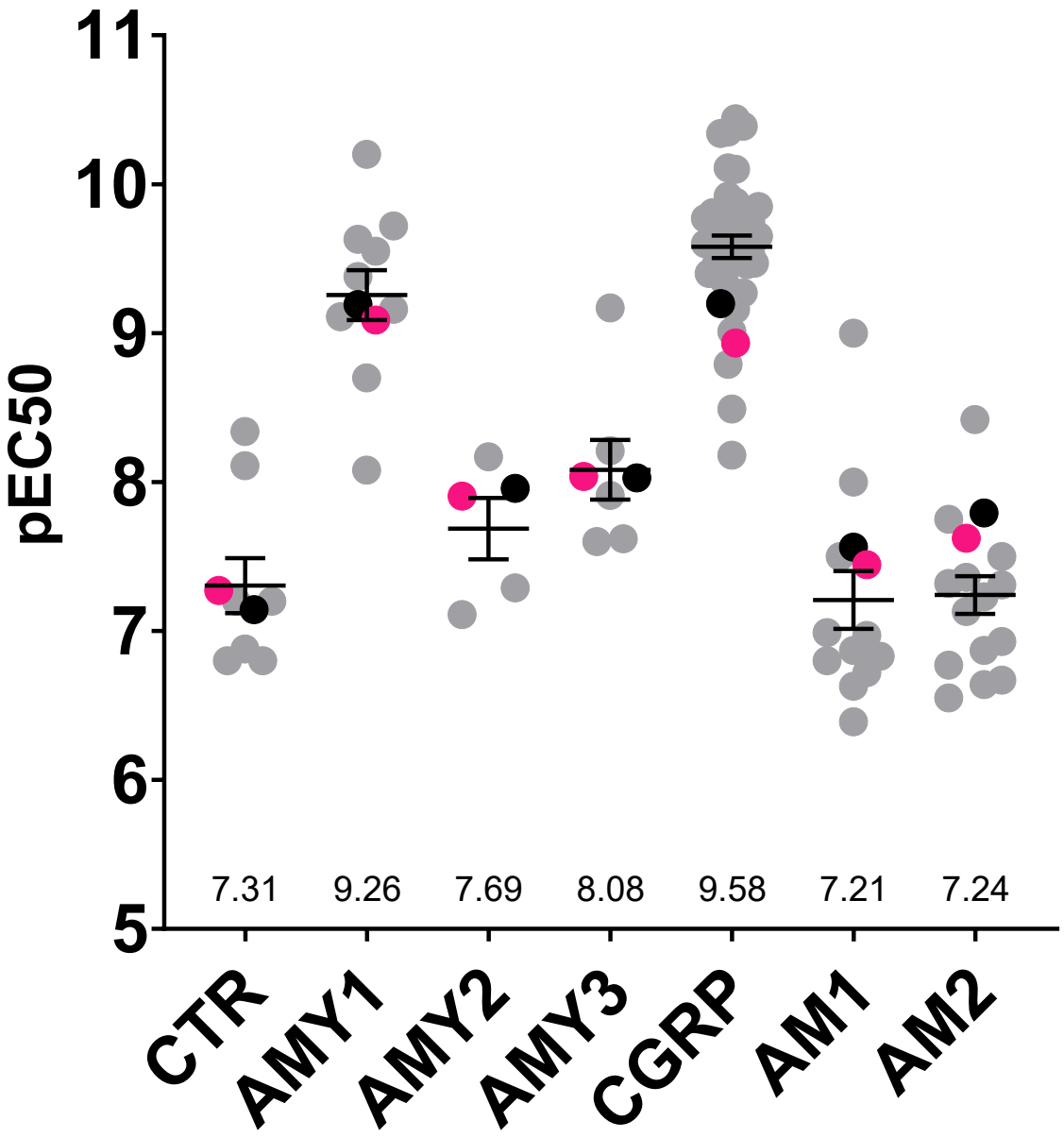

Figure 3.3.2.4. A Comparison of pEC50 Values for the Calcitonin Receptor Family when Stimulated with CGRP. pEC50 values reported in an update to the pharmacology of the calcitonin/CGRP peptide family (grey) (Hay *et al.*, 2018) compared to those determined in this body of work (black = untagged and pink = fluorescently tagged). Mean pEC50 values displayed at the bottom of each dataset and bars show the mean and standard error of the mean.

Overall, these results indicate that the receptor and RAMP constructs designed for the application to a FRET-based interaction assay were not only created successfully, but were able to be effectively transfected into the mammalian COS-7 cell line with the use of Lipofectamine 2000. Crucially, the constructs also demonstrated the expected pharmacological responses to CGRP, including when tagged with the fluorescent proteins needed for FRET itself, and thus were deemed appropriate for use in FRET experiments, especially in the context of drug screening.

3.4 – Detection of Receptor-RAMP Interactions by FRET

After successfully creating the fluorescently-tagged constructs by PCR and molecular cloning, optimising their transfection into COS-7 cells, and gaining an indication of receptor-RAMP functionality in terms of stimulated cAMP production, the constructs were deemed to be fit for purpose, and were therefore utilised in establishing and testing the FRET assay itself. As discovered by the transfection optimisation in section 3.3, COS-7 cells were again transfected for 48 hours in a 96-well format, to a total of 0.25 µg DNA per well, and analysed with a Mithras LB 940 microplate reader; this is based upon a 'gold-standard' method produced by Berthold Technologies to detect GPCR dimerization by FRET (Haase, Kamann and Brüx, 2011).

Cells were transfected with various combinations of either receptor or RAMP alone, both receptor and RAMP, and indeed both tagged and untagged versions of them; these transfection conditions not only act as controls for the various fluorescence measurements, but are also involved in the calculations associated with FRET, FRET efficiency and ultimately, the calculated distance between donor and acceptor. In addition to these 22 different transfection conditions, each of these had three different analyses performed, corresponding to the donor, acceptor and FRET channel wavelengths. The donor channel used CFP specific filters which excite samples at 430/10 nm and detect emission at 480/20 nm. The acceptor channel used YFP specific filters which excite samples at 485/14 nm and detect emission at 530/25 nm. And finally, the FRET channel used a combination of the CFP and YFP filters, to excite samples at 430/10 nm, and detect emission at 530/25 nm. These raw measurements of fluorescence were performed in duplicate, according to the Berthold method, background corrected against untransfected cells, and finally combined as an average of three independent experiments for each condition.

Once these average fluorescence values were obtained for each transfection condition, they were used in the following equations which calculate FRET efficiency (*E%*), and is detailed in section 2.2.3 and equations 2.2.3.2 to 2.2.3.5. Briefly, *d* calculates the fluorescence intensity

ratio for donor (CFP) bleed through, and *a* calculates the equivalent for acceptor (YFP) bleed through. These bleed through ratios are then used to calculate F_c which is the detected FRET signal, corrected for both donor and acceptor bleed through. Finally, the F_c value is then used along with the fluorophore quantum yields, Q_d and Q_a, to calculate the FRET efficiency between donor and acceptor, *E%*, which is the efficiency of energy transfer between CFP and YFP when in close proximity, which occurs due to the receptor-RAMP interaction. As shown in Table 3.4, the amylin 1 receptor (CTR-YFP and RAMP1-CFP) gave an average FRET efficiency of 88.32%, amylin 2 (CTR-YFP and RAMP2-CFP) gave 87.26% and amylin 3 (CTR-YFP and RAMP3-CFP) gave 86.91%. In addition, the CGRP receptor (CLR-YFP and RAMP1-CFP) gave an average FRET efficiency of 89.65%, adrenomedullin 1 (CLR-YFP and RAMP2-CFP) gave 86.74%, and adrenomedullin 2 (CLR-YFP and RAMP3-CFP) gave 87.78%. These results not only show that the tagged receptor and RAMP constructs were indeed well-expressed compared to background, but were also interacting with each other in the proximity needed to facilitate FRET events.

$$d = F_D / D_D$$

$$a = F_A / A_A$$

$$F_C = F_{DA} - (D_{DA} * d) - (A_{DA} * a)$$

$$E = 1 - [D_{DA} - (D_{DA} + F_c * \frac{Q_d}{Q_a})]$$

Equation 3.4.1. Calculating FRET Efficiency, *E*. These equations are repeated for ease of reading, please see section 2.2.3 and equations 2.2.3.2 to 2.2.3.5 for detail and explanation.

As FRET efficiency is relative to the inverse 6th power of distance (*r*) between donor and acceptor, this distance was then calculated using the equation below for each transfection condition, as an estimation of the average distance between the tagged receptor and RAMP. While the FRET efficiency values are obtained from the above calculation, the Förster distance, R_0, is 4.751 nm for the CFP-YFP pair, and thus rearrangement of this equation gives distance, *r*. Also shown in Table 3.4, the receptor and RAMP for the amylin 1 condition were estimated to be 3.391 nm in proximity, amylin 2 were 3.448 nm and amylin 3 were 3.466 nm. In addition, the receptor and RAMP for the CGRP condition were estimated to be 3.315 nm in proximity, adrenomedullin 1 were 3.523 nm and adrenomedullin 2 were 3.420 nm. These results therefore indicate that the FRET signals detected between the tagged receptors and RAMPs did not only occur, but did so on an appropriate scale of distance for these membrane proteins, approximately 3.4 nm on average.

$$E = \frac{1}{1 + (\frac{r}{R_0})^6}$$

$$r = R_0[\left(\frac{1}{E}\right) - 1]^{(\frac{1}{6})}$$

Equation 3.4.2. Calculating Donor-Acceptor Separation Distance, *r*. These equations are repeated for ease of reading, please see section 2.2.3 and equation 2.2.3.6 for detail and explanation.

Table 3.4. Detection of FRET as a Result of Receptor-RAMP Interactions. When the tagged constructs were co-transfected together into COS-7 cells, as described, CFP and YFP gave detectable FRET events, indicative of interactions between the calcitonin family of GPCRs and the RAMPs, on a scale of <10 nm. The efficiency of energy transfer during these FRET events were calculated (*E%*), giving an overall average of 87.78% which equates to an overall average distance (*r*) of 3.42 nm between CFP and YFP. Individual results are shown below, are averages of three independent experiments, and were first background corrected against untransfected cells.

Transfection Combination	Receptor Phenotype	FRET Efficiency (*E%*)	Distance *r* (nm)
CTR-YFP + RAMP1-CFP	AMY1	88.32	3.391
CTR-YFP + RAMP2-CFP	AMY2	87.26	3.448
CTR-YFP + RAMP3-CFP	AMY3	86.91	3.466
CLR-YFP + RAMP1-CFP	CGRP	89.65	3.315
CLR-YFP + RAMP2-CFP	AM1	86.74	3.523
CLR-YFP + RAMP3-CFP	AM2	87.78	3.420

Overall, the aim of creating a FRET-based assay to detect interactions between the calcitonin family of GPCRs, and the RAMPs was met with success, and the results obtained were of a proximity expected of these membrane proteins. This novel assay was therefore deemed ready for use in screening existing or potential drugs which may affect the receptor-RAMP interaction interface in an allosteric manner. Importantly, this assay was also confirmed to be appropriate for use in high-throughput approaches, as well as low-throughput methods such as confocal imaging. These results therefore suggest a promising FRET-based platform from which future experiments and projects may originate.

3.5 – *Drosophila melanogaster*: An Alternative Expression System for Mammalian Membrane Proteins?

The overall aim of this strategy was to express both CLR-YFP and RAMP1-CFP, making a fluorescently-tagged CGRP receptor, using *Drosophila melanogaster* as the expression system, and determine whether this receptor-RAMP complex could be extracted from the *Drosophila* intact, and retain applicability to the above FRET assay. As discussed in section 1.1.3, the rhabdomere membrane stacks of *Drosophila* photoreceptor cells offer ample surface area for targeted overexpression of membrane proteins, utilising the well-established GAL4-UAS system (Phelps and Brand, 1998; Duffy, 2002). If this were possible, this may prove to be an efficient and cost-effective stable expression line, overcoming the downsides of transient mammalian expression, and retaining strong applicability to high-throughput drug screening.

The GAL4-UAS system makes use of a yeast transcription factor (GAL4) which binds to an upstream activation sequence (UAS), and drives expression of target genes in a tissue specific manner (Busson and Pret, 2007). In order to achieve this, as shown in Figure 3.5a, two *Drosophila* strains are mated together: the driver strain which expresses GAL4 under a tissue-specific promoter or enhancer such as the glass multiple reporter (GMR), and the UAS strain in which the target gene has been cloned downstream of the GAL4 UAS. The resulting F1 offspring will therefore express the target gene in those cells which specifically express GAL4, and thus, a driver strain which expresses GAL4 in the *Drosophila* eye (GMR-GAL4) will restrict expression of the target membrane proteins to the photoreceptor cells. It should be noted here that firstly, the cloning and expression of the target gene makes these *Drosophila* transgenic, and secondly, this system has been used to effectively express a number of functional GPCRs alongside endogenous *Drosophila* opsins (Eroglu *et al.*, 2002; Panneels *et al.*, 2011).

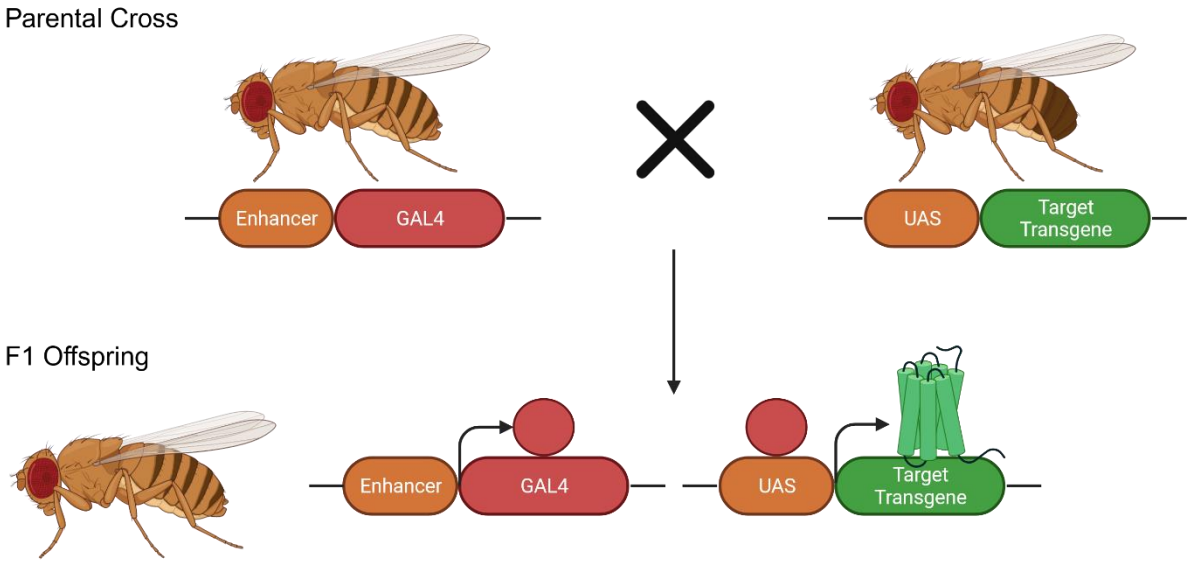

Figure 3.5a. The GAL4-UAS System for Targeted Transgene Expression in *Drosophila*. In order to express target transgenes in a tissue-specific manner, two different *Drosophila* strains are mated together in the parental cross. The driver strain expresses the yeast transcription factor GAL4, under the control of a tissue-specific enhancer or promoter, such as the glass multiple reporter (GMR). In addition, the UAS strain contains the target transgene, such as a GPCR or RAMP, downstream of the GAL4 upstream activating sequence (UAS). The resulting F1 offspring from this cross will therefore express GAL4 in a tissue-specific manner, driving expression of the target transgene within that tissue. As such, a driver strain which specifically expresses GAL4 within the eyes, such as GMR-GAL4, restricts heterologous expression of target transgenes to the photoreceptor cells. Created with Biorender.com.

Importantly, in order to make this GAL4-UAS system possible with transgenic *Drosophila*, the target gene must be cloned downstream of the UAS and inserted into one of the *Drosophila*'s chromosomes, thereby requiring a method to permanently incorporate the cloned target gene into the *Drosophila* genome. This can be achieved by the use of a UAST plasmid and manipulation of P elements, whereby the target gene is first cloned into the multiple cloning site of the plasmid, which itself is then incorporated into the genome via an endogenous germ-line-specific ΦC31 integrase (Bischof *et al.*, 2007). P elements are *Drosophila* transposons, or 'jumping genes', able to excise and insert themselves into various locations within the genome, and can be hijacked for the purpose of transgenesis. The plasmid pUAST-attB, along with its multiple cloning site cassette, contains the 285 bp attB fragment which is complementary to a specific attP landing site located on the *Drosophila* chromosome. Quite simply, as shown in Figure 3.5b, ΦC31 integrase catalyses the recombination of these attB and attP sites, incorporating the entire plasmid into that particular chromosome, resulting in the hybrid sites attR and attL. These new sites are refractory to the integrase enzyme and thus this P element system allows insertion, but not excision, and thus the target gene or

construct within pUAST-attB's multiple cloning site is permanently integrated into the *Drosophila* genome (Bischof *et al.*, 2007). Consequently, the target gene's expression can now be driven by the binding of GAL4 to the UAS sequence, as discussed. It should also be noted that the presence of the loxP sites eliminates any interference from sequences preceding or succeeding the integrated plasmid. Overall, the expression of target membrane proteins within the GAL4-UAS system requires the cloning of target construct sequences into the pUAST-attB plasmid, as evidenced by sections 3.5.1 to 3.5.3.

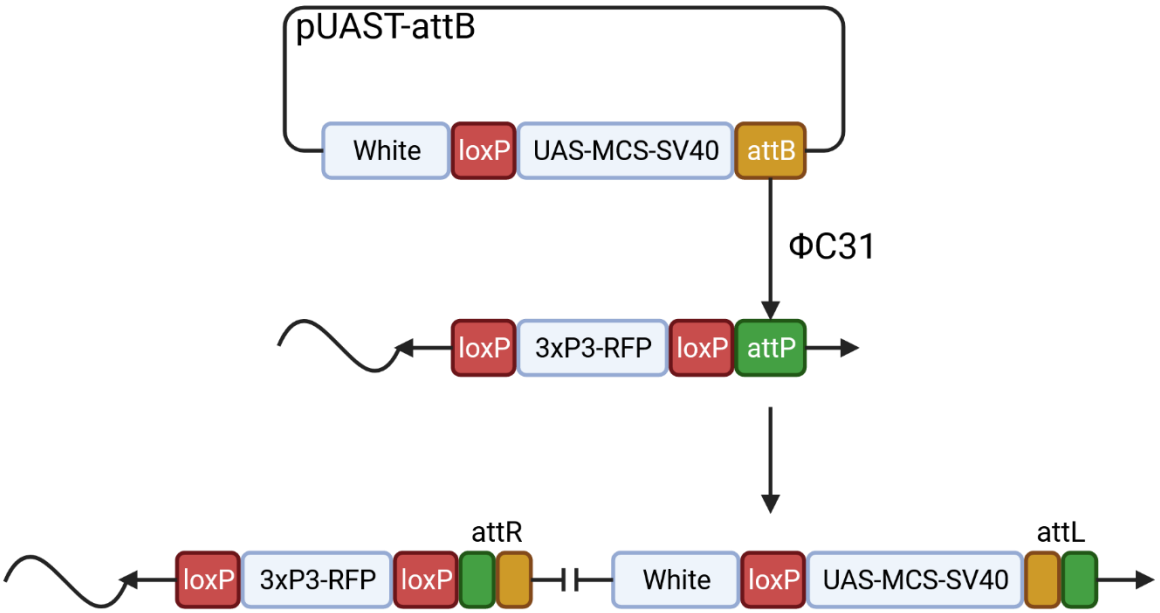

Figure 3.5b. ΦC31 Integrase-Mediated Integration of pUAST-attB into the *Drosophila* Genome. The attB fragment within the pUAST plasmid is able to recombine with a specific attP landing site on a *Drosophila* chromosome, with the use of an endogenous ΦC31 integrase, creating the hybrid sites attR and attL. The target gene or construct within the multiple cloning site (MCS) of the plasmid is thereby permanently integrated into the *Drosophila* genome, and can be driven to expression by the GAL4 transcription factor as shown in Figure 3.5a. The presence of loxP sites eliminates any interference by sequences either side of the integrated plasmid. Created with Biorender.com.

3.5.1 – Cloning Strategy for *Drosophila* Expression

In order to achieve the objective of this project, and express a fluorescently-tagged GPCR-RAMP system within the photoreceptor cells of *Drosophila melanogaster*, the DNA constructs designed in section 2.1 were modified in terms of their restriction sites, ensuring they would be complementary with this new plasmid, pUAST-attB. A new set of sense and anti-sense oligonucleotide primers were designed to achieve this strategy, listed in section 2.3.2. These new constructs were created by PCR, adding the necessary restriction sites, and enabling subsequent sub-cloning within the *Drosophila* vector, pUAST-attB, as illustrated by Figure

3.5.1. After considering the calcitonin family of GPCRs and the three RAMPs for inclusion in this trial experiment, the decision was made to focus solely upon expression of the CLR-RAMP1, or CGRP receptor, to provide proof of concept. As such, the RAMP1-CFP construct would be prepared by PCR and cloned into pUAST-attB, as evidenced by sections 3.5.2 and 3.5.3. In addition, due to the previous difficulties of cloning the CLR-YFP construct in section 3.2.3, its synthesis within pUAST-attB was again outsourced so as not to waste further time. Finally then, these two constructs, after confirmation by Sanger sequencing, would be ready for microinjection into *Drosophila* larvae, incorporation into their genome, and eventual expression trials.

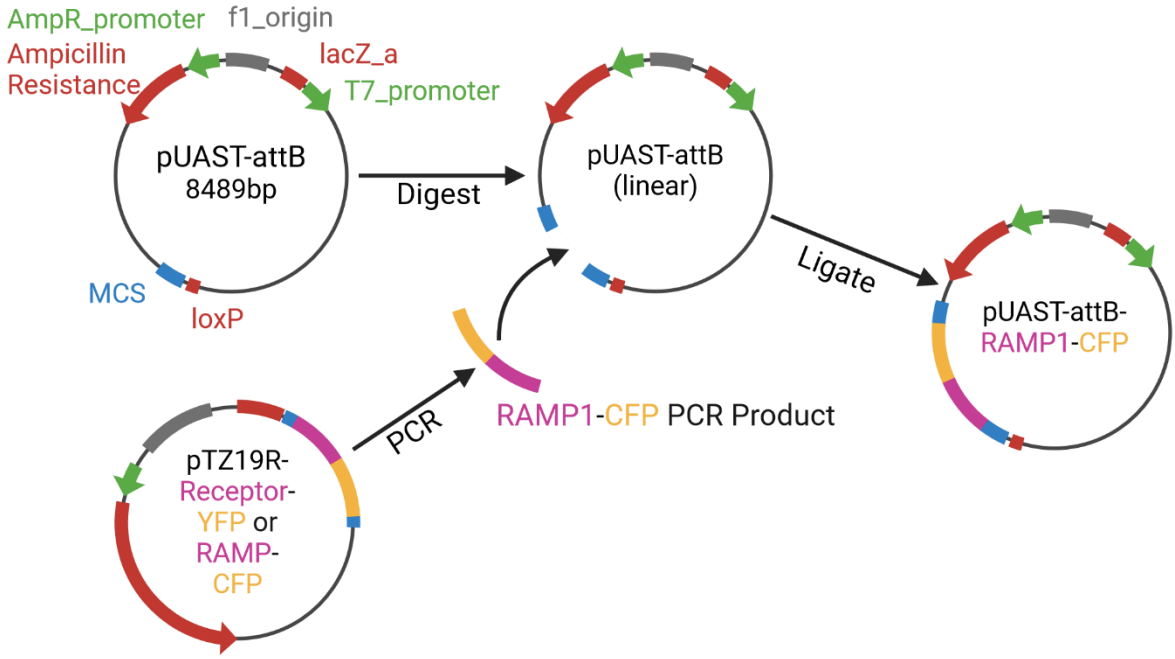

Figure 3.5.1. Cloning Strategy to Insert RAMP1-CFP into the *Drosophila* Expression Vector, pUAST-attB. The RAMP1-CFP construct cloned within pTZ19R would have its restriction sites modified by PCR, and subsequently ligated into the *Drosophila* expression vector pUAST-attB. Created with BioRender.com.

3.5.2 – Expansion of RAMP1-CFP by PCR

Taking the optimised PCR conditions initially performed in section 3.2.2 as a starting point, PCR conditions were then further optimised for expansion of the entire RAMP1-CFP construct. After many attempts, a pure PCR product was obtained as shown in Figure 3.5.2; the RAMP1-CFP construct (lanes 1-9) gave a band of approximately 1.2 kb, as expected, and was therefore selected for cloning attempts. This PCR step was ultimately required in order to modify the construct's restriction sites to be compatible with the pUAST-attB vector.

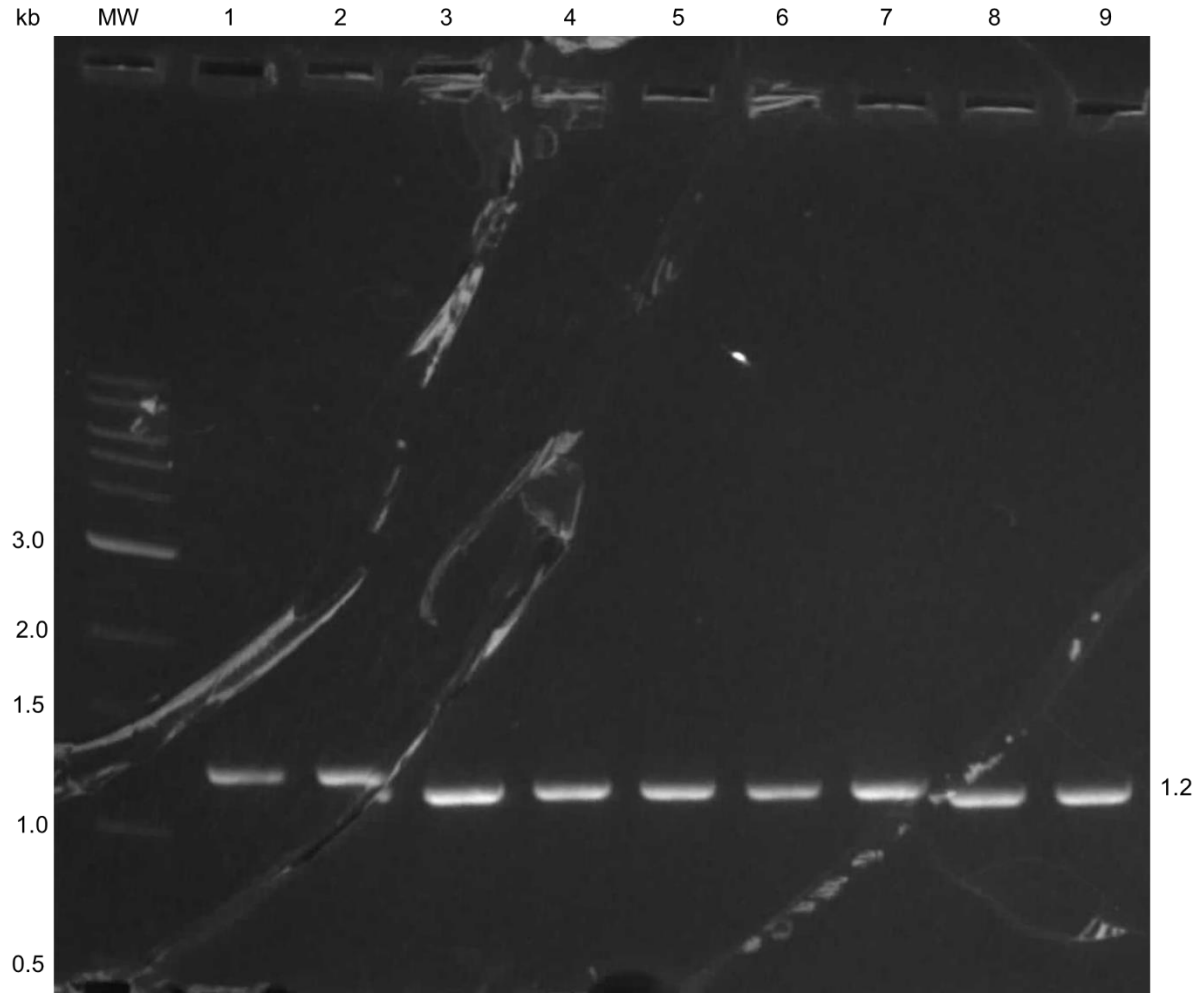

Figure 3.5.2. RAMP1-CFP PCR Products. When amplified using KOD polymerase and oligonucleotide primers listed in section 2.3.2, the RAMP1-CFP construct (lanes 1-9) resulted in an approximate band at 1.2 kb. MW is the NEB 1kb DNA ladder.

3.5.3 – Cloning RAMP1-CFP Into pUAST-attB

Prior to cloning attempts, the RAMP1-CFP PCR product was first purified through a PCR cleanup kit, digested with the appropriate restriction enzymes, run on an agarose gel, and purified a second time with a gel extraction kit. In tandem, the vector pUAST-attB was also digested with the corresponding restriction enzymes, run on an agarose gel, and purified with the gel extraction kit. Each of the purified PCR product and linearised plasmid were then combined into several ligation reactions with T7 DNA ligase, and subsequently transformed into DH5α *E. coli* cells. These were then grown overnight on agar plates at 37°C, and colonies were selected to be screened for successful clones by diagnostic digest and Sanger sequencing.

As shown in Figure 3.5.3, the RAMP1-CFP construct was successfully digested out of the vector pUAST-attB, resulting in two bands being present within each lane. RAMP1-CFP (lanes

3-6) gave a band of 1.2 kb, and the linearised pUAST-attB plasmid gave a band of 8.5 kb with each digest (lanes 3-6). Note that the upper band in lanes 2, 5 and 6 presented as a double-band which is explained by an incomplete digestion of the plasmid. Overall, these results mean the RAMP1-CFP construct fragment had been successfully cloned into the pUAST-attB plasmid, the sequence of which was then confirmed to be as expected by Sanger sequencing. This particular clone was then expanded to 0.5 mg/ml of recombinant plasmid DNA, in preparation for the incorporation into the *Drosophila* genome, along with the outsourced CLR-YFP construct.

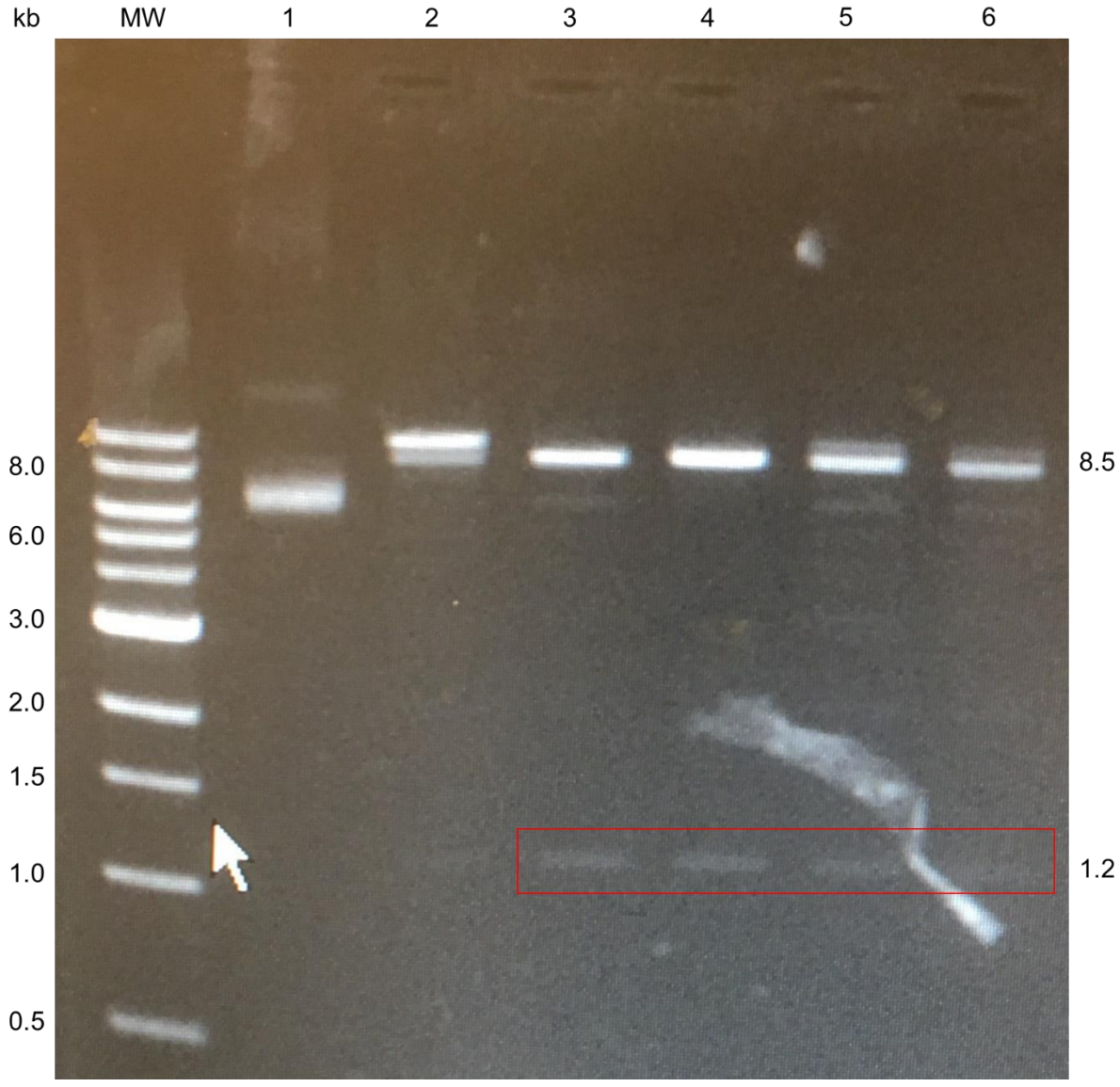

Figure 3.5.3. Restriction Digestion of the RAMP1-CFP Construct Assembled in pUAST-attB. Following its assembly in pUAST-attB, RAMP1-CFP (lanes 3-6) resulted in an approximate band at 1.2 kb when digested with the appropriate restriction enzymes listed in section 2.3.2. The linearised host vector pUAST-attB resulted in a band at 8.5 kb. MW is the NEB 1kb DNA ladder. The red box demarcates the lanes relevant to this experiment; lanes 1 and 2 are unsuccessful cloning attempts.

3.5.4 – From Transgenesis to Expression

Reiterating the aim of this project, to express the fluorescently-tagged CGRP receptor, both the CLR-YFP and RAMP1-CFP constructs must be incorporated into the same *Drosophila* strain, and be compatible with the transgenesis and GAL4-UAS systems described in section 3.5. Firstly, in terms of the P element-mediated transgenesis protocol, and combination of the attB and attP sites, the *Drosophila* genome possesses numerous P element landing sites across four different chromosomes. For simplicity, both constructs would be incorporated into two different strains, on two different chromosomes, which could then be crossed together at a later stage to produce a strain with both. The RAMP1-CFP construct within pUAST-attB would be incorporated into the attP40 site on chromosome 2, position 25C6, with a *yellow+* marker, and the CLR-YFP construct within pUAST-attB would be incorporated into the attP2 site on chromosome 3, position 68A4, also with a *yellow+* marker. In both cases, the endogenous ΦC31 integrase, needed for the recombination of the attB and attP sites, would be found on the X chromosome. Overall, this process of transgenesis was outsourced to a specialist *Drosophila* lab, the University of Cambridge Department of Genetics Fly Facility, whereby the constructs within pUAST-attB would be microinjected into *Drosophila* embryos along with a ΦC31 helper plasmid. Finally, separate tubes of these microinjected *Drosophila* larvae would be received, a proportion of which would have either CLR-YFP or RAMP1-CFP successfully incorporated into their genome.

Once this transgenesis protocol had been completed, and the resulting *Drosophila* received, they were initially cultured to adulthood, and subsequently used in a series of mating crosses, as shown in Figure 3.5.4, to achieve certain outcomes. Each of these crosses would take 10 days to complete, and thus the crossing schedule to attempt expression would be on a scale of approximately 8 to 10 weeks in total. Briefly, step 1 involves separately mating adult males of the transgenic strains received from the Fly Facility with white-eyed female virgins (WFV) of a $w^{+/-}$ host strain, moving the transgene construct from the germline to the somatic cell line, and reintroducing a red-eyed phenotype. The adult males of these red-eyed progeny were then mated with the WFV strain a second time, shown in step 2, to select for successful crosses, and boost stocks of them. In step 3, the transgenic adult males of these crosses were mated with a balancer strain called Curly derivative of Oster, or CyO (Miller *et al.*, 2018); this introduces an irradiated X chromosome to create a stable stock of *Drosophila* containing each transgene construct, both CLR-YFP/CyO and RAMP1-CFP/CyO. Their resulting phenotype is now red-eyed and curly wings.

At this point, these stable strains could be mated with GMR-GAL4 to individually drive expression of either CLR-YFP or RAMP1-CFP, however, in order to express both, the two

stable strains are first mated together as shown in step 4. This results in progeny which possess both transgene constructs, on chromosomes 2 and 3 (CLR-YFP/RAMP1-CFP/CyO), and can then be crossed with the GMR-GAL4 driver strain, as in step 5. Ultimately, the final progeny shown in step 6 would be expressing both target constructs (the fluorescently-tagged CGRP receptor) in the photoreceptor cells of the eyes, and can then be extracted for subsequent experiments. This final cross is characterised by a red-eyed and straight wing phenotype. Although this project could only be partially progressed towards expression in a RAMP1-CFP/GMR-GAL4 strain, with preliminary results shown in section 3.5.5, this offers a promising glimpse in the use of *Drosophila* as an alternate membrane protein expression system for biophysical and biochemical assay.

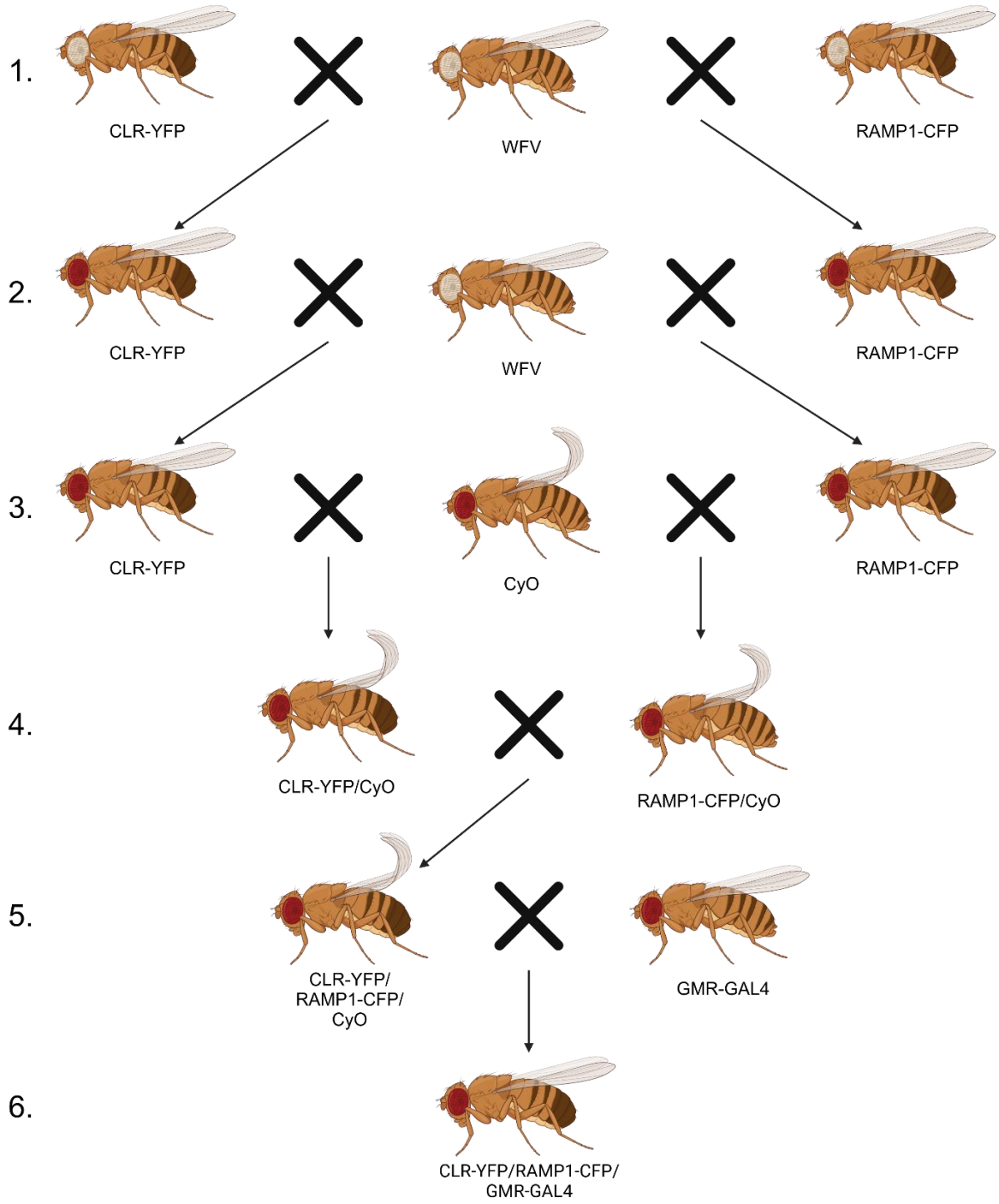

Figure 3.5.4. Crossing Schedule for Transgene Expression. Steps 1 to 6 show the process of creating stable *Drosophila* strains which will express the target constructs, CLR-YFP and RAMP1-CFP. The first two steps cross the transgenic strains with white-eyed female virgins (WFV) of a host $w^{+/-}$ strain to move the transgene, from the germline to the somatic cell line. This is repeated to select for successfully crossed progeny, and also reintroduces the red-eyed phenotype. Subsequent crossing with the CyO balancer strain makes this strain stable, rather than transient, and introduces a curly wing phenotype. Finally, crossing with the GMR-GAL4 strain drives expression of the target transgene constructs within the photoreceptor cells of the eyes, and is recognised by a red-eyed, straight wing phenotype. Created with Biorender.com.

3.5.5 – Preliminary Expression of RAMP1-CFP in *Drosophila* Photoreceptor Cells

Once the RAMP1-CFP/GMR-GAL4 strain had been obtained from the crossing schedule detailed in section 3.5.4, the expression of RAMP1-CFP should be under control of the GAL4-UAS system within the photoreceptor cells, and thus the next steps involved the detection or observation of this expression. As mentioned, the culture of *Drosophila* is easily scalable, and the number of flies to be used depends on the particular experiments of choice. Prior to more labour-intensive and time consuming techniques such as genome sequencing, SDS-PAGE, and Western blotting, two quicker and readily available methods were first performed to detect the expression of RAMP1-CFP.

The first was an extraction of RAMP1-CFP from *Drosophila* eyes, measuring the fluorescence of the sample given at the CFP emission wavelength (excitation 405 nm, emission 485 nm). The heads of 20 RAMP1-CFP/GMR-GAL4 *Drosophila* were collected and homogenised in a 25 mM Tris-Cl, 1% SDS lysis buffer, heated at 60°C for five minutes, centrifuged at 300 rcf, and the subsequent supernatant was harvested. This same process was also performed on the heads of 20 RAMP1-CFP/CyO *Drosophila*, as a non-expressing control; while this particular strain does contain the RAMP1-CFP transgene construct, it is not driven to express due to the lack of the GAL4 transcription factor. As shown in Figure 3.5.5.1, the non-expressing negative control, RAMP1-CFP/CyO gave a mean value for CFP fluorescence of 3389 relative light units (RLU) +/-936 standard error of the mean (SEM). Meanwhile, the expressing RAMP1-CFP/GMR-GAL4 strain gave a significantly greater mean value of 17040 RLU +/-1436 SEM; an unpaired two-tailed t-test between the two conditions gave a P value of 0.0013. Of course, the most specific method to detect membrane protein expression would be Western blotting, however, these preliminary fluorescence results suggest that RAMP1-CFP was indeed being driven to expression in the photoreceptor cells, within the RAMP1-CFP/GMR-GAL4 strain, as expected.

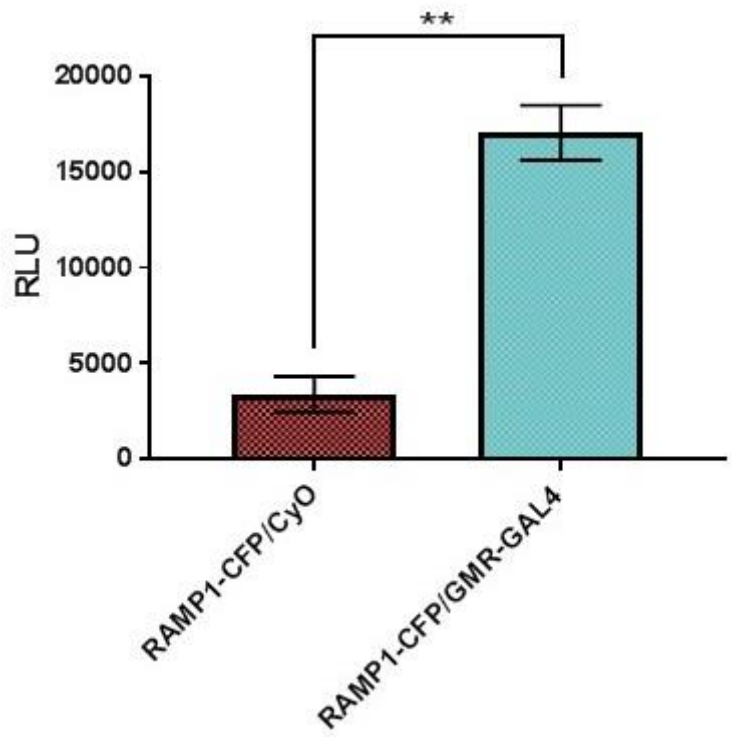

Figure 3.5.5.1. RAMP1-CFP Is Driven to Expression by GAL4. 20 heads of both the non-expressing RAMP1-CFP/CyO and the expressing RAMP1-CFP/GMR-GAL4 strains were harvested, homogenised and had their fluorescence measured, using the wavelengths appropriate for CFP (excitation 405 nm, emission 485 nm). The raw data was first background-corrected to a blank value of lysis buffer only. The non-expressing RAMP1-CFP/CyO strain gave a mean fluorescence value of 3389 RLU +/- 936 SEM, while the expressing RAMP1-CFP/GMR-GAL4 strain gave a mean of 17040 RLU +/- 1436 SEM. This was found to be a statistically significant difference with an unpaired two-tailed t-test giving a P value of 0.0013. Data is an average of n=3 with both biological and technical replicates.

The second available, and relatively quick, method to potentially detect CFP fluorescence within the eyes of the expressing strain was by fluorescent confocal microscopy. One single head of the RAMP1-CFP/GMR-GAL4 strain was dissected away from the body, mounted onto a slide and subsequently visualised with a Leica TCS SP8 confocal microscope. Following a wide-field baseline and checks for background auto-fluorescence, the confocal was calibrated to detect CFP fluorescence (excitation 405 nm, emission 485 nm) and was subsequently used at 10X, 40X and 60X magnification. As shown in Figure 3.5.5.2, image A shows the entire *Drosophila* head at 10X magnification, with CFP fluorescence detected within the eyes, as expected; scale bar represents 300 µm. As *Drosophila* eyes are an example of compound structures, being comprised of many cone-like ommatidia, much of the RAMP1-CFP expression in the photoreceptor cells would not be found at the apical surface of the eye. Instead, the photoreceptor cells are found underneath the apical cornea, deeper within the 3D structure of the ommatidia. As such, and shown in image B, a Z stack was created which

combines many optical slices of the sample, along the depth of the Z axis, detecting strong CFP fluorescence throughout the 3D structure of the ommatidia; this was performed at 40X magnification and the scale bar represents 50 µm. Finally, image C was taken at 60X magnification, and focusses on the junction between four ommatidia, showing a clear distinction between CFP fluorescence (cyan) and cellular auto-fluorescence (red) indicating a separate and therefore genuine presence of CFP; scale bar represents 2.5 µm.

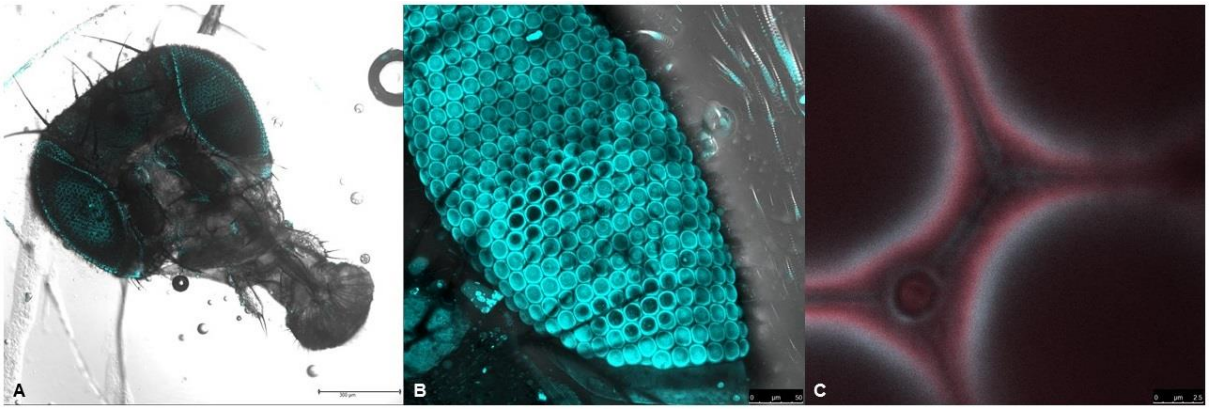

Figure 3.5.5.2. RAMP1-CFP Fluorescence Visualised in the Eyes of RAMP1-CFP/GMR-GAL4 *Drosophila*. A single head of the expressing strain, RAMP1-CFP/GMR-GAL4, was prepared for, and visualised under fluorescent confocal microscopy. The confocal microscope was calibrated for CFP fluorescence (excitation 405 nm and emission 485 nm) and assessed for any effects of cellular auto-fluorescence, ensuring any CFP fluorescence detected was genuine. Image A shows the entire *Drosophila* head with CFP fluorescence detected within the eyes; 10X, scale bar is 300 µm. Image B focusses upon one eye to create a Z stack, detecting strong CFP fluorescence throughout the compound ommatidia; 40X, scale bar is 50 µm. Image C focusses on four ommatidia, showing a clear distinction between CFP fluorescence (cyan) and auto-fluorescence (red); 60X, scale bar is 2.5 µm.

Overall, these two sets of fluorescence results give a strong indication that RAMP1-CFP was indeed being driven to expression by GAL4, suggesting that the expression of fluorescently-tagged membrane proteins within the photoreceptors of *Drosophila melanogaster* can indeed be useful for studying protein-protein interactions.

3.6 – Discussion

To reiterate, the modulation of the calcitonin family of GPCRs by the RAMP accessory proteins has been a well-characterised focal point over the last thirty years, but there remains a significant lack of effective, specific and comprehensive treatment options for many of the conditions involving these membrane proteins. Typically, therapeutic drugs for GPCRs largely target their orthosteric binding sites, but given that only 1% of drugs successfully pass through

clinical trials, alternative approaches must be discovered and adopted to strive towards greater chances of success (Hauser *et al.*, 2017). Given that the RAMPs themselves affect the pharmacology of GPCRs via allosteric modulation, an alternative approach to drugging GPCRs might involve targeting the receptor-RAMP binding interface itself, allosterically. Such drugs may act as inhibitors, by weakening or preventing receptor-RAMP interactions, but conversely may instead act as a bivalent 'adhesive' to strengthen or promote receptor-RAMP interactions (Kotliar *et al.*, 2023). In any case, there are also a lack of appropriate drug screening assays which would be able to detect these changes in protein-protein interactions; this body of work therefore sought to address this issue.

In selecting the platform able to detect protein-protein interactions, several techniques are highlighted as viable choices, however, their potential can often be opposed by their limitations. *In vitro* and *in vivo* options such as yeast-2-hybrid, co-immunoprecipitation, tandem affinity purification and crosslinking, for example, can indeed detect protein-protein interactions. However, aside from time and cost, their limitations often include 'noisy' datasets with unwanted background signals, a significant number of false positives, and a lack of direct applicability to drug screening approaches (Rao *et al.*, 2014). In addition, several *in silico* methods are also able to detect protein-protein interactions through a series of sequence-based or structure-based approaches; while these computational techniques have proven invaluable in further understanding the structure/function relationship of proteins, their applicability to drug screening in cellular or *in vivo* models may be limited or inappropriate, in some cases (Zhang, 2009).

In terms of the detection or analyses of specific receptor-RAMP interactions (summarised concisely in (Kotliar *et al.*, 2023)), many of the published interactions were initially detected *in vitro,* via immunofluorescent co-localisation at the cell surface (Christopoulos *et al.*, 2003; Bouschet, Martin and Henley, 2005). These experiments typically involved fluorescently tagging various GPCRs and the three RAMPs, with visualisation via fluorescent confocal microscopy, usually alongside a form of functional characterisation. More recent approaches to receptor-RAMP interactions seem to favour alternative fluorescent approaches including fluorescent activated cell sorting (FACS) and bioluminescence resonance energy transfer (BRET) (Bailey *et al.*, 2019; Mackie *et al.*, 2019; Shao *et al.*, 2022). In any case, these provide a prior basis of fluorescently tagging GPCRs and RAMPs to detect their interactions, offering justification for the methods selected in this body of work.

Given the importance of the structure/function relationship of membrane proteins, the conformational changes exhibited by GPCRs upon their activation, and the subtle ways in which this activation can be modulated allosterically by the RAMPs, drugs or otherwise, the

sensitivity of the techniques able to detect protein-protein interactions must be considered. Consequently, as the average diameter of a receptor-RAMP complex might be approximated to 5 nm (Gurevich and Gurevich, 2008; Padmanabha Das *et al.*, 2020), it was necessary to select a technique appropriate for drug screening, and also possesses the sensitivity necessary for such a scale. Despite the involvement of the confocal microscope in the initial detection of receptor-RAMP interactions, the typical resolution of a confocal lies in the range of 180-250 nm, and thus co-localised fluorescence indicates that these membrane proteins are within approximately 200 nm of each other; even the more specialised total internal reflection fluorescence (TIRF) microscopy only reaches a resolution of 60 to 100 nm (St. Croix, Shand and Watkins, 2005; Martin-Fernandez, Tynan and Webb, 2013). As a result of these limitations, FRET is commonly used to detect putative interactions between molecules on a scale of less than 10 nm (Shashkova and Leake, 2017); not only is this more appropriate for receptor-RAMP interactions, but can be performed with *in vitro* cell line models, is compatible with both fluorescent confocal microscopy and high-throughput fluorescence spectrophotometry, takes background fluorescence into account, and is very sensitive to subtle changes in distance between the interacting molecules. For all of these reasons, FRET was therefore deemed to be the best available option for this body of work.

Thus, after the initial *in silico* construct design, CFP-YFP became the selected FRET pair due to their popularity in overcoming the limitations of the BFP-GFP pair, their respectably high quantum yield, and previous use in the detection of GPCR oligomerisation (Haase, Kamann and Brüx, 2011; Bajar *et al.*, 2016). Although this body of work did not encounter them, the CFP-YFP pair does of course possess its own limitations, and therefore many different FRET pairs are available, depending on the context of their use. The fluorescently-tagged constructs were then created by optimised PCR and molecular cloning techniques, as described, before transfection and expression in mammalian cells, specifically the African green monkey kidney fibroblast-like COS-7 cells. These were selected not only because it is a commonly used immortalised cell line, but has been commonly utilised in the field of the calcitonin receptor family as these cells seemingly do not express significant levels of endogenous RAMPs (Bailey and Hay, 2006). In comparison, HEK293 cells and SK-NM-C neuroblastoma cells have been found to express RAMPs endogenously, for example (Aiyar *et al.*, 2002; Choksi *et al.*, 2002); by eliminating these endogenous proteins, their potential interactions with transfected constructs are minimised.

In any case, the receptor and RAMP constructs were transfected into COS-7 cells, and subsequently assessed for their ability to induce the production of cAMP in response to the agonist CGRP, as an indication of pharmacological functionality. Importantly, also ensuring that fluorescently-tagging the constructs did not significantly affect their signalling capabilities,

and thus remain applicable to drug screening approaches. This was indeed found to be the case, with no significant differences in the cAMP responses of the untagged and tagged receptor-RAMP complexes in response to CGRP, and in addition, the pEC50 values of the cAMP responses were also found to be comparable to those reported in the literature (Hay *et al.*, 2018). While this was reassuring overall, it should be noted that cAMP production in response to stimulation is not only agonist-, receptor- and RAMP-dependent, but can also be very cell-line dependent; the reported pEC50 values for the CGRP receptor for example, range from 8.18 to 10.44, and were determined in COS-7, HEK293/S/T, and CHO-K1 cell lines, after stimulation with CGRP. Interestingly, while the cAMP assays in this body of work did not significantly differ between the untagged and tagged constructs, the fluorescently-tagged receptor-RAMP complexes, across both CTR and CLR receptors and the three RAMPs, all gave dose-responses with a slightly higher efficacy than their untagged counterparts. One possible explanation for this may be due to a stabilising effect of the interaction between CFP and YFP, as this FRET pair is characteristically able to form weak dimers between themselves (Bajar *et al.*, 2016). Lastly, while cAMP production in response to CGRP was the focus here, the biased agonism of the constructs could be probed with several more calcitonin family ligands and various alternative signalling pathways, to further strengthen their pharmacological characterisation (Zhang and Xie, 2012).

Finally, after the target constructs demonstrated this expected functionality, they were used to establish the high-throughput FRET assay itself. Based upon the ‘gold standard’ method developed by Berthold Technologies to detect GPCR oligomerisation (Haase, Kamann and Brüx, 2011), FRET events were indeed detected between the fluorescently-tagged receptors and RAMPs, confirming their interactions, and giving an overall average distance of 3.2 nm between each fluorophore. While these results were within the hypothesised range for receptor-RAMP interactions, this method of experimentally determining FRET efficiency between fluorophores is subject to variation by its very nature. As the fluorophores are tagged to the receptor or RAMP with a flexible serine/glycine linker sequence, they are afforded a degree of mobility at the intracellular face of the protein complex, contributing to a very small amount of variation between each complex at the point of assay measurement. This particular method of determining FRET efficiency uses an ensemble average of all cells in the sample, taking all variation into account, however, this may result in unwanted background noise compared to single-cell or single-molecule techniques (Leavesley and Rich, 2016).

As discussed, FRET measurements are certainly compatible with fluorescent confocal microscopy, enabling focus on one particular cell or a portion of its cell membrane, but generally lack high-throughput capability. Likewise, FRET is also compatible with specialist single-molecule fluorescence techniques, which eliminate the background noise of working in

cells, but then lose cellular responses to drugs which may be an essential component of the desired screening approach. These single-molecule techniques are perhaps more suited to solely detecting changes in FRET efficiency, rather than coupling these molecular changes to pharmacology. On reflection, an alternative technique which bridges the gap between these requirements and limitations might be found in flow cytometry with fluorescence-activated cell sorting (FACS), often abbreviated to FACS-FRET (Lim *et al.*, 2022). This particular technique has major advantages compared to others; it is relatively high-throughput, is able to measure FRET in all cellular compartments, is not limited to mammalian cells (FACS-FRET has been successfully demonstrated in *E. coli* and the *Trypanosoma brucei* parasite, for example), and is suitable for the study of both inter- and intra-molecular interactions. While FACS-FRET is also subject to the limitations often attributed to FRET itself, such as the overexpression of transfected constructs leading to subcellular localisation and cytotoxicity, it has nevertheless become a very promising medium between spectrophotometry and confocal microscopy. It may be that a future evolution of the FRET-based receptor-RAMP interaction assay developed in this body of work would be well suited to a FACS-FRET approach.

Ultimately, following the optimisation and establishment of the receptor-RAMP interaction assay through the experimental work outlined in this chapter, it unfortunately could not be utilised for its intended purpose, due to the COVID-19 pandemic. It was planned to use this FRET assay as a basis for the discovery of experimental conditions or compounds which promote or inhibit the interaction between the calcitonin family of receptors and the RAMPs, influencing subtle changes in the distance between them, as detectable via FRET efficiency. Such compounds would have likely included known orthosteric ligands of the calcitonin family, therapeutics in current clinical use, and potential allosteric modulators (Hay *et al.*, 2018; Kotliar *et al.*, 2023). One of the main possibilities in drug discovery for receptor-RAMP complexes involves targeting not their orthosteric sites, but their known interactions between transmembrane helices as allosteric sites (Kotliar *et al.*, 2023). Given that the RAMPs modulate the pharmacological profiles of GPCRs, discovering inhibitors or promotors of their interactions may yield fruitful, and indeed blocking the interaction may significantly reduce or even eliminate detectable FRET events altogether, given the scale of <10 nm.

One additional point of contention within the field relates to the precise dynamics and stoichiometry of receptor-RAMP complex formation (Hilairet *et al.*, 2001; Udawela, Hay and Sexton, 2004; Héroux *et al.*, 2007; Harikumar *et al.*, 2009; Kotliar *et al.*, 2023). While interactions in a 1:1 ratio are a known paradigm, there exists conflicting evidence to suggest that RAMPs and even select receptors may be able to homodimerise, the purpose of which is as of yet unknown. In addition to this, investigation into the permanence of receptor-RAMP interactions has also garnered conflicting evidence. In theory, the interaction should be

relatively weak, however, in practice this is not observed to be the case. Receptor-RAMP complexes, as with the CGRP receptor in particular, have consistently produced experimental evidence to suggest that the complex remains associated together, throughout its journey from the endoplasmic reticulum, to the cell surface membrane, during internalisation and recycling, and finally internalisation to endosomes and eventual degradation. This is certainly considered strange for such supposed weakly interacting proteins; some hypotheses argue there is a stabilising effect from the RAMP itself, or perhaps from receptor component protein (RCP) (Routledge *et al.*, 2020), while some suggest the techniques used thus far have not been able to detect changes in complex dynamics. For example, if the CGRP receptor complex appears to remain associated during its lifetime, it is not to say that the original proteins are those being detected at each stage. Perhaps the RAMPs are able to associate and dissociate with receptors under varying conditions, and indeed even out-compete each other for opportunities to form complexes (Bühlmann *et al.*, 1999; Hay and Pioszak, 2016). In any case, this FRET assay forms a basis from which the dynamics of receptor-RAMP complex formation can be further probed, particularly in terms of the RAMP competition hypothesis.

As discussed in section 1.1.3, the varied systems available for the expression of membrane proteins are each characterised by both their advantages and disadvantages (Wiseman *et al.*, 2020); the eyes of *Drosophila melanogaster* offer a promising and balanced system, being relatively inexpensive, scalable, and offering a cellular environment more similar to mammalian cells, than perhaps *E. coli* or yeast, for example. While several GPCRs have been successfully expressed in *Drosophila* thus far, a novel extension of this existing work could involve the heterologous co-expression and analyses of both a receptor and RAMP within the same strain. This was one of the intended goals of this thesis, but was significantly impacted by the COVID-19 pandemic restrictions, and thus ended prematurely. The preliminary results obtained suggest that RAMP1-CFP was successfully expressed in the eyes of the GMR-GAL4 driver strain, but require further testing with additional controls to fully confirm that this was indeed the case. In addition, though the CLR-YFP construct had been prepared and used to generate transgenic *Drosophila*, they ultimately could not be mated with the RAMP1-CFP strain to create a strain expressing the fluorescently-tagged CGRP receptor.

Aside from the expression of receptor-RAMP complexes in *Drosophila*, which is a novel concept in and of itself, this approach also has direct applications to the use of FRET in the discovery of compounds which may promote or inhibit protein-protein interactions, as discussed previously. It is hypothesised that the expressed protein complexes are likely compatible with extraction and solubilisation into SMALPs for example, from which their FRET events would be detectable after purification (Dirnberger *et al.*, 2023). As mammalian expression systems tend to be relatively expensive, difficult to increase scale, and give a lower

yield, the initial expression of receptor-RAMP complexes in *Drosophila*, followed by this extraction may prove useful in providing a better yield of membrane proteins expressed in mammalian-like conditions. Furthermore, it is also feasible for such proteins within SMALPs to be subsequently reconstituted into appropriately composed synthetic membranes such as proteoliposomes, if such experimental conditions are favoured (Goers *et al.*, 2018).

To conclude then, the structural and pharmacological characterisation of GPCRs and RAMPs has progressed well over the last few decades, but there still remains a strong need to discover or develop effective treatment options for the plethora of conditions they are involved in. There is therefore a current need to develop drug screening assays which not only include both GPCRs and RAMPs, but are able to detect their interactions, and indeed detect drug-induced changes in these interactions. This body of work has therefore established a novel platform for such an assay, which satisfies these requirements, is compatible with a variety of low- and high-throughput techniques, and will additionally facilitate further probing of the dynamics of receptor-RAMP complex formation. Furthermore, preliminary results indicated that the novel expression of RAMP1 in transgenic *Drosophila melanogaster* was a success, and the heterologous expression of receptor-RAMP complexes in the eyes of *Drosophila* are more than likely possible using this methodology. Although these experiments were impacted by COVID-19 pandemic restrictions, the work detailed in this chapter did reach an advanced stage with a promising scope for future development.

Chapter 4: Elucidating the Structure-Based Signalling of GPCRs with the Novel use of Geometric Morphometrics

4.1 – Introduction

G protein-coupled receptors (GPCRs) form the largest 'superfamily' of membrane proteins in the human genome, and mediate cellular responses to a diverse array of signalling molecules including hormones, neurotransmitters, ions, odorants, and light (Bockaert and Pin, 1999; Weis and Kobilka, 2018). GPCRs are therefore key regulators of virtually every physiological, and consequently pathophysiological, process, making them highly attractive drug targets; over 30% of currently marketed therapeutics modulate GPCR activity (Santos *et al.*, 2017; Congreve *et al.*, 2020). However, developing safe and effective drugs which selectively modulate specific GPCR signalling pathways remains to be a formidable challenge. Addressing this particular challenge necessitates a comprehensive understanding of GPCR structure, and how this ultimately relates to intracellular signalling.

Over the last few decades, the field of GPCR structural biology has advanced dramatically through pivotal breakthroughs in X-ray crystallography (Congreve *et al.*, 2020), nuclear magnetic resonance (NMR) spectroscopy (Tapaneeyakorn *et al.*, 2011), and cryo-electron microscopy (cryo-EM) (Danev *et al.*, 2021). These techniques have yielded high-resolution snapshots of GPCR architecture, providing molecular insight into their ligand recognition sites (Lee, Booe and Pioszak, 2015), dynamic regions (Hilger, Masureel and Kobilka, 2018), and receptor-protein interfaces (Liang, Khoshouei, Deganutti, *et al.*, 2018; Liang, Belousoff, Fletcher, *et al.*, 2020; Liang, Belousoff, Zhao, *et al.*, 2020) such as the GPCR-RAMP interactions explored in Chapter 3. However, each of these methods have inherent strengths and limitations in capturing the nuanced intricacies of GPCR signalling mechanisms; the major limitation being the inability of current methods to 'see' the complete picture, and how this underpins the concept of structure-based signalling. This complex task necessitates a multidisciplinary approach that integrates structural data across techniques, in addition to biophysical assays and computational analyses, to construct a dynamic and holistic understanding of receptor function.

Against the increasing library of structural information, the emerging field of geometric morphometrics (Mitteroecker and Gunz, 2009; Wiseman *et al.*, 2021) has offered a transformative approach to GPCR structural biology. By enabling quantitative shape analyses, this approach facilitates the dissection of subtle yet functionally relevant conformational changes that underpin receptor activation and signalling bias.

4.1.1 – The Structural Study of GPCRs

The advent of X-ray crystallography has allowed for new insights into GPCR structural biology, providing the first high-resolution snapshots of receptor architecture. Seminal studies on bovine rhodopsin offered the earliest glimpses into GPCR organisation, revealing the characteristic seven-transmembrane helical bundle which is connected by flexible loops and stabilised by disulfide bridges (Palczewski *et al.*, 2000). These foundational insights highlighted common structural motifs among Family A GPCRs, providing a template for homology modelling of related receptors, and the start of pragmatic structural analyses. However, the inherent challenges in crystallising dynamic membrane proteins while preserving functional conformations continued to hamper the structural elucidation of ligand-activated states. A major breakthrough came with the stabilisation of the adrenergic receptor sub-family of GPCRs, either using additional fusion proteins for crystallisation efforts (Cherezov *et al.*, 2007), or thermostabilising methods (Vaidehi, Grisshammer and Tate, 2016), for example, which enabled their structure determination in complex with the partial inverse agonist carazolol (PDB 2RH1; (Cherezov *et al.*, 2007)). This landmark achievement highlighted key molecular details of the ligand binding pocket and intracellular regions involved in G protein coupling to the adrenergic receptors. Further optimisation of the crystallisation constructs subsequently culminated in the first structures of a GPCR bound to a fully activating agonist, providing snapshots of active-like receptor conformations critical for elucidating activation mechanisms (PDB 3SN6; (Søren G.F. Rasmussen *et al.*, 2011)).

Parallel advances in NMR spectroscopy offered complementary dynamic perspectives on GPCRs, overcoming key limitations in size and stability. Early NMR studies focussed on soluble protein domains, providing insights into microswitch conformational changes that enable receptor activation (Hu and Jin, 2022; Yang, Liu and Wüthrich, 2022). Recent advances have enabled NMR analyses of full-length receptor samples in synthetic liposomes, the key findings of which include the characterisation of ligand-specific conformational states of the β2-adrenergic receptor, revealing distinct structural signatures that could underlie biased signalling (Prosser and Alonzi, 2023).

While crystallography laid a strong initial foundation, the emergence of single-particle cryo-EM has revolutionised GPCR structural biology by enabling receptor characterisation in phospholipid nanodiscs. Cryo-EM has unlocked the ability to capture GPCR conformational ensembles, transitions between states, and complexes with binding partners, overcoming the limitations of requiring crystal contacts (Gusach, García-Nafría and Tate, 2023). Seminal cryo-EM studies visualised both active and inactive state transitions of the β2-adrenergic receptor, identifying precise structural rearrangements between states (Yanan Zhang *et al.*, 2020).

High-resolution analyses have also revealed unexpected ligand binding and dimerization interfaces, with implications for structure-based drug design (Gusach, García-Nafría and Tate, 2023). Moreover, cryo-EM facilitated the capture of transient GPCR-G protein complexes, offering unprecedented views of receptor-mediated nucleotide exchange and G protein state transitions critical for signalling (Zhang *et al.*, 2021). Most recently, innovative drug discovery platforms leverage cryo-EM to structurally characterise receptors bound to novel ligands, facilitating structure-guided design. Studies have also captured distinct ligand-specified states that likely represent structural intermediates underlying biased signalling phenomena (Yang *et al.*, 2021).

4.1.2 – The Limitations of Structural Techniques

While foundational techniques for structure determination have advanced understanding of static GPCR architecture, each method carries inherent limitations. A significant challenge is the multifaceted conformational heterogeneity of GPCRs, which is difficult to capture through traditional structural snapshots. Crystallography necessitates the stabilisation of receptors in specific ligand-bound states that likely misrepresent dynamically sampled conformations in native membranes (García-Nafría and Tate, 2021). In addition, NMR spectroscopy struggles to resolve complete structures of large, membrane-embedded receptors, and although powerful, even cryo-EM approaches still face technical hurdles in resolving heterogeneous structural states in receptor samples.

Critically, these techniques characterise receptors removed from cellular environments, often utilising stabilising mutations that fail to recapitulate native conformational ensembles. Capturing physiologically relevant receptor dynamics therefore requires emerging integrated approaches, combining structural data with computational simulations and functional assays to construct dynamic models. However, connecting high-resolution architecture with functional mechanisms remains non-trivial which underscores the need for quantitative analytical frameworks to decode conformational transitions.

4.1.3 – Elucidating the Link Between Structural Conformation and Signalling Bias

GPCR signalling bias, also termed ligand-directed stimulus trafficking, or functional selectivity, refers to the ability of structurally distinct ligands to preferentially activate specific downstream signalling pathways through the same receptor (Wootten *et al.*, 2018; Gurevich and Gurevich, 2020). This phenomenon enables the fine-tuning of physiological responses, and holds therapeutic promise by facilitating the development of pathway-specific pharmacological agents. However, traditional structural biology techniques face inherent barriers in capturing this conformational heterogeneity. Here, the emerging field of geometric morphometrics offers

a transformative analytical paradigm. By enabling the quantitative mapping of ligand-specific shapes, geometric morphometrics can elucidate the conformational determinants and dynamic structural transitions underlying the biased modulation of signalling pathways.

Further exploration of the biased signalling phenomenon will not only further contribute to understanding the structural basis of GPCR function, but also has clear application to biomedical therapeutics. Traditional opioid painkillers such as morphine act upon the μ-opioid receptor, to effectively relieve pain, however, they also activate the β-arrestin pathway, leading to adverse effects like respiratory depression and constipation (Kelly, Conibear and Henderson, 2023). Biased μ-opioid agonists, such as oliceridine (TRV130) and PZM21, were discovered to preferentially activate G protein signalling over β-arrestin recruitment, demonstrating potent analgesia with reduced side effects in pre-clinical models, offering promise for safer pain management options (Che *et al.*, 2021). Likewise, the angiotensin II type 1 receptor (AT1R) is a key regulator of blood pressure, and a target for treating hypertension and heart failure; traditional AT1R antagonists (ARBs) block all downstream signalling (Fatima, Patel and Hussain, 2021). However, recent studies have identified biased agonists that selectively activate beneficial signalling pathways, while avoiding those which contribute to disease progression (Delaitre *et al.*, 2021; Nivedha, Lee and Vaidehi, 2023). For example, TRV027 and TRV120027 are β-arrestin-biased AT1R ligands which promote cardiomyocyte contractility, and improve cardiac function in animal models of heart failure. Structural analyses have revealed distinct conformations of AT1R stabilised by these biased ligands, offering insights into the mechanistic basis of selective signalling (Wingler *et al.*, 2020).

One particularly salient example to highlight is that of the β2-adrenergic receptor and its inverse agonist, carazolol, which has been shown to exhibit biased signalling properties. Carazolol stabilises an inactive conformation of the β2-adrenergic receptor, decreasing basal signalling activity for cAMP production, however, carazolol preferentially activates the ERK1/2 pathway over the canonical G_s-cAMP-PKA pathway (van der Westhuizen *et al.*, 2014). This bias is mediated by β-arrestin recruitment, and is distinct from the effects of other inverse agonists like ICL-118, 551. In addition to this, further structural analyses reveal that carazolol stabilises a unique conformation of the β2-adrenergic receptor, which differs from the inactive states stabilised by other inverse agonists.

4.1.4 - Aims and Objectives

The aim of this body work was therefore to establish and implement a robust method for the novel application of geometric morphometric analyses to resolved GPCR structures. This would involve the selection of salient GPCRs to analyse, the definition of landmarks, a standardised data collection protocol, the geometric morphometric analysis itself, with subsequent principal component analysis, followed by rigorous statistical testing, all of which were successfully achieved. In order to quantitatively decode the conformational heterogeneity of GPCRs, and further elucidate their structure-based signalling, several receptors were selected for analysis, complementing their existing structural analyses published in the literature. Receptors selected from each of the three major families in humans (A, B, and C) included the β2-adrenergic, adenosine A2A, secretin-like and calcium-sensing receptors, which additionally emphasised the multifaceted nature of geometric morphometrics. Their analyses included individual receptor sub-families, a combination of several sub-families, analyses of monomeric and homodimeric structures, and are detailed in the case study results presented in this chapter; these form the foundational geometric morphometric analyses of GPCR structures, focussed on mapping global structural rearrangements between inactive- and active-state receptor crystals. As these analyses were proven to be effective and consistent, additional uses and applications of geometric morphometrics for GPCR structures were further explored, as detailed in Chapter 5.

4.2 – Geometric Morphometric Analyses of Shape

The computational framework of geometric morphometrics facilitates sophisticated analyses of shape variation, and transformations between geometrical configurations of objects. This analytical approach captures the spatial relationships between anatomical landmarks across structures, offering a comprehensive perspective of shape differences (Mitteroecker and Gunz, 2009; Polly *et al.*, 2016). In addition, statistical analysis tools allow for the robust quantitative analyses of these spatial arrangements; Procrustese superimposition enables alignment of shapes while retaining geometric information by using least squares methods to superimpose landmarks, and principal component analysis of coordinate data then identifies major patterns within sampled shape space, highlighting subtle yet functionally relevant features (Mitteroecker and Gunz, 2009; Polly *et al.*, 2016). These tools facilitate detailed visualisation of the structural dynamics enabling functional transitions, and when applied to GPCRs, these techniques can map activated state conformational changes, while higher

dimensional analysis may capture complex energy landscapes populated during signalling events (Wiseman *et al.*, 2021).

4.2.1 – What is geometric morphometrics?

Geometric morphometrics, abbreviated to GM, is a mathematical method of comparing shapes, and is most often found to be used in the disciplines of anthropology, palaeontology and vertebrate zoology (Mitteroecker and Gunz, 2009). Its initial conception began in the late 1800s with the mathematician Sir Francis Galton, leading to the description of correlation coefficients and multivariate statistics, and eventually the geometric morphometric techniques invented in the 1980s (O'higgins, 2000; Adams, Rohlf and Slice, 2004; Slice, 2005). Put simply, GM uses a shape's coordinate data, or Cartesian landmarks, to capture its morphology and enable comparisons between shapes (Webster and Sheets, 2010). GM is so named as the target shape's geometry is preserved throughout analyses, allowing results of statistical tests to represent the shape's morphology; the size, position and orientation of each shape are standardised during analysis, meaning any observations are solely based upon differences in shape morphology (Mitteroecker and Gunz, 2009). Overall, there are several approaches to GM analyses of which the Procrustese method is most popular and well-understood when compared to the Euclidean or Fourier methods, for example (Dryden and Mardia, 2016).

4.2.2 – Geometric Morphometrics in Practice

Each step of the geometric morphometric process will now be detailed, however, GM has been comprehensively summarised by Polly *et al*. in 2016, and a useful book was additionally written specifically for biologists by Zelditch *et al*. in 2004 (Zelditch *et al.*, 2004; Polly *et al.*, 2016).

Firstly, appropriate data for GM analysis must be gathered such as Cartesian coordinates, sometimes termed 'landmarks', to define the target shape; importantly, each shape must possess the exact same landmarks, otherwise comparison between them is not possible (Mitteroecker and Gunz, 2009). For example, the three vertices of a triangle and four vertices of a square means these two shapes cannot be compared. As GM is commonly used in anthropology, palaeontology and vertebrate zoology, the shapes of animals and skeletons can be defined by landmarks for comparison; Figure 4.2.2.1 demonstrates this principle with a Eurasian perch (*Perca fluviatilis*), in which its body is reduced into common landmarks from which coordinate data can be obtained (Bartels *et al.*, 2012). Furthermore, GM is compatible with both two-dimensional and three-dimensional data, as with XY and XYZ coordinates, respectively (Zhang and Schepartz, 2021). But how can this be applied to the molecular level?

As GPCRs all possess seven transmembrane domains, the ends of each helix can be considered as common landmarks as they are the defining structural characteristic of these serpentine receptors, and remain relatively conserved between them. As such, the amino acid coordinates at each end of the transmembrane helices were therefore selected to analyse the morphology of both the extra- and intracellular faces of the GPCR transmembrane bundle, as detailed in methods section 2.4.

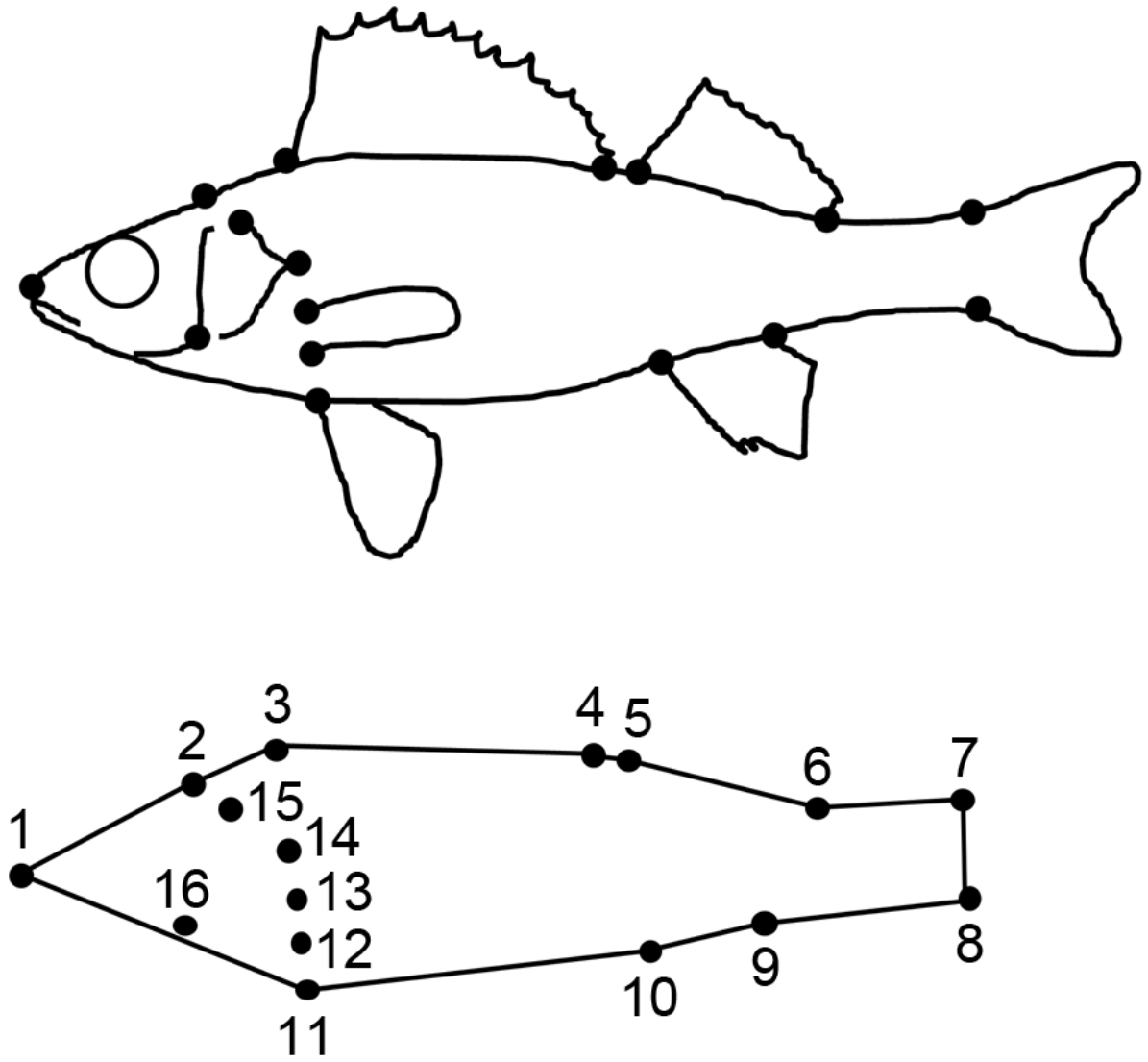

Figure 4.2.2.1. An Example of Selected Landmarks to Morphologically Analyse the Eurasian Perch. Importantly, all sixteen of these landmarks are shared commonly between fish of this species, enabling comparison by geometric morphometric techniques. Taken from Bartels *et al.*, 2012.

In any case, once the Cartesian coordinate data has been obtained from the appropriate shape landmarks, it must be standardised prior to analyses as raw coordinates are affected by size, position and orientation in space (Polly *et al.*, 2016). Most commonly, a generalised Procrustese superimposition is first performed as an orthogonal transformation to

superimpose, rescale, centre and rotate each set of landmarks, as demonstrated in Figure 4.2.2.2 (Mitteroecker and Gunz, 2009). This transformation results in Procrustese shape coordinates in which their average consensus gives a minimal sum of squared distances in order to estimate mean values needed for statistical analyses (Dryden and Mardia, 2016). Furthermore, the Euclidean distance between sets of Procrustese shape coordinates is called the Procrustese distance, and represents the dissimilarity between the sets of transformed landmark data. Essentially, the cumulative effects of size and orientation are eliminated during this process, and the Procrustese shape coordinates are therefore reflective of pure differences between shape morphology (Mitteroecker and Gunz, 2009).

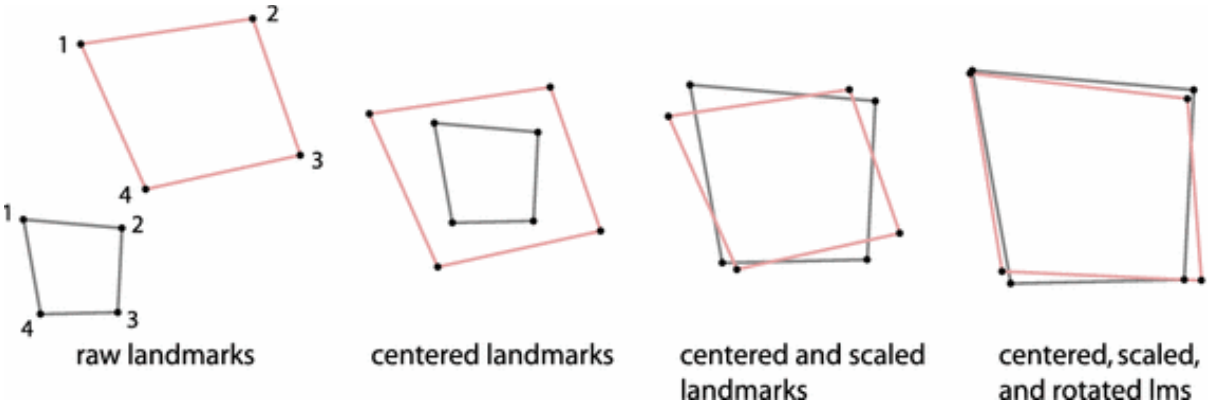

Figure 4.2.2.2. The Steps of Procrustese Superimposition. Raw landmark data are transformed into standardised Procrustese shape coordinates by superimposing, rescaling, centring and rotating each set of shapes. Taken from Mitteroecker and Gunz, 2009.

Finally, a covariance matrix can then be generated which displays the variance within datasets, and the covariance between pairs of datasets; variance measures the dispersion of data, and can be defined as the spread of data away from the mean of the dataset (Jolliffe, 2002). In other words, variance is the spread of individual shapes, or Procrustese residuals, from the calculated average of the Procrustese shape coordinates. Importantly, covariance matrices allow for the decorrelation of data, which optimises the basis from which complex data can be represented in a compact manner – one such method is a further technique called principal component analysis (Jolliffe, 2002).

4.2.3 – Principal Component Analysis

Principal component analysis, abbreviated to PCA, is a popular method of analysing very large datasets with many dimensions or variables, and reducing this data into a more compact, manageable package. By doing so, the data becomes much easier to interpret and enables the visualisation of multidimensional data, whilst importantly retaining the maximum amount of information from the original dataset (Mitteroecker and Gunz, 2009). This is achieved by a

linear or ‘whitening’ transformation, which transforms data into a new coordinate system with fewer dimensions, preserving the majority of the original variation within the dataset. Ultimately, PCA is a statistical method for dimensionality reduction, allowing complex data to be projected onto low-dimensional space, such as an XY graph, via its principal components.

The principal components themselves, when pertaining to coordinate data, are a sequence of *p* unit vectors, but can also be expressed as eigenvectors of the data’s covariance matrix, derived by eigendecomposition (Polly *et al.*, 2016). Eigenvectors are vectors which do not change direction after a linear transformation, and as such preserve this key information whilst reducing dimensionality, and are a key component to the PCA method. While the eigenvectors of a covariance matrix correspond to the principal components of the dataset, their eigenvalues represent the variance within the data (Dryden and Mardia, 2016). As such, the principal component scores are therefore projections of the Procrustese shape coordinates onto the low-dimensional space spanned by the eigenvectors, and can be represented in two- or three-dimensions such as an XY or XYZ scatter graph (Mitteroecker and Gunz, 2009). By doing so, the various sets of morphological data can be assessed for any differences, similarities or outliers of groups, based only upon shape morphology and not prior group affiliation. In the context of GPCRs then, any structural differences or similarities would be based upon their objective shape morphology, rather than affiliation to a particular family or receptor characteristic.

Finally, the principal component scores can be subjected to multivariate statistical analyses such as regression or analysis of variance, however, as shape variance rarely meets the requirements to assume normality, non-parametric tests based upon permutation are preferred (Polly *et al.*, 2016). Consequently, the permutational multivariate analysis of variance (PERMANOVA) and the analysis of similarities (ANOSIM) were selected as appropriate, non-parametric tests, and each assess the variation and dissimilarity of shapes within the dataset, respectively (Penrice and Deeming, 2020).

4.3 – Case Study Results

The foundational geometric morphometric analyses of GPCRs have focussed on mapping global structural rearrangements between inactive- and active-state receptor crystals; these studies utilised coordinate landmarks on transmembrane and intracellular loop regions that undergo key conformational changes during signalling. Statistical analyses reliably distinguish between functional states, while principal component analysis reveals shifts along specific structural modes that contribute to activation transitions (Wiseman *et al.*, 2021).

These quantitative architectures of activation uncover subtle conformational signatures within receptors that drive transitions into active states (Vogel *et al.*, 2008; Zhou *et al.*, 2019). Extending this framework across multiple receptor conformations, as well as different members of the GPCR superfamily, captures the breadth of diverse activated states adopted during signalling (Wiseman *et al.*, 2021). Comparative analysis of geometric shapes corresponding to different G protein-bound states will reveal divergent structural transitions that likely direct the engagement of specific downstream effector coupling (Mafi, Kim and Goddard, 2022). Identifying the impact of ligands in stabilising specific geometric phenotypes within this activation landscape will unravel key conformational determinants governing signalling bias (Wingler and Lefkowitz, 2020).

Detailed here are several case study examples, demonstrating these foundational analyses as a result of the novel application of geometric morphometrics to GPCR structures. Each of the receptors were selected to gain unique insights into the major GPCR families in humans, to build a more complete landscape of their structure/function relationship, to help in further understanding the diseases they are involved in as well as drug discovery efforts.

4.3.1 – Case Study 1: The Family A β2-Adrenergic Receptor

The β-adrenergic receptors elicit responses to natural catecholamines secreted by the adrenal glands, and are categorised into three sub-groups; the β2 receptors are predominantly found in airway smooth muscle, but are additionally found in other muscle tissues, epithelial cells, and various immune cells (Wallukat, 2002). The effects of the catecholamines are generally non-specific, however do display differing affinities to different adrenergic receptor sub-groups; adrenaline is the most effective agonist of the β2-adrenergic receptors, and noradrenaline less so, each leading to smooth muscle relaxation, for example (Abosamak and Shahin, 2024). As such, therapeutic agonists targeting the β2 receptors mainly treat the bronchospastic respiratory disorders linked to chronic obstructive pulmonary disease (COPD) and asthma (Barisione *et al.*, 2010). In terms of antagonistic options, the ‘beta blockers’ act generally on the β-adrenergic receptors to lower heart rate and reduce the physiological symptoms of anxiety, however, there are currently no β2-selective ‘blockers’ approved for therapeutic use (Cuesta *et al.*, 2019).

The β2-adrenergic receptor sub-group was selected to be the initial test case for geometric morphometric analysis for various reasons. As discussed, the β2-adrenergic receptor was one of the seminal structures which led to the elucidation of key molecular details of the ligand binding and G protein binding domains of the adrenergic receptors (Cherezov *et al.*, 2007); this was not only a landmark achievement in the structural biology of GPCRs, but has also significantly informed modelling efforts to date. Furthermore, the ligands of the β2-adrenergic

receptor have been demonstrated to show signalling bias, enabling an exploration of the structural basis of this phenomenon (Ippolito and Benovic, 2021). Moreover, the β2-adrenergic receptor sub-group also remains an effective therapeutic target, with clinically approved agonists and non-selective antagonists; there remains potential to develop safe and effective β2-selective ligands which are biased to preferentially activate beneficial pathways.

As such, at the time of sample selection, 36 human β2-adrenergic receptor structures were included in the dataset, with 23 described as inactive, 7 as active, and 6 as active bound to G_s. These structures were not filtered initially so as not to bias or limit the analysis, however, PDB 5JQH was not included due to heavily skewing the data; it was hypothesised that this was possibly due to 5JQH being allosterically nanobody stabilised (Staus *et al.*, 2016). Refer to Appendix B1 for a full list of structures, including PDB codes and reference to their publication. Overall, given the mixture of activation states in the dataset, and given that conformational changes are dominated by activation, it was hypothesised that activation state would be revealed to be the dominant source of variation in this dataset, represented by principal component 1.

PC Variance

The first two principal components accounted for 46% and 57%, of the variation in the extracellular and intracellular landmark coordinates, respectively, of the 36 β2-adrenergic receptor structures, as shown in Table 4.3.1.1. The first principal component (PC1) accounted for 29% (extracellular) and 42% (intracellular) of the variation.

Table 4.3.1.1. Principal Component Variance of the β2-Adrenergic Receptor Structures. The variance explained by the 14 principal components of 36 β2-adrenergic receptor structures, based on both the extra- and intracellular landmark coordinates. Variance is expressed as a percentage, alongside the total cumulative variance.

	Extracellular		Intracellular	
	%Variance	%Cumulative	%Variance	%Cumulative
PC1	28.487	28.487	42.071	42.071
PC2	17.554	46.041	14.602	56.673
PC3	13.333	59.373	12.142	68.815
PC4	9.039	68.412	10.423	79.238
PC5	7.02	75.433	5.202	84.44
PC6	6.473	81.906	3.611	88.051
PC7	5.22	87.125	3.211	91.263
PC8	4.366	91.491	2.688	93.95
PC9	3.267	94.758	2.039	95.99
PC10	2.091	96.85	1.609	97.599
PC11	1.835	98.685	1.432	99.031
PC12	0.554	99.239	0.473	99.504
PC13	0.512	99.751	0.333	99.836
PC14	0.249	100	0.164	100

PC1 and PC2 Scree Plots

Scree plots for the first two principal components, showing the location, direction and magnitude of variations are shown in Figures 4.3.1.1 and 4.3.1.2. For the extracellular landmarks, PC1 showed the greatest variation in the positions of transmembrane helices 1 and 6, with some variation in helices 3 and 4. PC2 showed the greatest variation in helices 1 and 5, and some variation in helices 3 and 7. For the intracellular landmarks, PC1 showed the greatest variation in helices 6 and 3, with some variation in helix 7. PC2 showed the greatest variation in helix 6, and some variation in helices 1, 2, 4, 5 and 7.

Overall, PC1 for both the extra- and intracellular landmarks highlighted variation in transmembrane helices 3 and 6, which likely corresponds to the conformational changes associated with the activation of GPCRs; as discussed in section 1.2.3, the upwards shift of TM3, and the outward rotation of TM6 form key components of the common mechanism of GPCR activation. Moreover, the variation displayed in both PC1 Scree plots for TM6 is directed away from the transmembrane helix bundle, which again reflects the outward rotation of TM6 during activation, and indicates principal component analysis determined this to be the greatest source of variation in the dataset. In addition to this, the inward movements of TM1, TM5 and TM7 were also demonstrated which further suggests the variation in this dataset is dominated by receptor activation, and therefore the conformational differences between structures resolved in active or inactive snapshots.

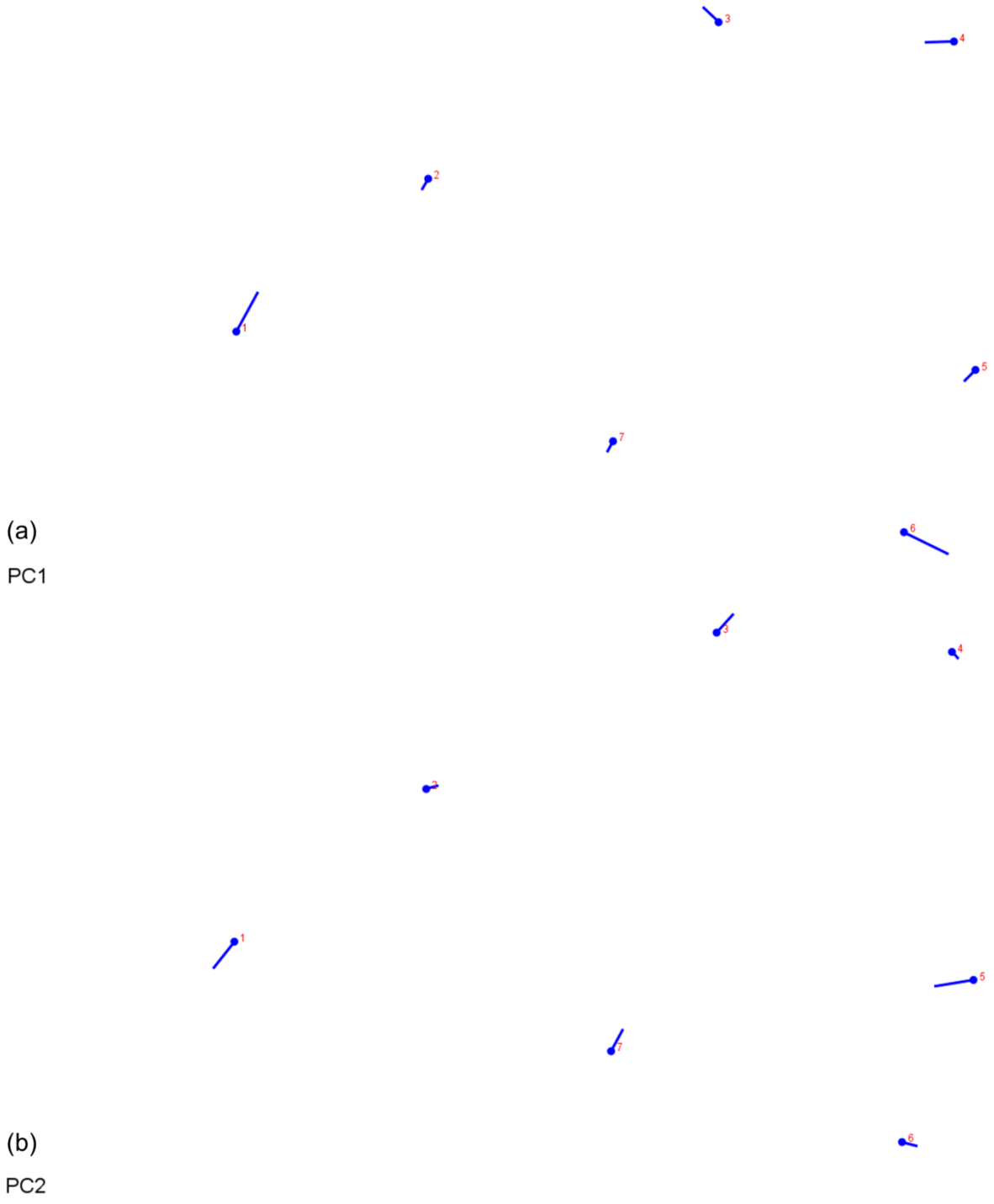

Figure 4.3.1.1. Scree Plots for both PC1 and PC2 of the Extracellular β2-Adrenergic Receptor Landmarks. Scree plots for PC1 (a) and PC2 (b) are shown for the 7 extracellular landmarks of the 36 β2-adrenergic receptor structures.

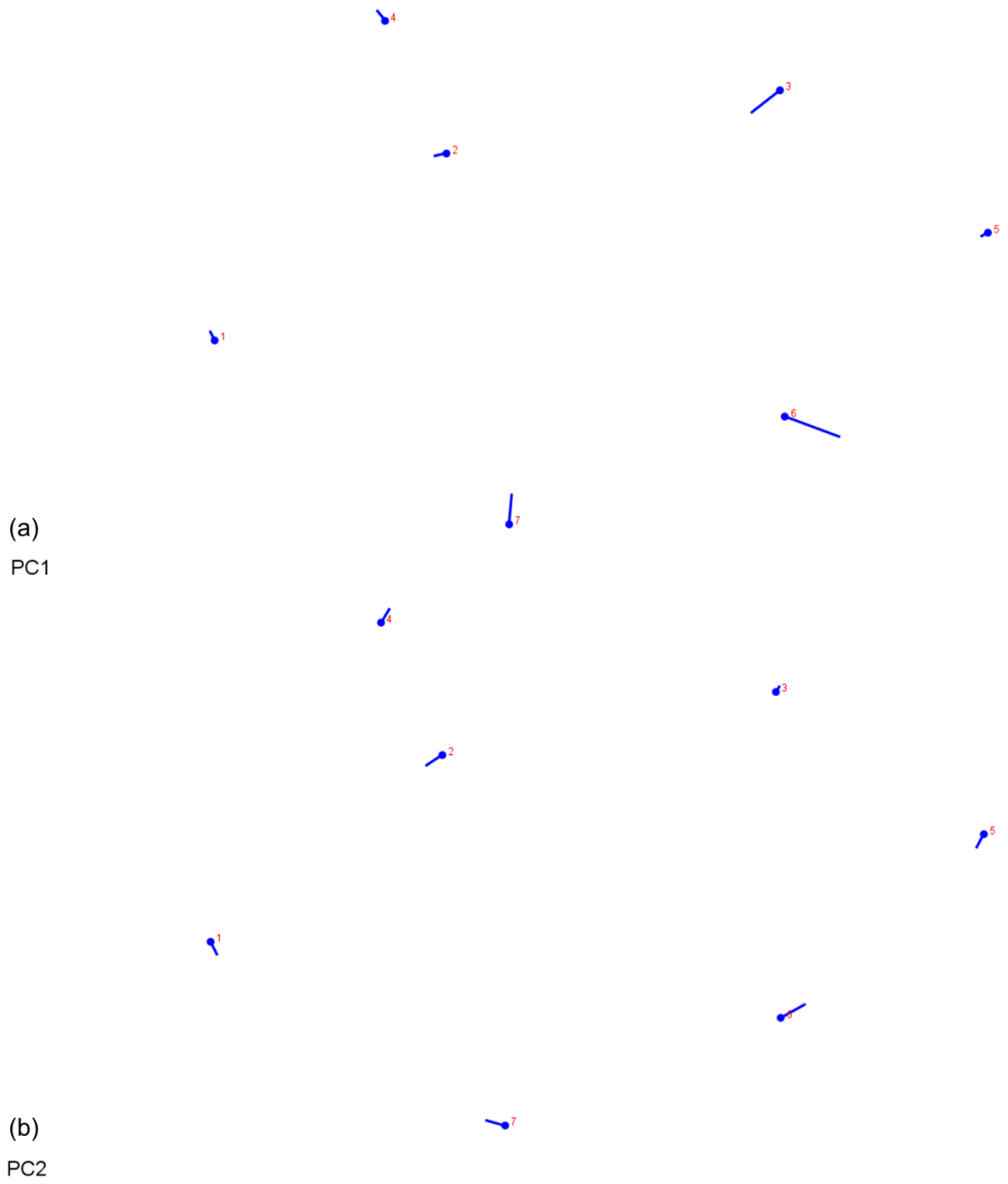

Figure 4.3.1.2. Scree Plots for both PC1 and PC2 of the Intracellular β2-Adrenergic Receptor Landmarks. Scree plots for PC1 (a) and PC2 (b) are shown for the 7 intracellular landmarks of the 36 β2-adrenergic receptor structures.

PC1 vs PC2 Morphospace

Scatterplots comparing PC1 and PC2 show a spread of data across both the extracellular (Figure 4.3.1.3a) and intracellular (Figure 4.3.1.3b) principal components, with evidence of both clustering and outliers. The intracellular landmark coordinates (b) revealed a relatively

compact cluster of points around the -0.10 PC1, -0.00 PC2 morphospace, with the majority of outliers varying towards the positive range of PC1. In contrast, the extracellular landmarks (a) revealed a reasonable cluster of points around the 0.05 PC1, 0.03 PC2 morphospace, but the data were more generally distributed between PC1 and PC2. Importantly, it should be noted that the PC axes for the extracellular data (a) were half the range of the intracellular PC axes (b), reflecting a greater degree of variance in the intracellular landmarks, additionally highlighted by the greater cumulative percentage variation in Table 4.3.1.1.

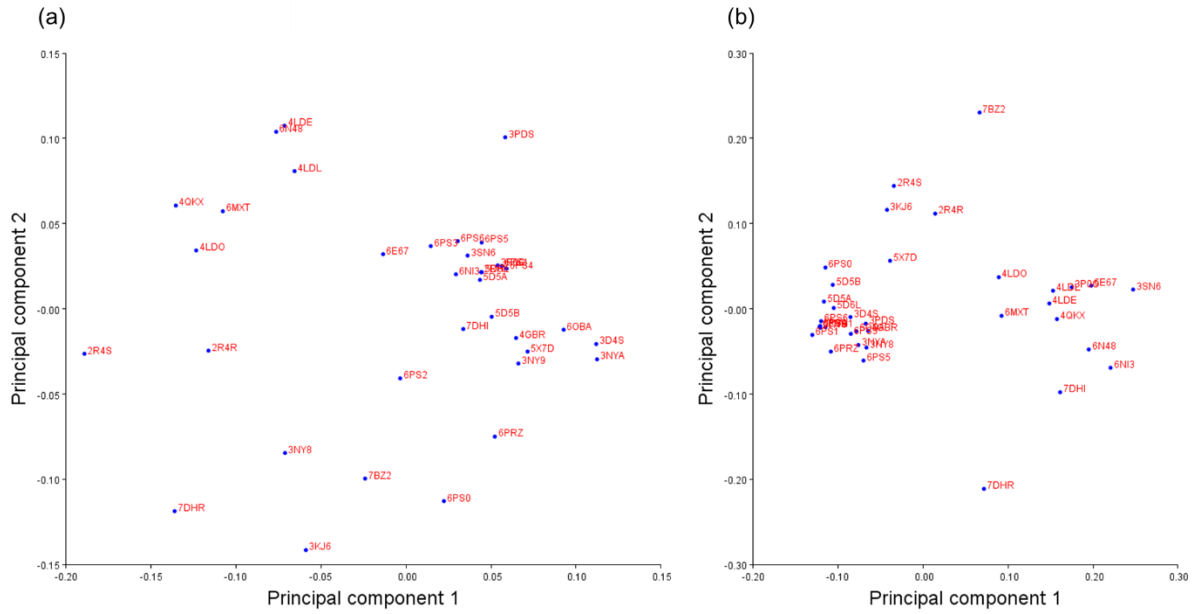

Figure 4.3.1.3. Morphospace of the β2-Adrenergic Receptor Structures. Position in morphospace of 36 β2-adrenergic receptor structures along PC1 and PC2, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Each structure is labelled with its PDB code.

Categorisation by Activation State

For the β2-adrenergic receptor structures then, PC1 represents the dominant source of variation in the dataset, and is likely caused by the differences in structural conformations adopted by these receptors resolved in varying states of activation. If this is indeed the case, one would expect to see active structures aligned to the more positive (righthand) scale of PC1, and the inactive structures aligned to the more negative (lefthand) scale of PC1. Thus, upon categorisation and assignment of group affiliation, this is shown to be the case in Figure 4.3.1.4. Structures described to be inactive are red, active are green, and active with G_s are blue, based on information gathered from the *GPCRdb*. While there is general overlap between these three categories for the extracellular morphospace (Figure 4.3.1.4a), the intracellular morphospace (Figure 4.3.1.4b) reveals a clear distinction between the structures affiliated to the inactive or active groups.

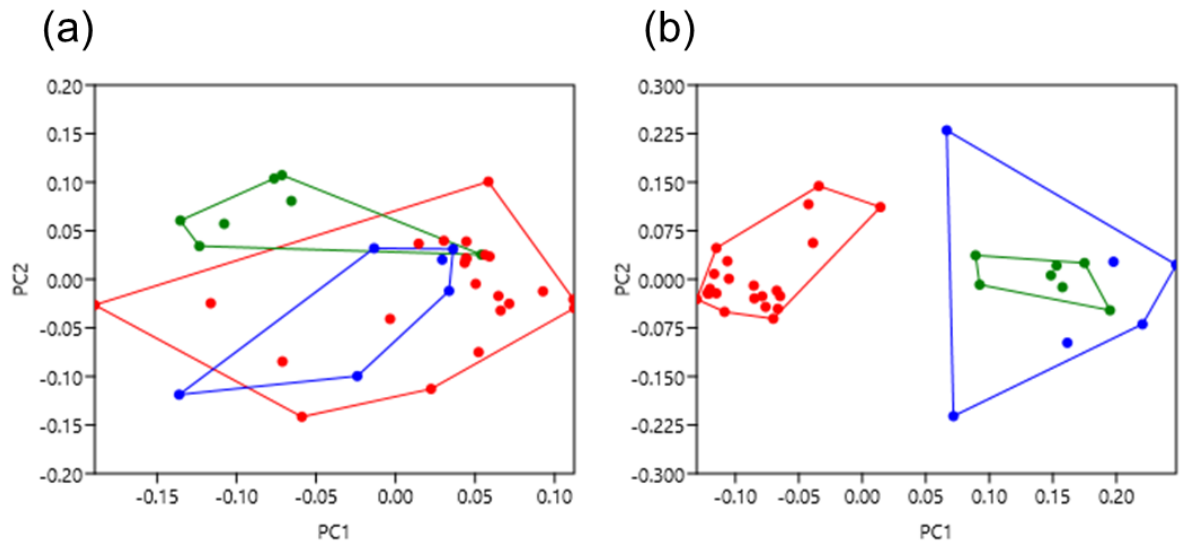

Figure 4.3.1.4. Morphospace of the β2-Adrenergic Receptor Structures, Categorised by Activation State. Position in morphospace of 36 β2-adrenergic receptor structures, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Convex hulls delimit the structures as categorised, indicating group affiliation. Inactive is red, active is green, active with G_s is blue.

PERMANOVA (extracellular F = 6.237, p = 0.0005; intracellular F = 30.84, p = 0.0001) and ANOSIM (extracellular R = 0.2784, p = 0.0069; intracellular R = 0.8904, p = 0.0001) showed that there were significant differences between the categorised groups, summarised in Table 4.3.1.2. For the extracellular morphospace, both PERMANOVA and ANOSIM showed statistically significant differences between the inactive and active groups, and between the active and active with G_s groups. In contrast, the inactive and active with G_s groups were not found to be significantly different, highlighted by the overlap between the two in Figure 4.3.1.4a. For the intracellular morphospace, PERMANOVA showed that the inactive group varies significantly from both the active and active with G_s groups, while ANOSIM showed that all three groups were significantly dissimilar from each other. PERMANOVA did not find the active and active with G_s groups to significantly vary from each other, highlighted by their overlap in Figure 4.3.1.4b. Finally, the greater F and R values for the intracellular morphospace indicate greater shape variation and less dissimilarity, respectively, which suggests a greater degree of variance in the intracellular morphospace compared to the extracellular.

Table 4.3.1.2. Statistical Analyses of the Activation States of the β2-Adrenergic Receptor Structures. Pairwise comparisons of PC1 and PC2 scores of both the extra- and intracellular landmarks of the β2-adrenergic receptor structures, in terms of defined activation states. Structures were categorised as either inactive, active, or active with G_s based on information gained from the *GPCRdb*. PERMANOVA *p* values are above the diagonal, and ANOSIM *p* values are below the diagonal. The Bonferroni sequential corrected significant *p* values are indicated in bold.

Extracellular				Intracellular			
	Inactive	Active	Active with G_s		Inactive	Active	Active with G_s
Inactive		**0.0004**	0.396	Inactive		**0.0001**	**0.0001**
Active	**0.0011**		**0.0056**	Active	**0.0002**		0.8714
Active with G_s	0.3839	**0.0055**		Active with G_s	**0.0001**	**0.0081**	

Categorisation by Bound Ligand

Upon categorisation by bound ligand, the β2-adrenergic receptor structures morphospaces reveal significant differences between the agonist (green), antagonist (blue), inverse agonist (red), and unbound structures (gold), despite their overlap in places. This is more apparent in the intracellular morphospace, with more distinct clustering between the agonist-bound structures, and the antagonist- and inverse agonist-bound structures which are not significantly different from each other. Moreover, these morphospaces reveal general variation in the β2-adrenergic receptor structures, including both PC1 and PC2, which may be consistent with subtle differences in their conformational changes when resolved in complex with these classes of ligand.

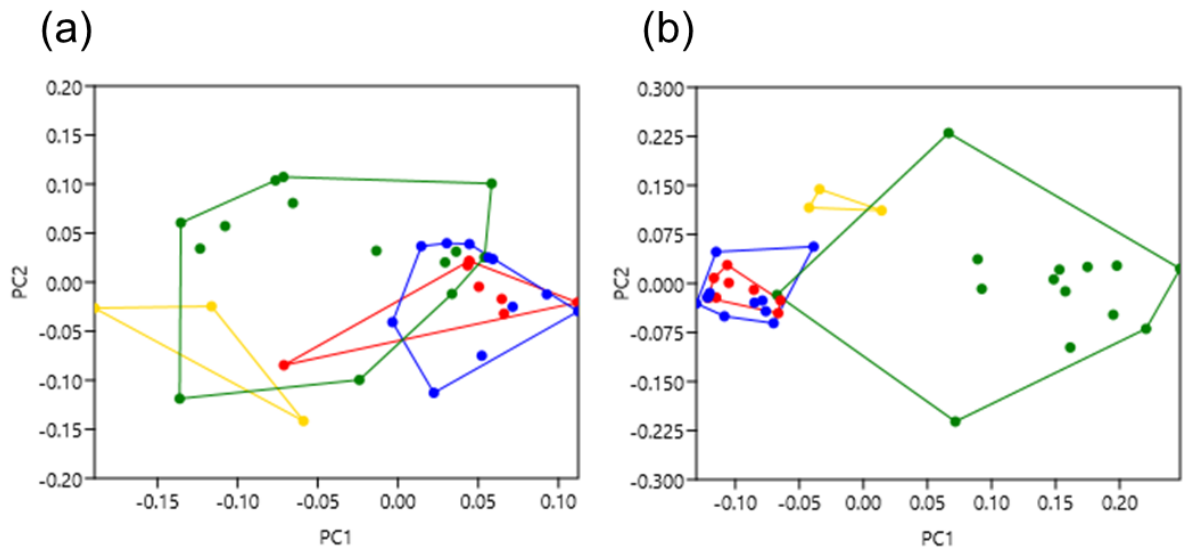

Figure 4.3.1.5. Morphospace of the β2-Adrenergic Receptor Structures, Categorised by Bound Ligand. Position in morphospace of 36 β2-adrenergic receptor structures, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Convex hulls delimit the structures as categorised, indicating group affiliation. Inverse agonist is red, agonist is green, antagonist is blue, and unbound is gold.

PERMANOVA (extracellular F = 6.862, p = 0.0001; intracellular F = 22.1, p = 0.0001) and ANOSIM (extracellular R = 0.2312, p = 0.0019; intracellular R = 0.5096, p = 0.0001) showed that there were significant differences between the categorised groups, summarised in Table 4.3.1.3. For the extracellular morphospace, PERMANOVA showed all bound ligand groups to be significantly different from each other, aside from inverse agonist- compared to antagonist-bound structures. ANOSIM showed the unbound structures to be significantly different from the inverse agonist- and antagonist-bound structures, but not agonist-bound, which were significantly different to the antagonist-bound structures. For the intracellular morphospace, both PERMANOVA and ANOSIM revealed significant differences between all groups aside from the inverse agonist- compared to antagonist-bound groups. Finally, the greater F and R values for the intracellular morphospace also indicate greater shape variation and less dissimilarity, respectively.

Table 4.3.1.3. Statistical Analyses of the Bound Ligands of the β2-Adrenergic Receptor Structures. Pairwise comparisons of PC1 and PC2 scores of both the extra- and intracellular landmarks of the β2-adrenergic receptor structures, in terms of their bound ligand. Structures were categorised as being unbound or agonist-, inverse agonist-, or antagonist-bound, based on information gained from the *GPCRdb*. PERMANOVA *p* values are above the diagonal, and ANOSIM *p* values are below the diagonal. The Bonferroni sequential corrected significant *p* values are indicated in bold.

Extracellular					Intracellular				
	Unbound	Inverse Agonist	Antagonist	Agonist		Unbound	Inverse Agonist	Antagonist	Agonist
Unbound		**0.0073**	**0.0029**	**0.0232**	Unbound		**0.0073**	**0.0029**	**0.0057**
Inverse Agonist	**0.0058**		0.9627	**0.0071**	Inverse Agonist	**0.0058**		0.9033	**0.0001**
Antagonist	**0.0027**	0.8512		**0.0008**	Antagonist	**0.0027**	0.6482		**0.0001**
Agonist	0.0267	0.0658	**0.0077**		Agonist	**0.0182**	**0.0001**	**0.0001**	

4.3.2 - Case Study 2: The Family A Adenosine A2A Receptor

The adenosine receptors elicit responses to extracellular adenosine, and mediate many physiological processes, including the regulation of pain, cerebral blood flow, respiration and sleep (Jaakola *et al.*, 2008). Each of the four adenosine receptor sub-groups primarily couple to the cAMP pathway, but are characterised by unique pharmacological profiles. The adenosine A2A receptor preferentially signals through both $G\alpha_s$ and $G\alpha_{olf}$, increasing cAMP production, and is antagonised by methylxanthine molecules such as caffeine, with moderate affinity (Jacobson *et al.*, 2022). Interestingly, there is strong epidemiological evidence to link the consumption of coffee to a reduced risk of Parkinson's disease, attributed to caffeine's antagonism of the adenosine A2A receptor (Hernán *et al.*, 2002). The development of selective therapeutic compounds which preferentially activate specific adenosine receptor sub-groups would therefore find application in the treatment of pain, asthma, Huntington disease, Parkinson's disease, seizure, and various other neurological disorders, for example (Franco *et al.*, 2022; Mori *et al.*, 2022).

The A2A receptor is one of the most well-crystallised receptors in the GPCR superfamily (Araya *et al.*, 2024), providing structures in various conformational states for exploration with geometric morphometrics. Seminal studies of the A2A receptor closely followed those of rhodopsin and the β2-adrenergic receptor, being crystallised in complex with the high affinity selective antagonist ZM241385 (Jaakola *et al.*, 2008). Not only did this enable comparison to other Family A GPCR structures, but reiterated the concept of ligand selectivity and the associated mechanisms of structure based signalling. Interestingly, the crystal structure of the adenosine A2A receptor in complex with ZM241385 revealed several key differences to previously reported GPCR structures. Firstly, the extracellular loop organisation was observed to be quite different to those of the adrenergic receptors and bovine/squid rhodopsin

(Palczewski *et al.*, 2000; Rasmussen *et al.*, 2007; Rosenbaum *et al.*, 2007; Murakami and Kouyama, 2008; Warne *et al.*, 2008). In addition, the binding of antagonist ZM241385 was observed to be extended perpendicular to the membrane, and co-linear with transmembrane helix 7, which differed from the initial homology modelling based on rhodopsin and the β2-adrenergic receptor structures (Martinelli and Tuccinardi, 2008; Yuzlenko and Kieć-Kononowicz, 2009). Finally, a subtle difference in helical orientation was observed in the A2A receptor when compared to rhodopsin and the β2 receptor, redefining A2A's antagonist binding domain to be closer to TMs 6 and 7, and more limited interaction with TMs 3 and 5 (Jaakola *et al.*, 2008).

As such, the adenosine A2A receptor makes for an interesting comparison to the β2-adrenergic receptor with geometric morphometrics, building on previously reported structural analyses of these Family A GPCRs. At the time of sample selection, 58 human adenosine A2A receptor structures were included in the dataset, with 49 described as inactive, and 9 as active, 2 of which were coupled to G_s but were included in the active category for statistical analyses. No structures were filtered or removed to capture a wide landscape of conformational heterogeneity, and no structures were observed to disproportionately skew the morphospaces of this dataset. Refer to Appendix B2 for a full list of structures, including PDB codes and reference to their publication. Given that more traditional structural analyses have revealed key differences between the A2A receptor and other Family A receptors, it was hypothesised that differences may also be revealed by geometric morphometrics. In addition to this, it was also hypothesised that shape variation would likely remain dominated by the conformational differences between activation states.

PC Variance

The first two principal components accounted for 63% and 81%, of the variation in the extracellular and intracellular landmark coordinates, respectively, of the 58 adenosine A2A receptor structures, as shown in Table 4.3.2.1. The first principal component (PC1) accounted for 47% (extracellular) and 58% (intracellular) of the variation.

Table 4.3.2.1. Principal Component Variance of the Adenosine A2A Receptor Structures. The variance explained by the 14 principal components of 58 adenosine A2A receptor structures, based on both the extra- and intracellular landmark coordinates. Variance is expressed as a percentage, alongside the total cumulative variance.

	Extracellular		Intracellular	
	%Variance	%Cumulative	%Variance	%Cumulative
PC1	47.425	47.425	57.686	57.686
PC2	15.05	62.475	23.614	81.3
PC3	12.222	74.697	8.307	89.607
PC4	9.912	84.61	3.036	92.643
PC5	6.072	90.682	2.473	95.116
PC6	3.508	94.19	2.069	97.184
PC7	2.121	96.311	1.084	98.268
PC8	1.345	97.656	0.951	99.219
PC9	0.986	98.642	0.326	99.545
PC10	0.682	99.324	0.167	99.712
PC11	0.317	99.641	0.148	99.859
PC12	0.251	99.892	0.07	99.929
PC13	0.065	99.957	0.064	99.993
PC14	0.043	100	0.007	100

PC1 and PC2 Scree Plots

Scree plots for the first two principal components, showing the location, direction and magnitude of variations are shown in Figures 4.3.2.1 and 4.3.2.2. For the extracellular landmarks, PC1 showed the greatest variation in the positions of transmembrane helices 1, 5 and 2, with some variation in helices 6 and 7. PC2 showed the greatest variation in helices 1, 2 and 3, and some variation in helices 4 and 7. For the intracellular landmarks, PC1 showed the greatest variation in helix 6, with some variation in helices 1, 3, 5 and 7. PC2 showed the greatest variation in helices 5 and 6, and some variation in helices 1, 2, 3 and 7.

Overall, PC1 for both the extra- and intracellular landmarks again highlighted variation in transmembrane helices 5 and 6, likely corresponding to the common GPCR activation mechanism, and further supporting the results of the β2-adrenergic receptor case study. These results therefore also suggest the structural variation in the adenosine A2A receptor snapshots are dominated by their activation state.

Figure 4.3.2.1. Scree Plots for both PC1 and PC2 of the Extracellular Adenosine A2A Receptor Landmarks. Scree plots for PC1 (a) and PC2 (b) are shown for the 7 extracellular landmarks of the 58 adenosine A2A receptor structures.

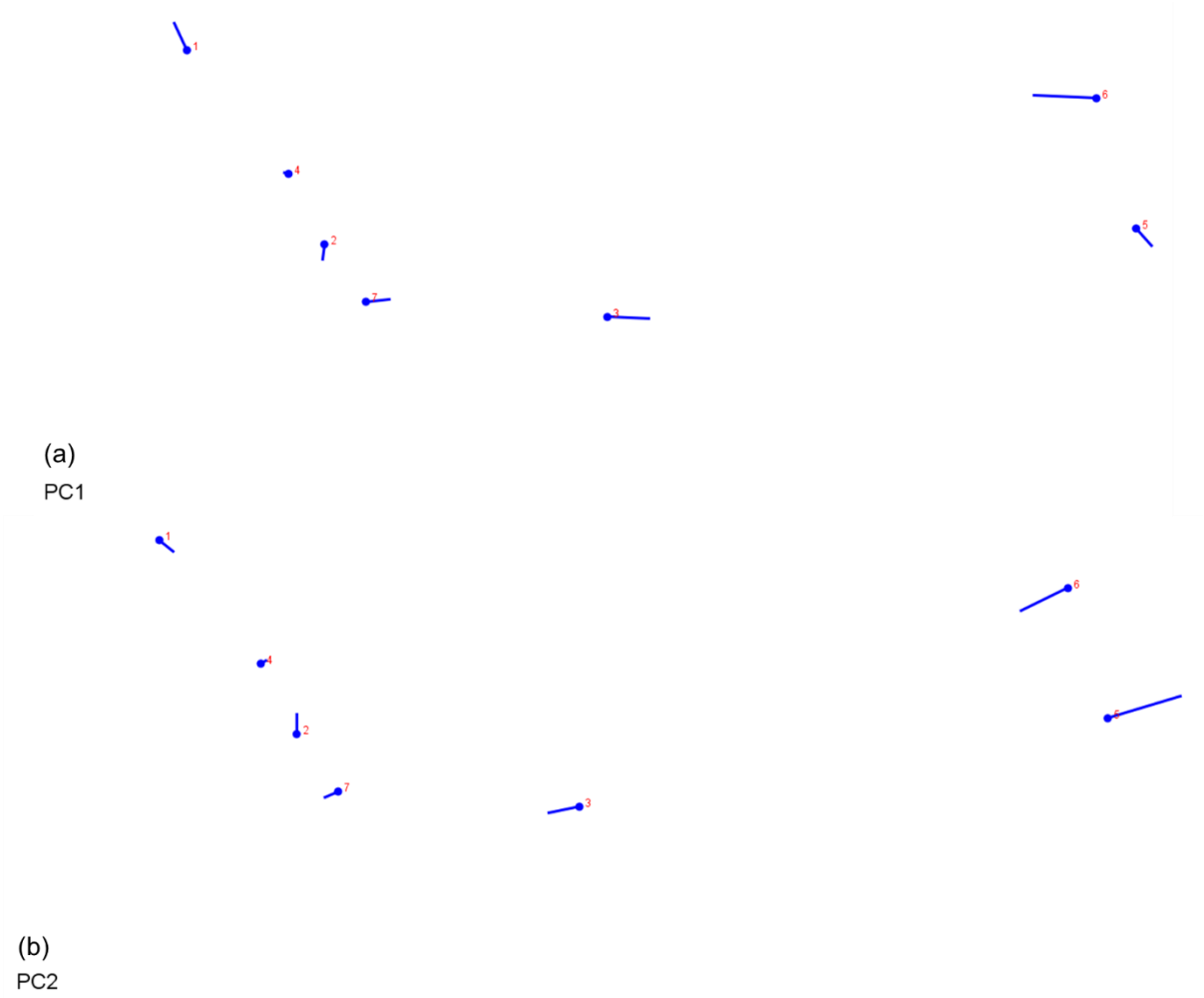

Figure 4.3.2.2. Scree Plots for both PC1 and PC2 of the Intracellular Adenosine A2A Receptor Landmarks. Scree plots for PC1 (a) and PC2 (b) are shown for the 7 intracellular landmarks of the 58 adenosine A2A receptor structures.

PC1 vs PC2 Morphospace

Scatterplots comparing PC1 and PC2 show a spread of data across both the extracellular (Figure 4.3.2.3a) and intracellular (Figure 4.3.2.3b) principal components, with evidence of very strong clustering and outliers. The intracellular landmark coordinates (b) revealed a distinct cluster of points around the -0.15 PC1, -0.00 PC2 morphospace, and several smaller compact clusters towards the positive range of PC1. In contrast, the extracellular landmarks (a) revealed a very strong cluster of points around the 0.05 PC1, 0.00 PC2 morphospace, with data points distributing mainly towards the negative range of PC1 and the positive range of PC2. Similarly to the β2-adrenergic receptor structures, the PC1 axis for the extracellular data (a) was half the range of the intracellular PC1 axis (b), again reflecting a greater degree of variance in the intracellular landmarks, and again highlighted by the greater cumulative percentage variation in Table 4.3.2.1.

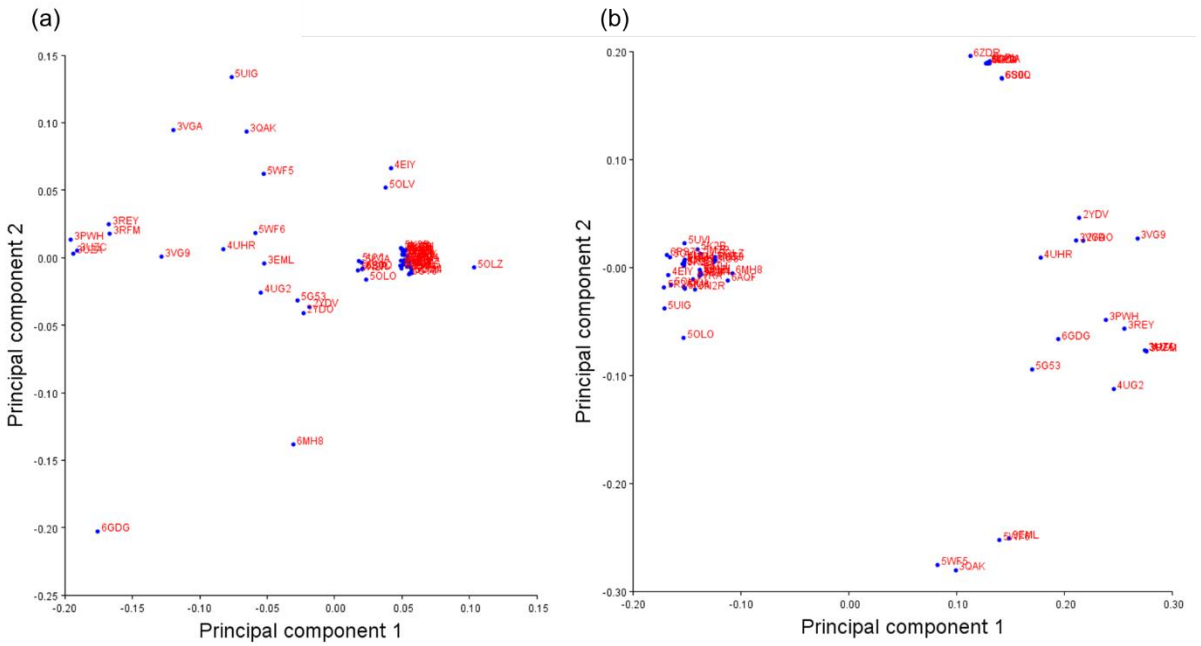

Figure 4.3.2.3. Morphospace of the Adenosine A2A Receptor Structures. Position in morphospace of 58 adenosine A2A receptor structures along PC1 and PC2, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Each structure is labelled with its PDB code.

Categorisation by Activation State

For the adenosine A2A receptor structures then, PC1 also likely represents the differences in structural conformations adopted by these receptors resolved in varying states of activation. Thus, upon categorisation and assignment of group affiliation, this is shown to be the case in Figure 4.3.2.4. Structures described to be inactive are red and active are green, based on information gathered from the *GPCRdb*. While there is a general overlap between these two categories for both the extracellular morphospace (Figure 4.3.2.4a), and the intracellular morphospace (Figure 4.3.2.4b), statistical analyses reveal significant differences between the structures affiliated to the inactive or active groups. Similarly to the β2-adrenergic receptor structures, the adenosine A2A intracellular landmarks of the active structures are also shown to skew towards the positive range of PC1. However, in contrast, the intracellular landmarks of the inactive structures cluster strongly to the negative range of PC1, but also display outliers to the positive range, which differs from the pattern displayed by the β2-adrenergic receptor structures.

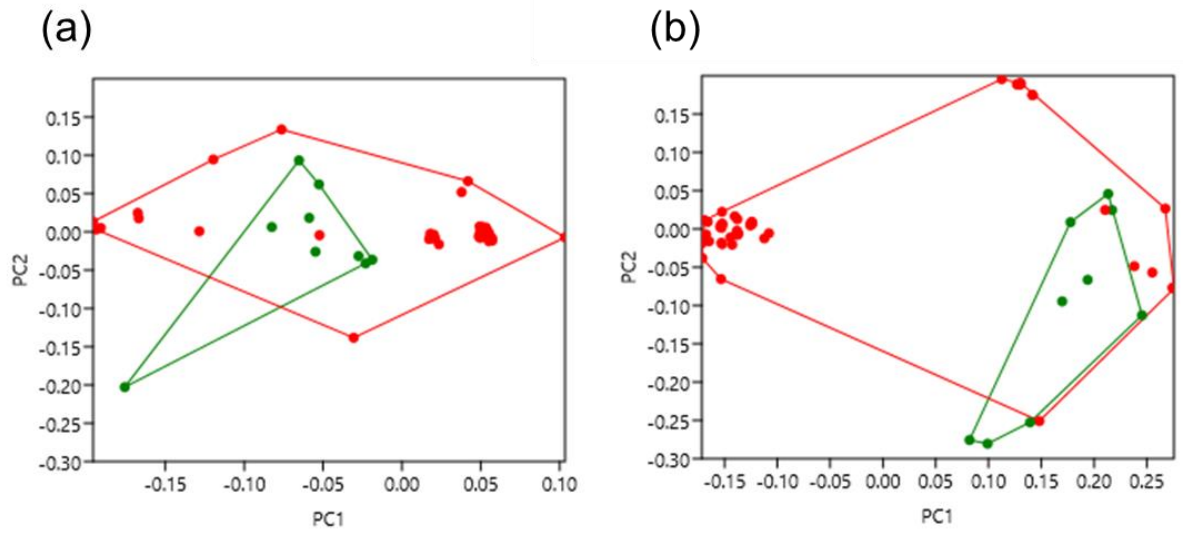

Figure 4.3.2.4. Morphospace of the Adenosine A2A Receptor Structures, Categorised by Activation State. Position in morphospace of 58 adenosine A2A receptor structures, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Convex hulls delimit the structures as categorised, indicating group affiliation. Inactive is red and active is green.

PERMANOVA (extracellular F = 5.768, p = 0.0116; intracellular F = 13.59, p = 0.0003) and ANOSIM (extracellular R = 0.4012, p = 0.0029; intracellular R = 0.3791, p = 0.0001) showed that there were significant differences between the categorised groups, summarised in Table 4.3.2.2. For both the extracellular and intracellular morphospaces, both PERMANOVA and ANOSIM showed statistically significant differences between the inactive and active groups, however, as these groups overlapped, both the F and R values indicated less variation and less dissimilarity across the adenosine A2A receptor structures in general, despite significant pairwise differences.

Table 4.3.2.2. Statistical Analyses of the Activation States of the Adenosine A2A Receptor Structures. Pairwise comparisons of PC1 and PC2 scores of both the extra- and intracellular landmarks of the adenosine A2A receptor structures, in terms of defined activation states. Structures were categorised as either inactive or active based on information gained from the *GPCRdb*. PERMANOVA p values are above the diagonal, and ANOSIM p values are below the diagonal. The Bonferroni sequential corrected significant p values are indicated in bold.

Extracellular			Intracellular		
	Inactive	Active		Inactive	Active
Inactive		**0.0101**	Inactive		**0.0002**
Active	**0.0018**		Active	**0.0001**	

Categorisation by Bound Ligand

When categorising the adenosine A2A receptor structures by bound ligand, both the extracellular and intracellular morphospaces reveal significant differences between the agonist- (green) and antagonist-bound (blue) structures, despite their overlap. In this dataset, there is a spread of datapoints and lack of clear clustering, especially for the antagonist-bound A2AR structures, which likely indicates a variety of subtle conformational differences which could be further associated with specific antagonist ligands. It would be interesting to probe whether these are true 'antagonists', or are actually inducing biased signalling through alternative pathways; geometric morphometrics may therefore be revealing differences here, consistent with the concept of structure-based biased signalling.

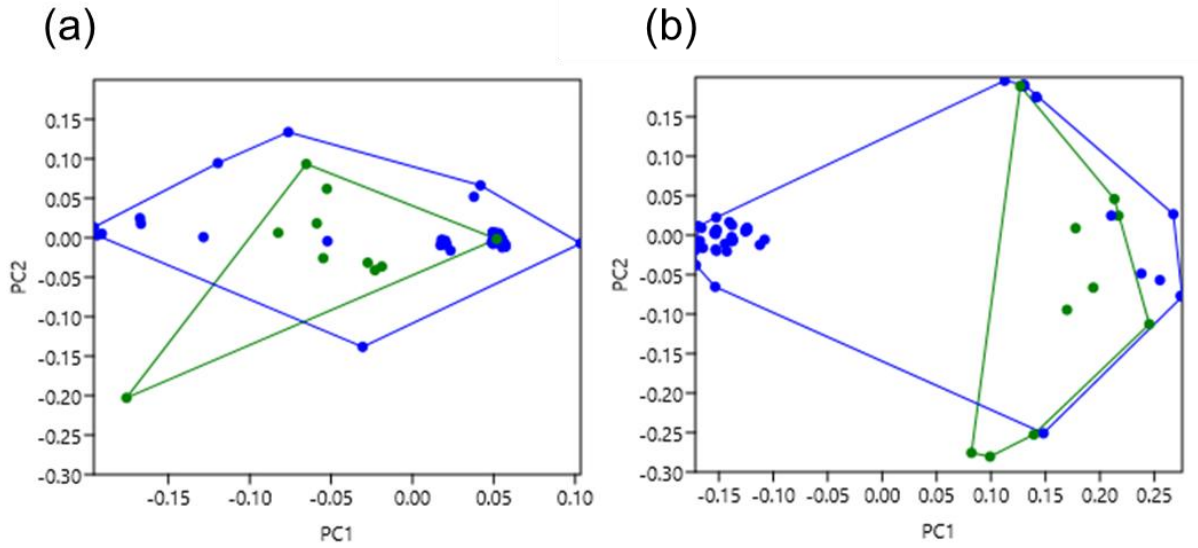

Figure 4.3.2.5. Morphospace of the Adenosine A2A Receptor Structures, Categorised by Bound Ligand. Position in morphospace of 58 adenosine A2A receptor structures, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Convex hulls delimit the structures as categorised, indicating group affiliation. Agonist is green and antagonist is blue.

PERMANOVA (extracellular F = 4.337, p = 0.0234; intracellular F = 12.53, p = 0.0004) and ANOSIM (extracellular R = 0.3182, p = 0.0079; intracellular R = 0.351, p = 0.0005) showed that there were significant differences between the categorised groups, summarised in Table 4.3.2.3. For both the extracellular and intracellular morphospaces, both PERMANOVA and ANOSIM revealed significant differences between the agonist- and antagonist-bound structures, the intracellular morphospace more so. Again, these results were significant, despite the overlap of groups.

Table 4.3.2.3. Statistical Analyses of the Bound Ligands of the Adenosine A2A Receptor Structures. Pairwise comparisons of PC1 and PC2 scores of both the extra- and intracellular landmarks of the adenosine A2A receptor structures, in terms of their bound ligand. Structures were categorised as being either agonist-, or antagonist-bound, based on information gained from the *GPCRdb*. PERMANOVA *p* values are above the diagonal, and ANOSIM *p* values are below the diagonal. The Bonferroni sequential corrected significant *p* values are indicated in bold.

Extracellular			Intracellular		
	Agonist	Antagonist		Agonist	Antagonist
Agonist		**0.0191**	Agonist		**0.0003**
Antagonist	**0.0073**		Antagonist	**0.0003**	

4.3.3 – Case Study 3: The Family B1 Secretin-Like Receptors

The relatively small B1, or secretin-like, family of GPCRs is composed of 15 members in humans, and are characterised by a relatively large extracellular domain and activation by peptide hormones; there are no orphan receptors in family B1 as each receptor has had at least one endogenous agonist identified (Cary *et al.*, 2023). As well as the calcitonin family of receptors discussed in this thesis, family B1 also contains receptors for corticotropin, glucagon, parathyroid hormone, pituitary adenylate cyclase-activating polypeptide (PACAP) and vasoactive intestinal peptide (VIP), to name a few. As family B1 mediates such a variety of peptide hormones, they are also targeted clinically for a wide range of diseases including diabetes, obesity, migraine, hypoglycaemia and osteoporosis, for example (Bortolato *et al.*, 2014). While some effective therapeutics are in use, the broad physiological and clinical relevance of family B1 makes their structural mechanisms of ligand binding and activation a salient topic of interest to the field.

Similarly to both the β2-adrenergic and adenosine A2A receptors, the secretin-like receptors of family B1 preferentially couple to G_s to increase the production of cAMP, however, numerous biased agonists have been identified (Wootten *et al.*, 2017). Despite this, insight into the physiological effects and therapeutic application of the biased agonism of secretin-like receptors is limited (Cary *et al.*, 2023). For example, an engineered biased agonist of the parathyroid receptor, PTH^{7D}, with an increased preference for the cAMP pathway compared to β-arrestin recruitment relative to PTH, did not show an increase in bone tissue growth in mice (White *et al.*, 2021). In contrast, the PTH analogue [$D\text{-}Trp^{12}$, Tyr^{34} – PTH(7-34)] was found to preferentially activate the ERK1/2 pathway via β-arrestin recruitment, stimulating bone growth and decreasing bone degradation in mice (Gesty-Palmer *et al.*, 2006, 2009). These results indicate that the treatment of osteoporosis for example, may be facilitated by these β-arrestin-biased agonists of the PTH receptor, although some results were notably cell-

line dependent (van der Lee *et al.*, 2013). Perhaps the most well-studied receptor of family B1, the glucagon-like peptide 1 receptor (GLP-1R) is pleiotropically activated by several ligands, and is therefore thought to be more disposed to biased agonism. Structures of GLP-1R resolved in complex with biased agonists including semaglutide, oxyntomodulin and exendin 4, for example, have collectively provided insight into the structural mechanisms of GLP-1R, and could lead to the rational structure-based design of biased agonists for these receptors, and the wider B1 family in general (Cary *et al.*, 2023).

As such, the family B1 secretin-like receptors make for an interesting comparison to the family A β2-adrenergic and adenosine A2A receptors, highlighting the multifaceted capabilities of geometric morphometrics across GPCR families. At the time of sample selection, 61 human secretin-like receptors were included in the dataset, with 16 described as inactive, and 45 described as active. No structures were filtered or removed to capture a wide landscape of conformational heterogeneity, and no structures were observed to disproportionately skew the morphospaces of this dataset. Refer to Appendix B3 for a full list of structures, including PDB codes and reference to their publication. It was hypothesised that activation state would again underlie the greatest source of variation in the secretin-like receptors. In addition, observations may also reveal an effect of including several sub-families in this analysis, which was not the case for the β2-adrenergic or adenosine A2A receptor structures.

PC Variance

The first two principal components accounted for 54% and 64%, of the variation in the extracellular and intracellular landmark coordinates, respectively, of the 61 secretin-like receptor structures, as shown in Table 4.3.3.1. The first principal component (PC1) accounted for 35% (extracellular) and 50% (intracellular) of the variation.

Table 4.3.3.1. Principal Component Variance of the Secretin-Like Receptor Structures. The variance explained by the 14 principal components of 61 secretin-like receptor structures, based on both the extra- and intracellular landmark coordinates. Variance is expressed as a percentage, alongside the total cumulative variance.

	Extracellular		Intracellular	
	%Variance	%Cumulative	%Variance	%Cumulative
PC1	34.934	34.934	49.489	49.489
PC2	18.577	53.511	14.178	63.667
PC3	10.953	64.464	10.06	73.727
PC4	8.062	72.526	5.072	78.799
PC5	5.793	78.319	4.053	82.852
PC6	4.364	82.683	3.76	86.612
PC7	3.929	86.612	3.363	89.974
PC8	3.151	89.763	2.695	92.669
PC9	2.632	92.395	2.307	94.977
PC10	2.172	94.568	1.653	96.629
PC11	2.031	96.599	1.311	97.94
PC12	1.498	98.097	1.107	99.047
PC13	1.207	99.304	0.574	99.621
PC14	0.696	100	0.379	100

PC1 and PC2 Scree Plots

Scree plots for the first two principal components, showing the location, direction and magnitude of variations are shown in Figures 4.3.3.1 and 4.3.3.2. For the extracellular landmarks, PC1 showed the greatest variation in the positions of transmembrane helix 6, with some variation in helices 1, 2, 5 and 7. PC2 showed the greatest variation in helix 1, and some variation in helices 2, 4, 5, 6 and 7. For the intracellular landmarks, PC1 showed the greatest variation in helix 6, with some variation in helices 3 and 5. PC2 showed the greatest variation in helices 3, 4 and 6, and some variation in helix 2.

Overall, PC1 for both the extra- and intracellular landmarks again highlighted variation in transmembrane helices 5 and 6, with greater variation observed in the other TM helices, likely due to the inclusion of several receptor sub-types in this dataset. In any case, it is reassuring to consistently observe the greatest magnitudes of variation being attributed to underlying conformational changes between the activation states of crystal snapshots.

Figure 4.3.3.1. Scree Plots for both PC1 and PC2 of the Extracellular Secretin-Like Receptor Landmarks. Scree plots for PC1 (a) and PC2 (b) are shown for the 7 extracellular landmarks of the 61 secretin-like receptor structures.

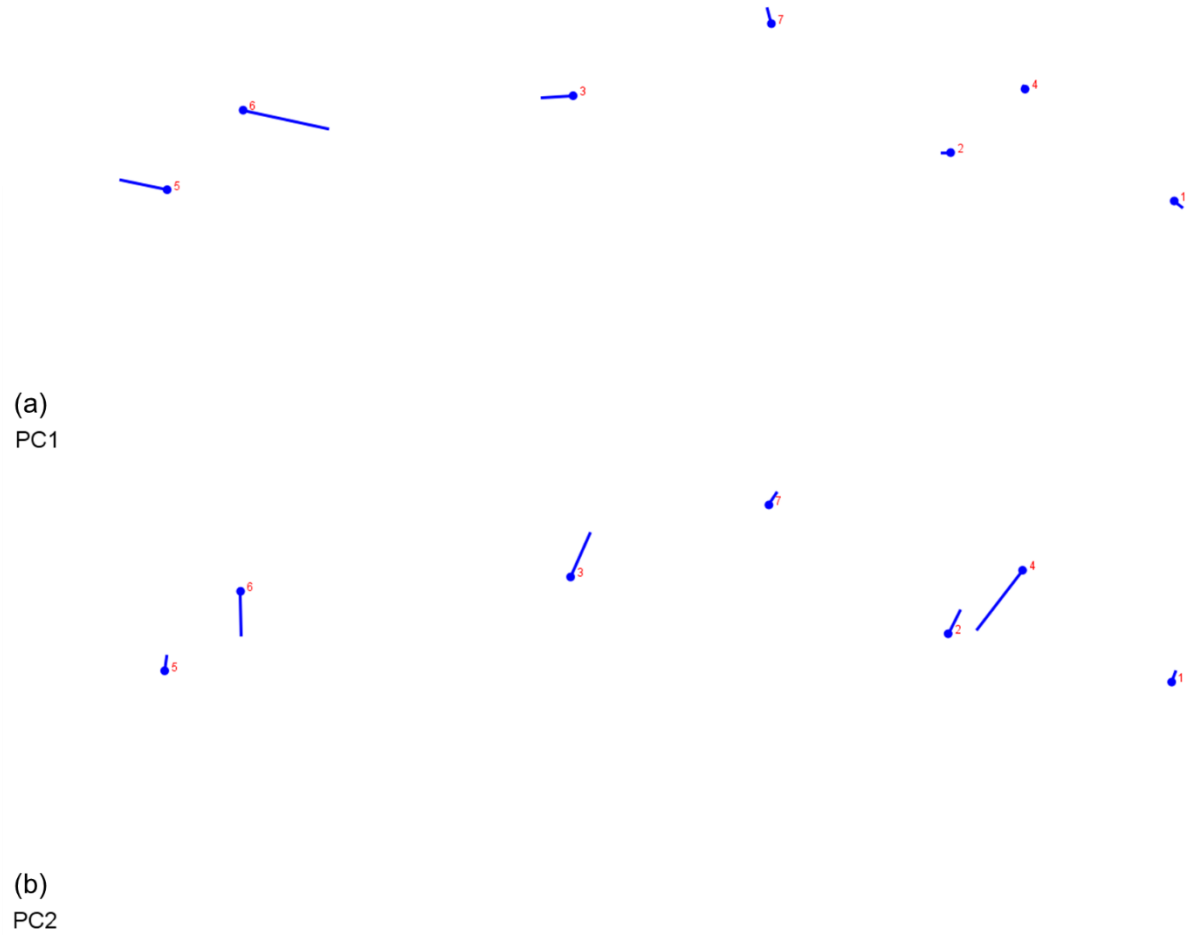

Figure 4.3.3.2. Scree Plots for both PC1 and PC2 of the Intracellular Secretin-Like Receptor Landmarks. Scree plots for PC1 (a) and PC2 (b) are shown for the 7 intracellular landmarks of the 61 secretin-like receptor structures.

PC1 vs PC2 Morphospace

Scatterplots comparing PC1 and PC2 show a spread of data across both the extracellular (Figure 4.3.3.3a) and intracellular (Figure 4.3.3.3b) principal components, with evidence of both clustering and outliers. The intracellular landmark coordinates (b) revealed a general clustering of points between the -0.10 PC1, -0.10 to 0.10 PC2 morphospace, and a more distinct cluster towards the positive range of PC1 (0.30, -0.00). In contrast, the extracellular landmarks (a) revealed a small cluster of points around the -0.05 PC1, -0.00 PC2 morphospace, with data points distributing more generally between both PC1 and PC2. Similarly to both the β2-adrenergic receptor and adenosine A2A receptor structures, the PC1 axis for the extracellular data (a) were less positive than the intracellular PC1 axis (b), highlighted by the greater cumulative percentage variation in Table 4.3.3.1.

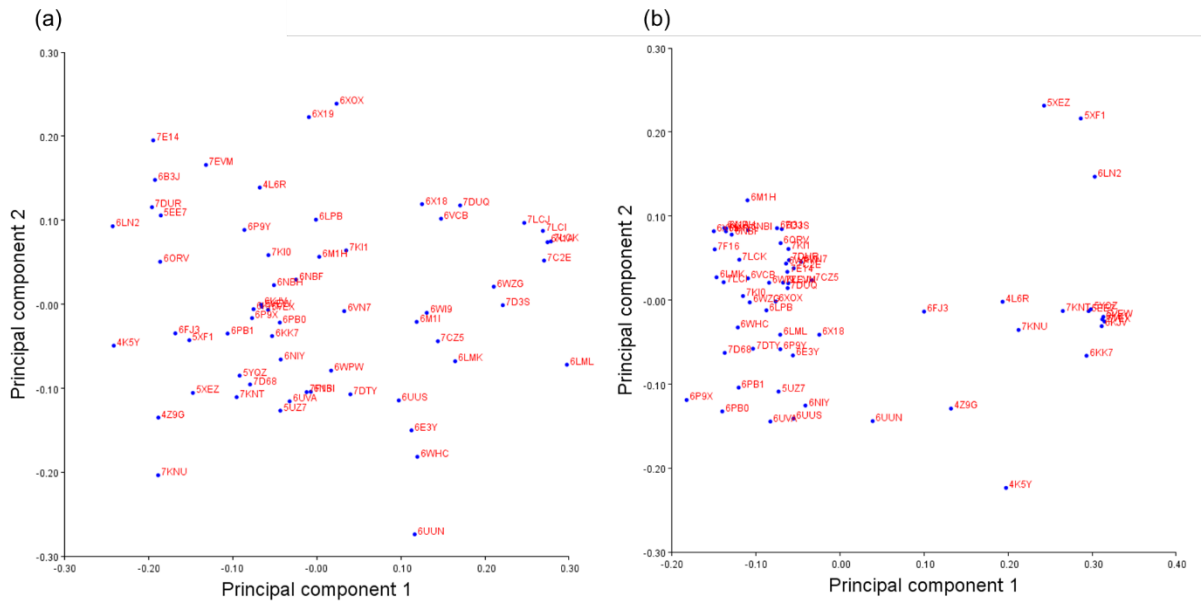

Figure 4.3.3.3. Morphospace of the Secretin-Like Receptor Structures. Position in morphospace of 61 secretin-like receptor structures along PC1 and PC2, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Each structure is labelled with its PDB code.

Categorisation by Activation State

For the secretin-like receptor structures then, PC1 also likely represents the differences in structural conformations adopted by these receptors resolved in varying states of activation. Thus, upon categorisation and assignment of group affiliation, this is shown to be the case in Figure 4.3.3.4. Structures described to be inactive are red and active are green, based on information gathered from the *GPCRdb*. While there is some overlap between these two categories for the extracellular morphospace (Figure 4.3.3.4a), the intracellular morphospace forms two distinct groups with the active structures falling in the negative range of PC1, and the inactive structures to the positive range of PC1 (Figure 4.3.3.4b). This contrasts to the β2-adrenergic and adenosine A2A receptors in which their active structures were skewed to the positive range of PC1. This possibly suggests that while PC1, and therefore activation state, is the dominant source of variation in these datasets, perhaps the positivity or negativity of the groupings along the principal components differs for Family B1 receptors when compared to Family A. Nonetheless, these groupings were also significantly different from each other.

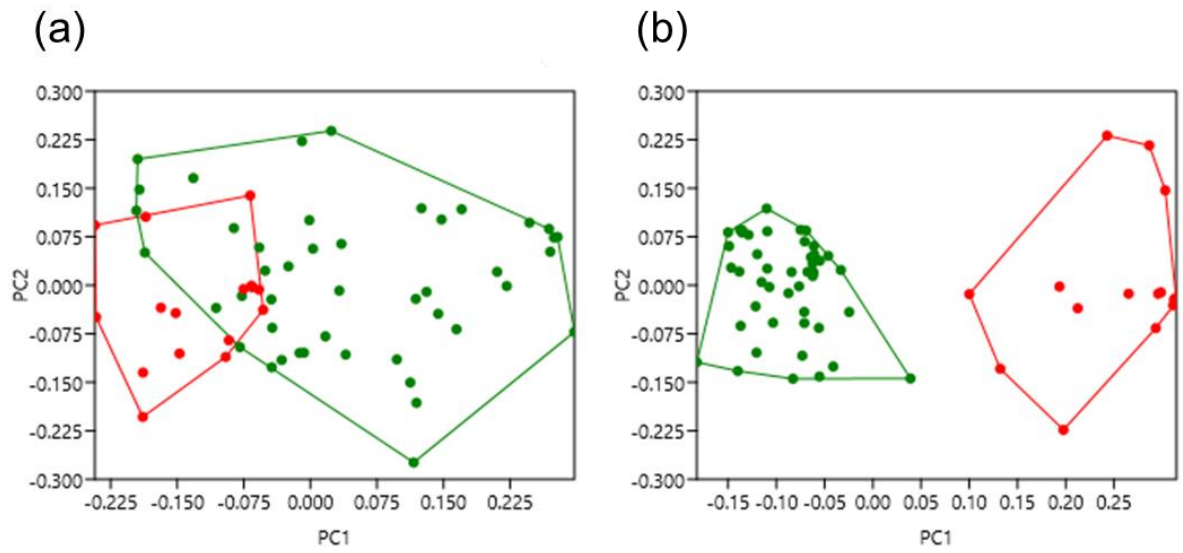

Figure 4.3.3.4. Morphospace of the Secretin-Like Receptor Structures, Categorised by Activation State. Position in morphospace of 61 secretin-like receptor structures, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Convex hulls delimit the structures as categorised, indicating group affiliation. Inactive is red and active is green.

PERMANOVA (extracellular F = 14.48, p = 0.0001; intracellular F = 139.7, p = 0.0001) and ANOSIM (extracellular R = 0.13, p = 0.0232; intracellular R = 0.9696, p = 0.0001) showed that there were significant differences between the categorised groups, summarised in Table 4.3.3.2. For the extracellular morphospace, PERMANOVA showed statistically significant differences between the inactive and active groups while ANOSIM did not; this indicates significant variation but a lack of clustering. The F and R values were both low, indicating lesser variation and dissimilarity between groups. However, for the intracellular morphospace, both PERMANOVA and ANOSIM showed very statistically significant differences between the inactive and active groups, with very high F and R values indicating a great degree of variation and dissimilarity between the active and inactive structures due to their clear distinction from each other.

Table 4.3.3.2. Statistical Analyses of the Activation States of the Secretin-Like Receptor Structures. Pairwise comparisons of PC1 and PC2 scores of both the extra- and intracellular landmarks of the secretin-like receptor structures, in terms of defined activation states. Structures were categorised as either inactive or active based on information gained from the *GPCRdb*. PERMANOVA *p* values are above the diagonal, and ANOSIM *p* values are below the diagonal. The Bonferroni sequential corrected significant *p* values are indicated in bold.

Extracellular			Intracellular		
	Inactive	Active		Inactive	Active
Inactive		**0.0001**	Inactive		**0.0001**
Active	0.0235		Active	**0.0001**	

Categorisation by Bound Ligand

Upon categorising the secretin-like receptors by their bound ligand, the extracellular morphospace reveals a general spread of data points across both PC1 and PC2, with overlap of groups, and a lack of significant clustering aside from a difference between the agonist- (green) and antagonist-bound (blue) structures. It was hypothesised that perhaps the inclusion of several sub-families in this dataset might have overpowered the effects of bound ligand, and categorisation by sub-family may therefore reveal greater differences. This was not the case, with sub-family categorisation also giving no significant differences for both morphospaces (data not shown). The intracellular morphospace however, not only showed clearer groupings of bound ligand, but PC1 also revealed a gradient from most active (green) through to antagonised (blue), and then least active being negatively allosterically modulated (fuschia), with unbound structures falling in between (gold). These differences could also reflect the conformational heterogeneity of the family B1 receptors when resolved in complex with these different classes of ligand.

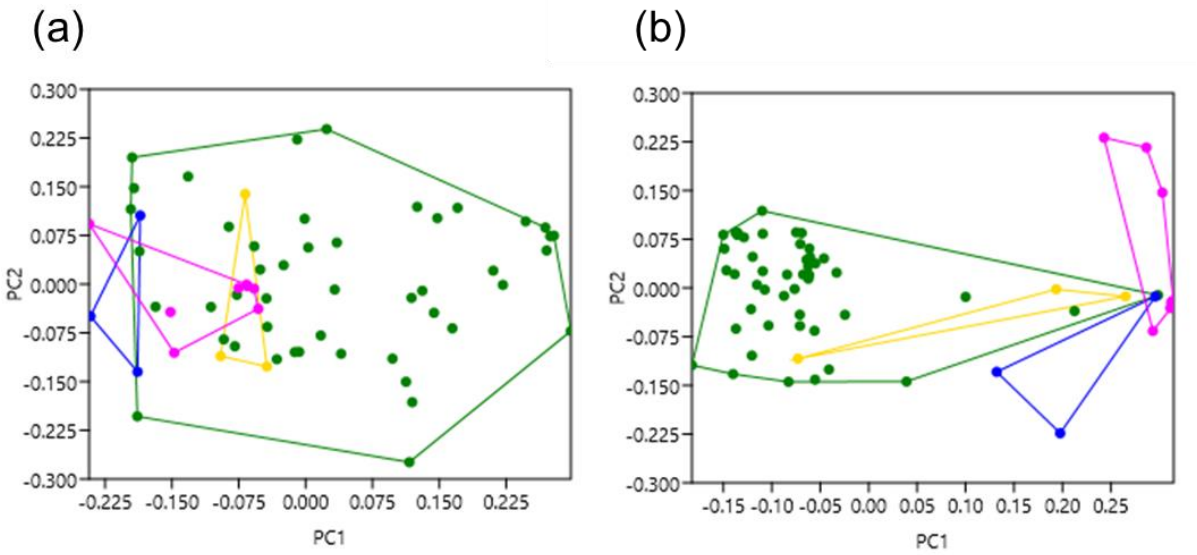

Figure 4.3.3.5. Morphospace of the Secretin-Like Receptor Structures, Categorised by Bound Ligand. Position in morphospace of 61 secretin-like receptor structures, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Convex hulls delimit the structures as categorised, indicating group affiliation. Agonist is green, antagonist is blue, fuschia is negative allosteric modulator (NAM) and unbound is gold.

PERMANOVA (extracellular F = 3.428, p = 0.05; intracellular F = 28.75, p = 0.0001) and ANOSIM (extracellular R = 0.01417, p = 0.5494; intracellular R = 0.8005, p = 0.0001) showed that there were significant differences between the categorised groups, summarised in Table 4.3.3.3. For the extracellular morphospace, PERMANOVA showed a significant difference between the agonist- and antagonist-bound structures only, and ANOSIM did not show any significant differences. For the intracellular morphospace, both PERMANOVA and ANOSIM revealed significant differences between the agonist- and unbound, antagonist-, and NAM-bound structures. Defined groups in the intracellular morphospace therefore varied more and formed more distinct clusters, as indicated by the greater F and R values.

Table 4.3.3.3. Statistical Analyses of the Bound Ligands of the Secretin-Like Receptor Structures. Pairwise comparisons of PC1 and PC2 scores of both the extra- and intracellular landmarks of the secretin-like receptor structures, in terms of their bound ligand. Structures were categorised as being unbound or agonist-, antagonist-, or negative allosteric modulator (NAM)-bound, based on information gained from the *GPCRdb*. PERMANOVA *p* values are above the diagonal, and ANOSIM *p* values are below the diagonal. The Bonferroni sequential corrected significant *p* values are indicated in bold.

Extracellular					Intracellular				
	Unbound	Agonist	Antagonist	NAM		Unbound	Agonist	Antagonist	NAM
Unbound		0.3557	0.3944	0.7537	Unbound		**0.009**	0.4989	0.0174
Agonist	0.5748		**0.0048**	0.0123	Agonist	**0.0038**		**0.0004**	**0.0001**
Antagonist	0.4964	0.0287		0.137	Antagonist	0.5958	**0.0007**		0.0401
NAM	0.0978	0.7707	0.0648		NAM	0.1041	**0.0001**	0.1119	

4.3.4 – Case Study 4: The Family C Calcium-Sensing Receptor

The family C calcium-sensing receptor detects the fluctuating levels of calcium ions in the blood, maintaining their homeostasis in conjunction with the reabsorption of Ca^{2+} in the kidneys, and the secretion of PTH from parathyroid cells (Brown, 2013). These receptors belong to the family C group of GPCRs, along with the $GABA_B$ and metabotropic glutamate receptors, functioning as disulphide-linked obligate homodimers. They possess a much larger extracellular domain, with a Venus flytrap module (VFT), a cysteine-rich domain, and the 7TM bundle which relays activation signals from the VFT to the intracellular G proteins (Geng *et al.*, 2016; Zhang *et al.*, 2016). Dysfunction of the calcium-sensing receptor are therefore involved in Ca^{2+} homeostasis disorders such as autosomal dominant hypocalcaemia and familial hypocalciuric hypercalcaemia, for example (Ward *et al.*, 2012). Current therapeutics include calcium mimetics such as cinacalcet, etelcalcetide, and evocalcet, which are positive allosteric modulators, utilised during dialysis for kidney disease (Leach *et al.*, 2016).

The recent resolution of several full-length family C receptor structures has provided an invaluable framework to further elucidate their activation mechanisms. It is proposed that both Ca^{2+} and L-Trp co-activate the calcium-sensing receptor, causing closure of the VFT module and subsequent conformational change of the 7TM bundle which leads to signalling (Ling *et al.*, 2021). The hallmark characteristic of calcium-sensing receptor activation mechanisms is a rearrangement of a TM5-TM5 dimer interface in the inactive state, to a TM6-TM6 interface in the active state (Chen *et al.*, 2021). In terms of their biased agonism, well-known orthosteric agonists have been explored across several signalling pathways, revealing a bias in the potency of barium, spermine, neomycin, and tobramycin via $G_{q/11}$, $G_{i/o}$ and ERK1/2 signalling pathways (Thomsen, Hvidtfeldt and Bräuner-Osborne, 2012). Moreover, the polyamine and aminoglycoside agonists were also found to preferentially signal via ERK1/2, indicating that

the calcium-sensing receptor displays ligand-dependent conformational heterogeneity, which may lead to clinically effective biased therapeutics for the calcium-sensing receptor.

In any case, this receptor makes for an interesting comparison to the other case studies due to its structural nature as an obligate homodimer, which again reinforces the multifaceted nature of geometric morphometrics, as well as completing the analyses of all three major GPCR families for humans - families A, B and C. At the time of sample selection, 11 human calcium-sensing receptor structures were included in the dataset, with 7 described as inactive, and 4 described as active. No structures were filtered or removed to capture a wide landscape of conformational heterogeneity, and no structures were observed to disproportionately skew the morphospaces of this dataset. Refer to Appendix B4 for a full list of structures, including PDB codes and reference to their publication. It was again hypothesised that the greatest source of variation in these structures would be underpinned by their activation state, however, as this was a completely novel use of this technique with dimeric structures, it was unclear if any significant variation would be detected in a similar manner to the previous three case studies.

PC Variance

As the calcium-sensing receptor is an obligate homodimer, each analysis was performed on 14 landmarks instead of the usual 7; as a result, the dimensionality of the data was 35 which is greater than the sample size, resulting in 10 principal components instead of the usual 14. The first two principal components accounted for 73% and 80%, of the variation in the extracellular and intracellular landmark coordinates, respectively, of the 11 calcium-sensing receptor structures, as shown in Table 4.3.4.1. The first principal component (PC1) accounted for 57% (extracellular) and 52% (intracellular) of the variation, whereas PC2 accounted for 16% and 28%, indicating that the variance represented by PC2 is much more prevalent in the intracellular landmarks.

Table 4.3.4.1. Principal Component Variance of the Calcium-Sensing Receptor Structures. The variance explained by the 10 principal components of 11 calcium-sensing receptor structures, based on both the extra- and intracellular landmark coordinates. Variance is expressed as a percentage, alongside the total cumulative variance.

	Extracellular		Intracellular	
	%Variance	%Cumulative	%Variance	%Cumulative
PC1	56.824	56.824	51.536	51.536
PC2	16.219	73.043	28.305	79.841
PC3	10.845	83.888	6.175	86.016
PC4	5.513	89.401	4.459	90.475
PC5	3.027	92.428	3.525	94
PC6	2.675	95.103	2.2	96.199
PC7	2.169	97.272	1.674	97.873
PC8	1.439	98.711	0.959	98.832
PC9	0.915	99.627	0.865	99.697
PC10	0.373	100	0.303	100

PC1 and PC2 Scree Plots

Scree plots for the first two principal components, showing the location, direction and magnitude of variations are shown in Figures 4.3.4.1 and 4.3.4.2. Within the dimeric structure, the chain A monomer is landmarks 1 to 7, and the chain B monomer is landmarks 8 to 14. For the extracellular landmarks, PC1 showed the greatest variation in the positions of transmembrane helices 1, 5, 6, 8, and 13, with some variation in helices 4, 7, 11, 12, and 14. PC2 showed the greatest variation in helices 5, 11, and 12, and some variation in helices 3, 6, 7, 10, 13, and 14. For the intracellular landmarks, PC1 showed the greatest variation in helices 3, 5, 6, 12, 13, and 14, with some variation in helices 1, 4, 7, 8, 10, and 11. PC2 showed the greatest variation in helices 5 and 12, and some variation in helices 3, 4, 6, 7, 9, 10, 11, 13, and 14.

Overall, PC1 for both the extra- and intracellular landmarks again highlighted some of the greatest variation in transmembrane helix 6 for both monomers of the dimer which is a recurring observation in the morphometric analyses of GPCR structures. Compared to the other case study data, the calcium-sensing receptor structures did reveal variance in most of the 14 landmarks, for which there may be several explanations. Firstly, this case study is the first example of the use of geometric morphometrics with an obligate homodimer, and the inclusion of the 14 landmarks instead of the usual 7 may contribute to this variation. Secondly, this analysis was performed on a smaller sample size of 11 structures; subtle variations are amplified in these smaller datasets, making them appear greater than they perhaps would in a larger dataset. Finally, the complete activation mechanism of the calcium-sensing receptor is yet to be fully elucidated but the presence of the Venus fly trap module, the concept of

transactivation, and the conformational heterogeneity observed in their structures may all contribute to the structural changes underlying this variation in shape morphology.

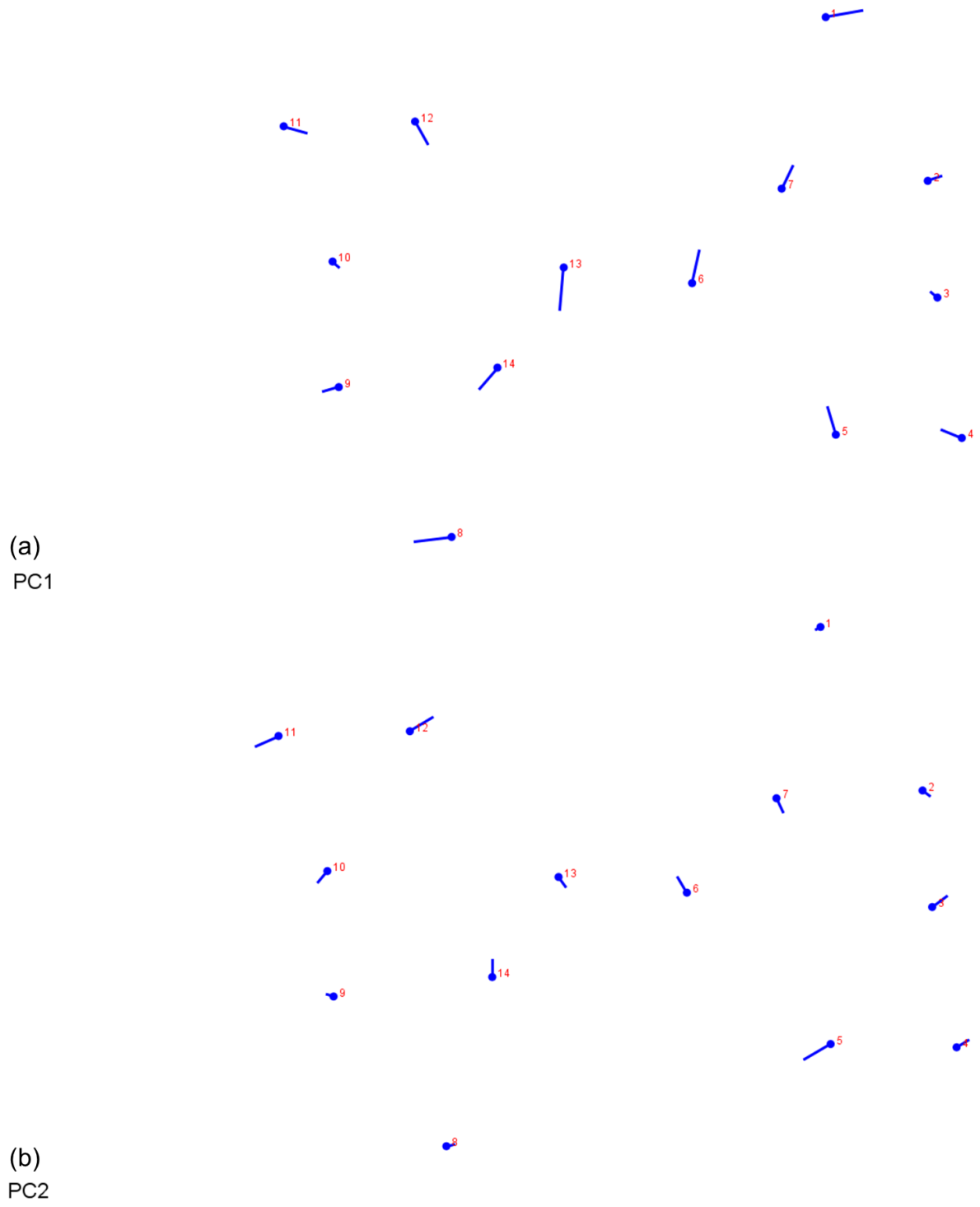

Figure 4.3.4.1. Scree Plots for both PC1 and PC2 of the Extracellular Calcium-Sensing Receptor Landmarks. Scree plots for PC1 (a) and PC2 (b) are shown for the 14 extracellular landmarks of the 11 calcium-sensing receptor structures. Chain A of the dimer is 1 to 7, chain B is 8 to 14.

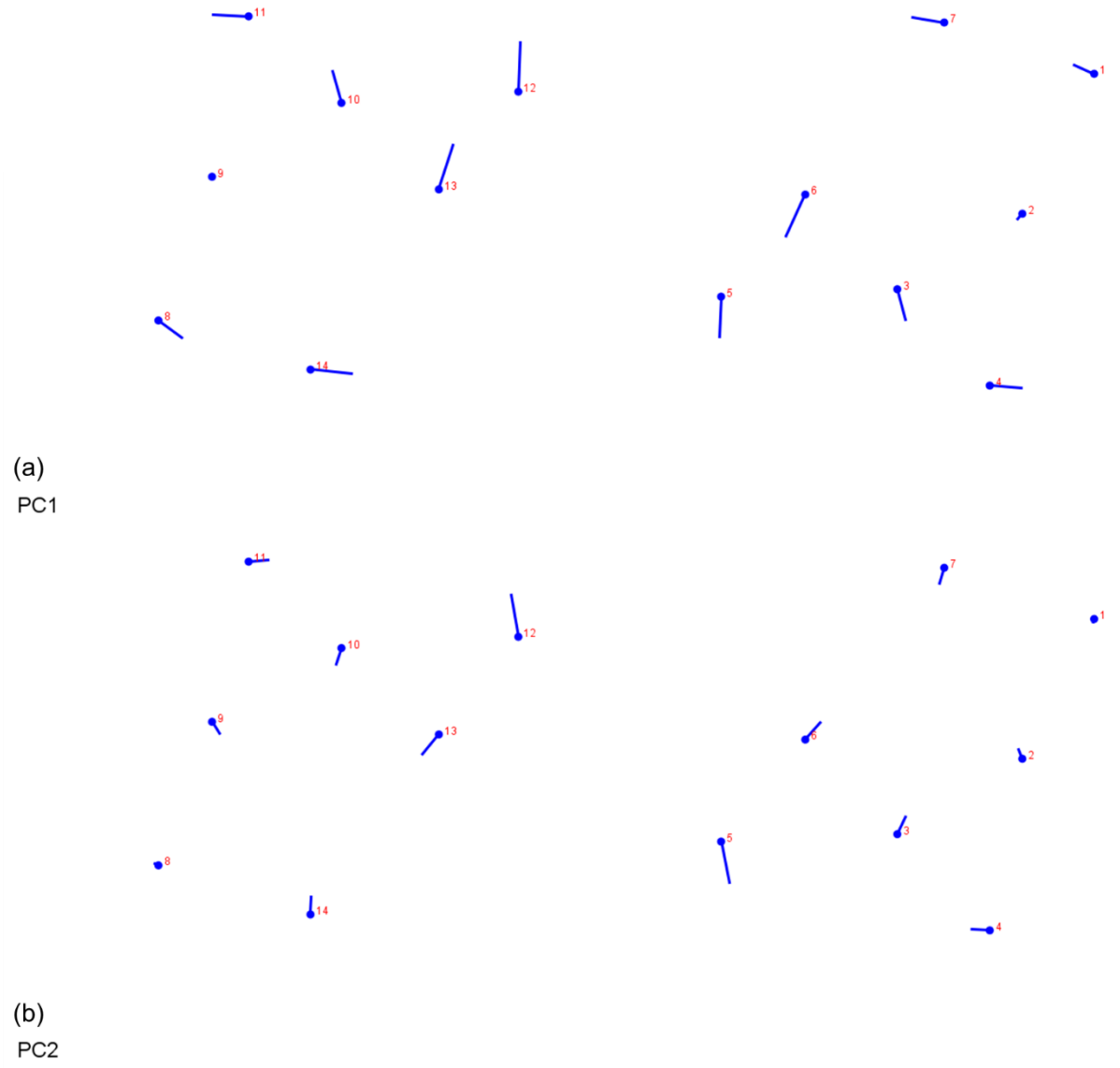

Figure 4.3.4.2. Scree Plots for both PC1 and PC2 of the Intracellular Calcium-Sensing Receptor Landmarks. Scree plots for PC1 (a) and PC2 (b) are shown for the 14 intracellular landmarks of the 11 calcium-sensing receptor structures. Chain A of the dimer is 1 to 7, chain B is 8 to 14.

PC1 vs PC2 Morphospace

Scatterplots comparing PC1 and PC2 show a general spread of data across both the extracellular (Figure 4.3.4.3a) and intracellular (Figure 4.3.4.3b) principal components, with some indication of clustering, albeit less obvious than the other case study morphospaces. The data points are fairly distributed across both PC1 and PC2, although one particular cluster is observed for the intracellular landmarks (b) at the 0.10 PC1, -0.05 PC2 morphospace. In contrast to the other three case studies, the axes for the extracellular data (a) were more positive than the intracellular axes (b), indicating a greater range of variation at the extracellular face of the calcium-sensing receptor.

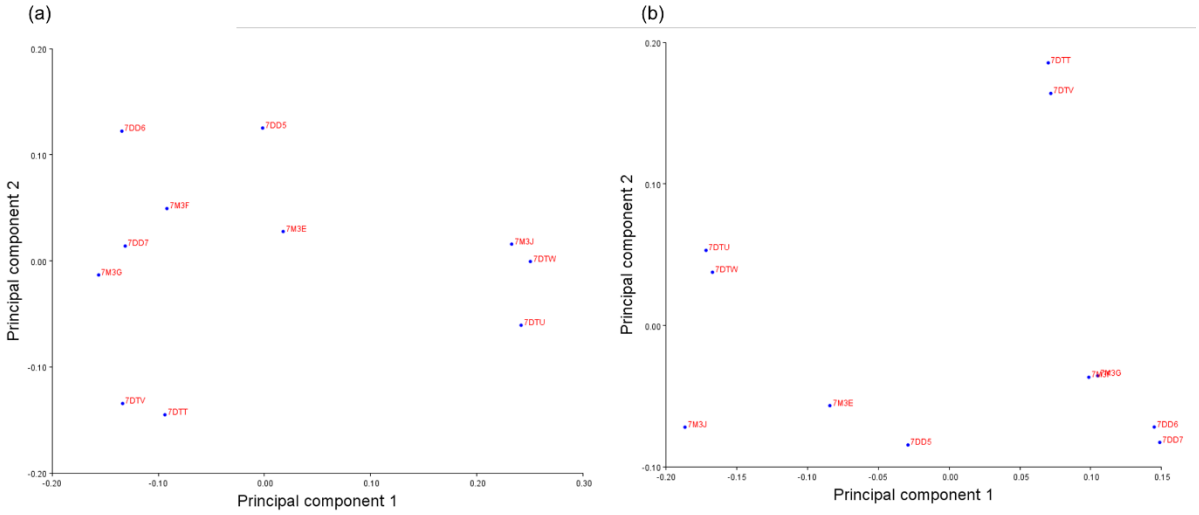

Figure 4.3.4.3. Morphospace of the Calcium-Sensing Receptor Structures. Position in morphospace of 11 calcium-sensing receptor structures along PC1 and PC2, based on the 14 extracellular landmarks (a) and 14 intracellular landmarks (b). Each structure is labelled with its PDB code.

Categorisation by Activation State

For the calcium-sensing receptor structures then, both PC1 and PC2 play a role in representing the differences in structural conformations adopted by these receptors resolved in varying states of activation. Upon categorisation and assignment of group affiliation, Figure 4.3.3.4 shows complete separation between the groups of inactive and active structures, for both extra- and intracellular morphospaces. Structures described to be inactive are red and active are green, based on information gathered from the *GPCRdb*. While the active structures were observed to cluster relatively closer together, the inactive structures occupied a larger morphospace which may underly the conformational heterogeneity in these receptors; indeed, the calcium-sensing receptors are known to adopt intermediate conformations between activation states. In addition to this observation, the groupings revealed along PC1 were mirrored between the extra- and intracellular morphospaces; active structures in the extracellular morphospace (Figure 4.3.4.4a) fell in the negative range of PC1, but then fell in the positive range in the intracellular morphospace (Figure 4.3.4.4b), for example. In any case, the calcium-sensing receptor groupings were shown to be significantly different from each other.

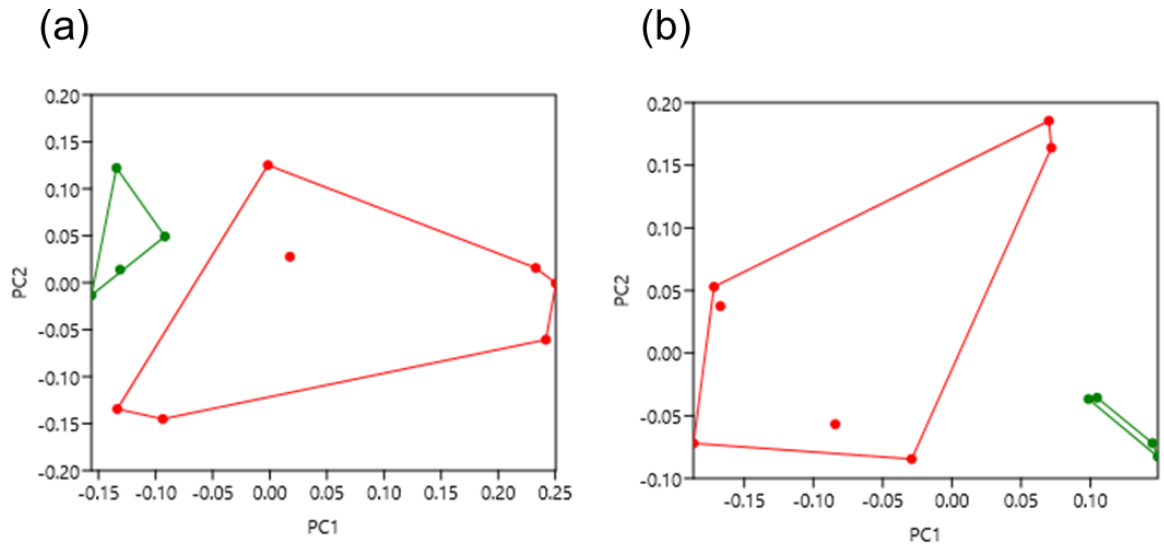

Figure 4.3.4.4. Morphospace of the Calcium-Sensing Receptor Structures, Categorised by Activation State. Position in morphospace of 11 calcium-sensing receptor structures, based on the 14 extracellular landmarks (a) and 14 intracellular landmarks (b). Convex hulls delimit the structures as categorised, indicating group affiliation. Inactive is red and active is green.

PERMANOVA (extracellular F = 4.464, p = 0.0368; intracellular F = 6.922, p = 0.0084) and ANOSIM (extracellular R = 0.2513, p = 0.0498; intracellular R = 0.4709, p = 0.0166) showed that there were significant differences between the categorised groups, summarised in Table 4.3.4.2. For the extracellular morphospace, PERMANOVA showed statistically significant differences between the inactive and active groups while ANOSIM did not; this indicates significant variation but a lack of clustering, similar to the secretin-like receptors. The F and R values were both relatively low, indicating lesser variation and dissimilarity between groups. However, for the intracellular morphospace, both PERMANOVA and ANOSIM showed statistically significant differences between the inactive and active groups, with higher F and R values indicating greater variation and dissimilarity between the active and inactive structures due to their clear distinction from each other, again, similar to the secretin-like receptors.

Table 4.3.4.2. Statistical Analyses of the Activation States of the Calcium-Sensing Receptor Structures. Pairwise comparisons of PC1 and PC2 scores of both the extra- and intracellular landmarks of the calcium-sensing receptor structures, in terms of defined activation states. Structures were categorised as either inactive or active based on information gained from the *GPCRdb*. PERMANOVA *p* values are above the diagonal, and ANOSIM *p* values are below the diagonal. The Bonferroni sequential corrected significant *p* values are indicated in bold.

Extracellular			Intracellular		
	Inactive	Active		Inactive	Active
Inactive		**0.0364**	Inactive		**0.0084**
Active	0.0525		Active	**0.0171**	

Categorisation by Bound Ligand

Due to the lower sample size of this dataset, statistical testing could not be performed in terms of grouping by bound ligand, however, untested observations reveal some differences. The most active and most inactive calcium-sensing receptor structures were found to be at the two extreme opposites of PC1, with agonist-PAM-PAM-bound structures first (bright green), followed by agonist-PAM-bound (bright blue) and agonist-bound (dark green). On the other extreme of PC1 are the apo- (silver), PAM- (orange), and NAM-bound structures (red), with an absence of agonist. Agonist-NAM-PAM-bound structures (dark blue) fell in between the two. Overall, these results are indicative of a gradient of activation across PC1, depending on the presence or absence of agonist and allosteric modulators in the resolved structures. In addition, there also appears to be a spread of data across PC2, even between the active structures which may correspond to subtle differences in the conformational states of the calcium-sensing receptor. As discussed, this variation may appear to be more prevalent due to the lower sample size, and would need more calcium-sensing receptor structures to be resolved to enable statistical testing, and determine if these observations remain consistent in a larger sample size.

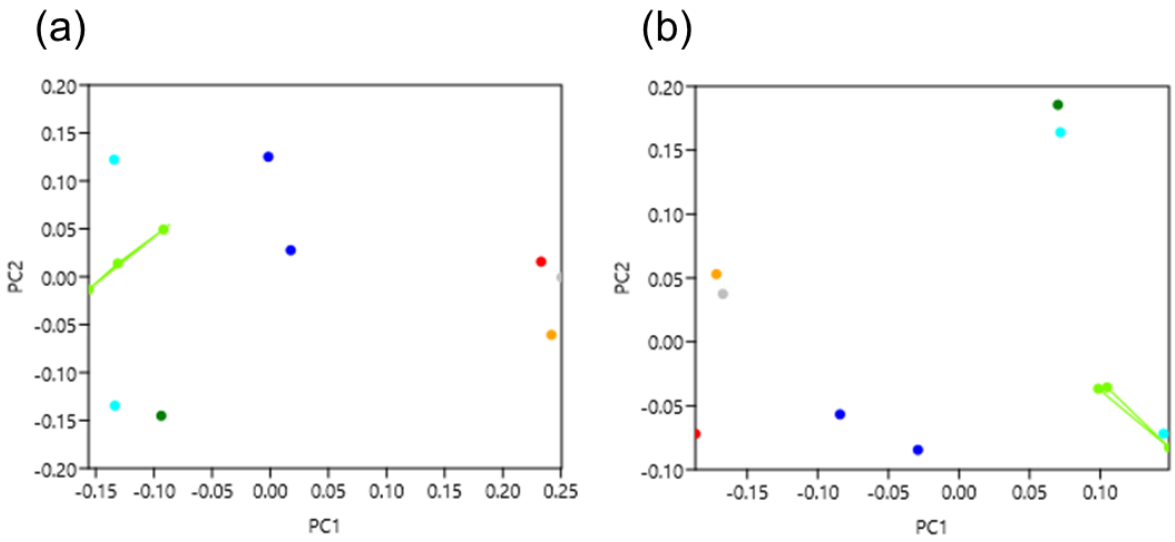

Figure 4.3.4.5. Morphospace of the Calcium-Sensing Receptor Structures, Categorised by Bound Ligand. Position in morphospace of 11 calcium-sensing receptor structures, based on the 14 extracellular landmarks (a) and 14 intracellular landmarks (b). Convex hulls delimit the structures as categorised, indicating group affiliation. Agonist is dark green, ago-PAM-PAM is bright green, ago-PAM is bright blue, ago-NAM-PAM is dark blue, PAM is orange, NAM is red, and apo is silver.

4.4 – Discussion

To reiterate, the wide involvement of GPCRs in physiological processes makes them very attractive drug targets, though the development of safe and effective drugs is often hindered by the challenging nature of these membrane proteins. This is particularly true when considering the possibility of therapeutic agents able to selectively modulate specific GPCR signalling pathways, in which a robust understanding of the structure-function relationship of GPCRs is essential. While the techniques available for structure determination and analysis have produced high-resolution snapshots of GPCR architecture, providing invaluable molecular insights, their inherent limitations prevent visualisation of the complete landscape of the nuanced intricacies of GPCR signalling. Moreover, the elucidation of structure-based biased signalling remains a formidable challenge, and requires a multidisciplinary approach of biophysical, computational and analytical techniques to holistically resolve. This body of work therefore sought to quantitatively decode the conformational heterogeneity in GPCR signalling via the novel application of geometric morphometric and principal component analyses to resolved GPCR structures.

Historically, the most commonly used quantitative method to compare protein structures is root mean square deviation, or RMSD, which gives a similarity metric based on the

superimposition of atomic coordinates (Carugo and Pongor, 2001; Kufareva and Abagyan, 2012). Despite its popularity, RMSD has several inherent limitations. Firstly, RMSD calculations are sensitive to initial structural alignment, and thus differences in superimposition can lead to significantly different RMSD values. The calculation itself also requires a reference structure, which if poorly chosen or unrepresentative of the comparison, may give RMSD values which do not accurately reflect the structural differences observed. As RMSD measures global structural similarity and all atoms are treated equally, more local differences may not be captured accurately; this is particularly relevant when considering the subtle conformational heterogeneity of GPCRs. Importantly, RMSD is unable to distinguish between functionally or biologically relevant structural differences, as opposed to trivial or irrelevant artifacts, and can also be influenced by the resolution of structures, even for similar conformations (Kufareva and Abagyan, 2012). To address these limitations, additional metrics are often employed in tandem, including global distance testing, template modelling and sequence-based comparisons.

A superimposition-independent method of comparing protein structures is Quality scoring, or the Q score, and has been suggested to be an alternative to the global distance test, or GDT (Krissinel and Henrick, 2004). The Q score estimates the similarity between protein structures, based on their internal residue distances but similarly to RMSD, also has inherent limitations (McGuffin and Roche, 2010). Q score relies on the definition of residue contacts derived from the native structure, and therefore if incorrect, may not provide a reliable score of structural similarity. Q score is also sensitive to small changes in the definition of contacts, leading to significant changes in score, in response to the resolution of techniques used to define contacts. In contrast to RMSD, Q score primarily focusses on local interactions and therefore may not capture global structures effectively, including those exhibited by conformational changes. Furthermore, and similarly to RMSD, Q score cannot distinguish between different types of structural rearrangements, regardless of their functional or biological relevance (Kufareva and Abagyan, 2012). Finally, as Q score is primarily applicable to the evaluation of models determined through simulations or *de novo* prediction, it is perhaps less suitable for the comparison of experimentally resolved protein structures.

Based on these two examples then, RMSD may struggle to accurately capture local, and perhaps subtle, differences in protein structures, while Q score cannot capture global differences such as conformational changes between activation states. Moreover, both of these techniques are unable to distinguish between functionally or biologically relevant structural differences, making them particularly unsuitable for this body of work. Geometric morphometrics was therefore selected for its advantages in analysing shape variation between biological structures.

First of all, geometric morphometrics is a quantitative method of shape analysis, allowing precise, reproducible and comparable measurements of structures; the method is grounded in robust statistical principles which enables rigorous analyses between shape differences and patterns of variation (Bookstein, 1984, 1997). GM can simultaneously analyse both shape and size variation, offering powerful visual tools to represent such variation including deformation grids and the principal component plots used in this body of work. As discussed, GM techniques often use procrustean superimposition to remove differences in position, orientation and scale, making comparisons between structures independent of these factors (Mitteroecker and Gunz, 2009). GM can not only handle complex shapes with irregular or non-linear structures, such as proteins, but is also able to provide insights into functionally or biologically relevant shape differences. Whilst this is certainly true for the anthropological study of bones, teeth and skulls, the novel application of GM to GPCRs in this body of work provides evidence that is consistent with functionally relevant conformational landscapes (Wiseman *et al.*, 2021). Finally, GM also allows for both multivariate and comparative analyses, facilitating the exploration of complex patterns of shape variation and covariation across multiple dimensions, as well as comparative analyses of shape across different receptor families. The justification for the novel use of geometric morphometrics to help elucidate the conformational landscapes of GPCRs was therefore evidence-based, enabling robust testing of hypotheses and rigorous statistical analyses, as detailed in section 4.3, and was concluded to be a successful application of this technique to GPCR structural biology.

The foundational geometric morphometric analyses presented in each of the case studies demonstrated consistent quantitative mapping of global structural rearrangements between inactive- and active-state receptor crystals, in addition to revealing ligand-stabilised conformational heterogeneity which is consistent with the breadth of diverse activated states adopted during signalling. Extending this framework across multiple receptor conformations, as well as different members of the GPCR superfamily, highlights the power of this multifaceted, transformative structural tool, and its ability to contribute towards the ‘complete picture’ of structure-based signalling. Ultimately, a holistic understanding of the biased ligand-receptor conformational landscape will lead to dynamic structure/function models from which safe, effective and selective therapeutics can be developed.

However, despite its usefulness, geometric morphometrics does of course possess its own challenges and limitations which require consideration going forwards. This technique heavily relies upon landmark selection to define the shapes to be analysed, which can be a subjective and time-consuming process, especially for more complex or three-dimensional structures such as membrane proteins. In addition, alignment of coordinate data by Procrustese superimposition may not always yield biologically meaningful alignments, and the

interpretation of shape variation can be challenging, especially when the underlying biological factors driving those variations are not well understood. Moreover, the high-dimensional datasets produced by geometric morphometrics can also be challenging to visualise and interpret, and their analyses rely on certain statistical assumptions which may not always be appropriate, violations of which may impact the accuracy and reliability of the results.

One particularly important aspect to highlight in the use of geometric morphometrics with GPCR structures, is that of landmark selection. Landmarks are specifically defined to be standardised positions which are located exactly in all shapes to be compared (Webster and Sheets, 2010). In this body of work, the selection of residues at the ends of each transmembrane helix are relatively standardised and reasonably well conserved between receptors, but do require consideration. Firstly, given that GPCRs exhibit variability in the sequence and length of helices between different families, the selection of the ends of helices enables the capture of global conformational changes, despite these differences between family members. In this context, the definition of landmarks by residue name or sequence number may have been too restrictive, and thus the ends of helices were deemed more appropriate. The selection of GPCR landmarks is of course dependent on the quality and completeness of crystal structures, and thus selecting these landmarks also accommodates the variability of maintained helices, especially when problematic crystal contacts or loop interactions are encountered. Furthermore, the selection of end of helix landmarks additionally allowed for the robustness of the geometric morphometric method to be tested, and was shown to be resilient against this helical variability between the GPCR families. Indeed, the comparison of GPCR structures by the ends of their transmembrane helices is a proven concept, linking conformational heterogeneity across GPCR sub-families to structure-based G protein coupling selectivity (Kruse *et al.*, 2012); this concept is further explored with geometric morphometrics in Chapter 5.

In any case, the manual selection and definition of landmarks remains non-trivial, open to subjectivity, and has the potential to miss more subtle shape variations in receptor conformation; the refinement of the landmark selection strategy will help to alleviate these challenges, and further enhance the precision, efficiency and scope of landmarks which are biologically meaningful and reproducible. It may well be that the use of artificial intelligence and machine learning in landmark selection will prove to be advantageous; machine learning involves computers ‘learning’ from experience, whereby a series of rules map example data to expected outcomes, and a subsequent application of those rules to unseen data provides probable answers (Chen and Siu, 2020). Machine learning has already been implemented in the prediction of protein structures, and the assessment of their quality, with applications to predictive structure-based drug discovery, for example (Jianlin Cheng, Tegge and Baldi, 2008;

Chen and Siu, 2020; Schauperl and Denny, 2022). It likely makes sense to use machine learning to refine the mapping of receptor landmarks to the conformational landscape, and link structure to functional outcomes, in other words, elucidating the structure-based signalling of GPCRs. Crucially, this refinement of landmark selection also relies upon high-quality, high-resolution structure determination from X-ray crystallography and cryo-EM, each of which necessitate their own continual refinement, as discussed (Danev *et al.*, 2021).

In order to truly bridge the divide between structure and function, it will be critical to directly link morphological states identified through geometric morphometrics to specific signalling outputs. The enhancement of the resolution and accuracy of the geometric morphometric analyses of GPCR structures will enable the consistent capture of subtle conformational differences, crucial to understanding their pharmacology. This is especially the case in terms of structure-based signalling bias, by which the morphological heterogeneity observed in these foundational case studies is consistent with the diversity of conformations adopted by GPCRs in their various states of activation. The link between ligand binding, conformation and distinct signalling profiles can therefore be decoded by geometric morphometrics and correlated with biochemical and cellular assays, to ensure that the structural groupings are functionally relevant.

As the case study results in this body of work indicate that the principal components representing the greatest variation in each dataset are dominated by global structural rearrangements between inactive- and active-state receptor crystals, further investigation of the lower eigenvalues may be needed to probe the more subtle structural variations. Referring back to the example of the β2-adrenergic receptor and carazolol (Cherezov *et al.*, 2007; van der Westhuizen *et al.*, 2014), principal components 1 and 2 might be dominated by carazolol's effects on major signalling pathways as a partial inverse agonist, however, the lower ranked principal components may reveal its biased agonism of more minor pathways such as ERK1/2. Again, bridging the gap between these more subtle conformational landscapes and their biased signalling outcomes necessitates the implementation of sophisticated, temporally relevant assays, in combination with quantitative morphological analyses. Ultimately, the detailed mapping of GPCR conformational landscapes can inform the design of novel, selective therapeutics which target specific receptor states for improved efficacy and reduced side effects. Indeed, this method is not solely limited to GPCRs, but has the potential for broad applicability to other clinically relevant targets such as kinases, ion channels and transcription factors, for example (Wang *et al.*, 2020).

Importantly, in its current iteration, the use of the geometric morphometrics method with principal component analysis and multivariate statistics with GPCR structures is certainly

effective and reproducible, however, there are several improvements which could be made to enhance its user-friendliness. As discussed, the landmark selection process can be enhanced by sequence alignments, and the use of artificial intelligence and machine learning, and the collection of coordinate data from these landmarks would also benefit from some level of automation. Currently, data collection remains a manual process, requiring each PDB file to be opened as text, each residue located, and their XYZ coordinates recorded; the automated extraction of this data using Python script, for example, would be a significant improvement to the user experience. This would enable analyses of much larger datasets, without the onerous task of manual data collection, but would rely on the quality of PDB annotations, which do vary, especially for older submissions.

Similarly, the analyses of coordinate data and the statistical testing process require the import and export of various data across several software programs; the development of one coherent software package which integrates all of these aspects seamlessly would no doubt significantly improve the user experience. Aside from the utility of this, if subtle morphological differences are to be explored in the lower ranked principal components, their comparisons, morphospace groupings and statistical testing should also be easily accessible in one coherent package. In the current method, this is not the case, and therefore means the morphological analyses of more subtle structural differences can be missed if they are not specifically focussed upon. Finally, the integration of the geometric morphometric analyses of GPCRs with commonly used bioinformatic and structural services such as the *GPCRdb*, *mpstruc* and PDB depositories would link receptor morphology and structure-based signalling to the existing information provided, and more widely share this technique with the structural biology community.

Conclusion

While highly promising, challenges remain in comprehensively mapping relationships between dynamic structure and function. The accuracy of comparisons relies on precise, three-dimensional models, requiring continuously improving high-resolution experimental structures to enable capturing the nuanced features within activation landscapes. Improvements in cryo-EM will expand high-fidelity structural catalogues of physiologically relevant receptor states for precise analyses. The number and selection of geometric landmarks also requires careful consideration regarding biological relevance over perceived signal, to avoid overinterpreting tangential shape changes lacking functional correlates. Statistical rigor in testing observed associations is necessary to avoid sculpting models overfitted to a narrow landscape that fail to represent results observed in the laboratory. Interpreting multidimensional structural landscapes also benefits from visualisation approaches that intuitively communicate complex

morphological relationships. Finally, comprehensively integrating structure with functional outcomes remains limited by challenges in correlating dynamic shapes captured through *in silico* simulations, with relevant signalling measurements. Improving cellular assays with higher dimensionality and temporal precision promises more robust mapping of structure-function relationships by enabling precise quantification of signalling phenotypes for exact morphological states identified through modelling.

In summary, geometric morphometric methodologies enable the quantitative decoding of structural mechanisms directing functional plasticity in GPCR signalling. Comparative analyses of biased ligand-receptor shape spaces links conformational heterogeneity with downstream signalling effects, delivering dynamic models of biased signalling modulation. Integrating robust analyses of members of the GPCR superfamily, along with improved experimental biophysical techniques such as those detailed in Chapter 3, promises unprecedented structure-function resolution.

Chapter 5: Progress to date in the Further Exploration of Geometric Morphometrics and GPCR Structure

5.1 – Introduction

The novel application of geometric morphometrics to experimentally resolved GPCR structures, as detailed in Chapter 4, was met with success. These foundational analyses focussed upon global structural differences between inactive- and active-state receptor crystals, reliably mapping these functional states to receptor morphology. In addition, the subtle conformational heterogeneity which underpins the breadth of diverse activated states during signalling was also revealed, along with receptor morphologies stabilised by various classes of ligand during structure determination. Principal component analyses revealed shifts along specific structural modes which are consistent with activation transitions, and rigorous statistical analyses reliably determined significant differences between defined categories. Overall, the case study results presented in Chapter 4 provided consistent evidence for proof of concept, and suggest that the use of geometric morphometrics provides unique insight into the structural landscape of GPCRs. Consequently, while activation states are certainly a dominant aspect of variation in GPCR conformation, it was hypothesised that geometric morphometrics could be further applied to several other structural aspects of GPCR biology. By doing so, the morphological relationship between structure and function can be further explored, and its architecture quantitatively decoded, and in addition, further test the usefulness and multifaceted nature of geometric morphometrics.

The aim of this chapter therefore, was to explore the further applicability of geometric morphometrics to GPCRs, and determine whether any significant or interesting observations could be revealed from their structures, in terms of the determinants of G protein coupling, the possible consequences of thermostabilisation and the use of fusion proteins, and finally whether geometric morphometrics is useful in the comparison of AlphaFold models to experimentally resolved structures. Each of these concepts are introduced in more detail in sections 5.2 to 5.5.

5.2 – The Structural Determinants of G Protein Coupling Specificity

Historically, it was thought that each GPCR could couple to only one sub-type of G protein, eliciting a single cellular response, however, it is now known that this relationship is pleiotropic, and thus the selectivity or specificity of G protein coupling to GPCRs is now a key concept in the structure/function relationship of these receptor complexes (Masuho *et al.*, 2023). Just as GPCRs respond to a wide variety of ligands, transducing their signals across the cell membrane and into the cell, so do they subsequently couple to a variety of heterotrimeric G proteins, ultimately activating a variety of effector molecules and signalling pathways. The G proteins themselves are composed of three subunits, alpha, beta, and gamma, and the sixteen alpha subunits in the human genome can be classified into four sub-types – $G\alpha_s$, $G\alpha_{i/o}$, $G\alpha_{q/11}$, and $G\alpha_{12/13}$ (Kamato *et al.*, 2015). Each of the G proteins initiate signalling cascades differently, resulting in unique pharmacological profiles, possibly even over different timescales (Masuho *et al.*, 2015; Klein Herenbrink *et al.*, 2016); this necessitates both a pharmacological and temporal understanding of G protein activity, alongside physiological outcomes, which is currently not well understood overall.

The muscarinic acetylcholine receptor sub-family was one of the first explorations into preferential G protein coupling; each of the five members showed a similar degree of homology, but differed in their coupling preferences, and subsequent physiological responses (Caulfield and Birdsall, 1998). Eventually, the structures of both $G_{i/o}$-coupled M2 and $G_{q/11}$-coupled M3 muscarinic receptors offered insight into the structural basis of G protein coupling preference, and enabled a novel comparison between receptors coupled to these G proteins (Kruse *et al.*, 2012). Interestingly, it was found that the most significant difference was found at the intracellular end of transmembrane helix 5 and ICL2, in which a conserved tyrosine points towards the M2 receptor, but away from the M3 receptor. In support of this, mutagenesis approaches discovered a motif in both receptors at the intracellular end of transmembrane helix 6 which directly interacts with TM5, and is critical in determining G protein selectivity (Blin, Yun and Wess, 1995; Liu *et al.*, 1995). Even more interestingly, the equivalent motif in the β2AR-G_s complex makes direct contact with the terminal helix of $G\alpha_s$ (Søren G.F. Rasmussen *et al.*, 2011). These findings therefore reveal that the intracellular ends of TM5 and TM6 are particularly involved in the structural determination of G protein coupling.

Following on from these findings, Kruse *et al.*, 2012 discovered that by mapping the interhelical distances at the intracellular ends of the transmembrane helices, unique inactive GPCR structures formed distinct groupings which correlated with their G protein coupling specificity (Kruse *et al.*, 2012). As shown in Figure 5.2, measurements of distance in Angstroms between the Cα atoms of TM3 residue 3.54 and TM5 residue 5.62, and between TM5 residue 5.62 and

TM6 residue 6.37, reveal these groupings for 10 unique GPCR sub-families. The exception to these distinct groupings was an overlap, and therefore similarity, between $G_{i/o}$-coupled receptors (yellow) and G_t-coupled bovine rhodopsin (red). Consequently, aside from these interesting results which give insight into the structural determinants of G protein coupling specificity, several key concepts are here proven to be effective and translate directly into the geometric morphometrics method in this thesis. For example, although there is a difference between measuring interhelical distances, and the geometric landmarks for morphometric analyses, both are facilitated by the Cα atom of selected residues. In addition, focus was placed on the intracellular ends of the transmembrane helices, revealing groupings through differences between TMs 3, 5, and 6, which are also generally consistent sites of variation observed through geometric morphometrics.

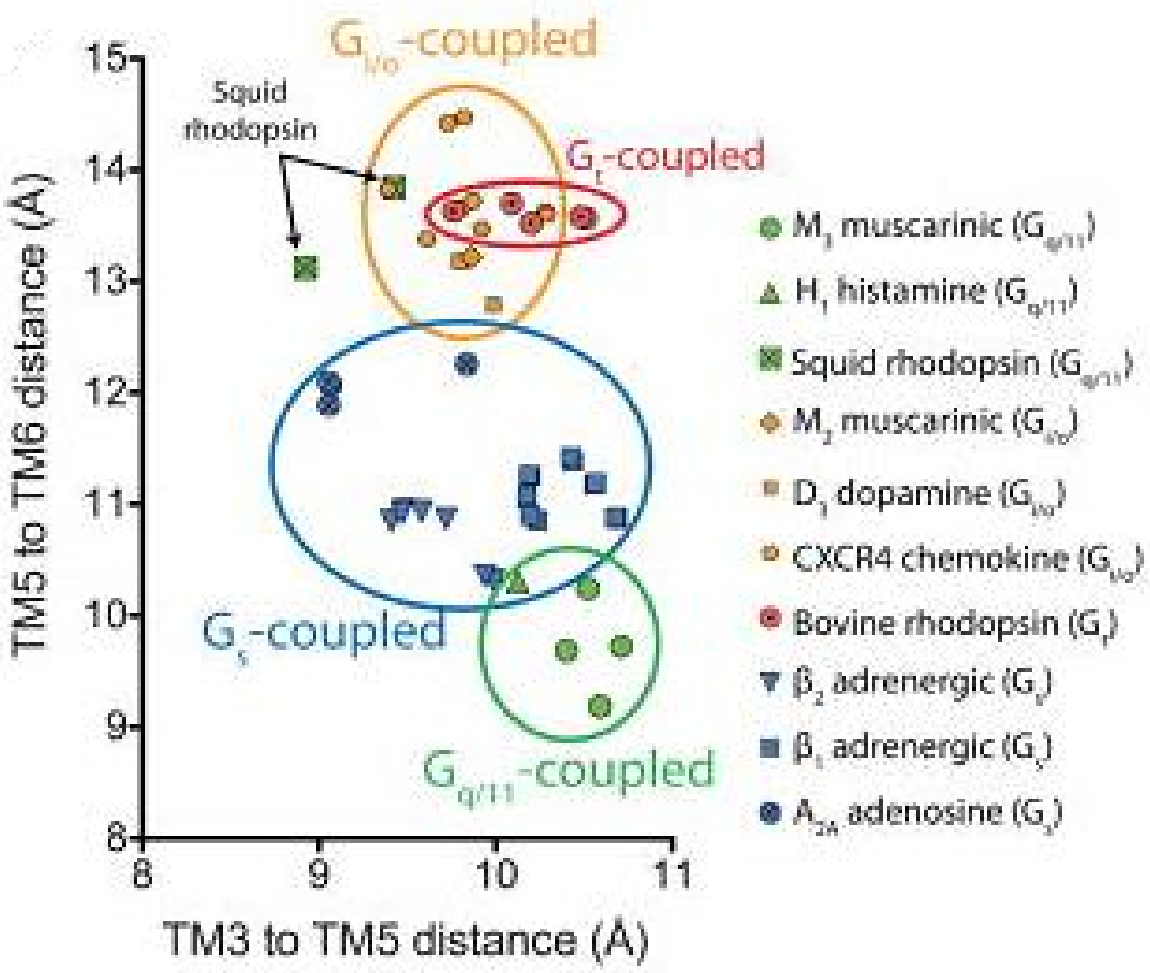

Figure 5.2. Interhelical Distances Reveal G Protein Coupling Specificity. Angstrom measurements between the Cα atoms of TM3 residue 3.54 and TM5 residue 5.62, and between TM5 residue 5.62 and TM6 residue 6.37 reveal groupings at the intracellular face of 10 GPCR sub-families. A2AR is blue circle, β1AR is blue square, β2AR is blue triangle, bovine rhodopsin is red circle, CXCR4 is yellow circle, D3R is yellow square, H1 is green triangle, M2 is yellow diamond, M3 is green circle, and squid rhodopsin is green square. In addition, receptors coupled to G_s are blue, $G_{q/11}$ are green, $G_{i/o}$ are yellow, and G_t are red. Taken from Kruse *et al*., 2012.

However, despite the interesting link between structure and G protein coupling specificity revealed by Kruse *et al*., 2012, their method did not take into account the effects of transmembrane helices 1, 2, 4, and 7 on groupings, and did not explore the extracellular face of the TM bundle at all. While TMs 3, 5 and 6 certainly represent dominant conformational changes associated with the receptor activation mechanism, there may well be more subtle conformational effects conferred by the other helices which result in functionally unique, or biased, states. Geometric morphometrics was therefore used in this body of work to explore this missing information, and to complement and advance the understanding of the structural determinants of G protein coupling specificity. Two experiments were performed; the first used the exact same structures explored by Kruse *et al*., 2012 to determine whether geometric

morphometrics could replicate their observations and reveal statistically significant differences between groups, taking all helices into account, at both extra- and intracellular ends. Secondly, this experiment was subsequently repeated to include more recently solved structures, for the 10 GPCR sub-families to determine whether these observations remained consistent or varied over time.

As such, the first experiment included 59 pre-2013 receptor-G protein structures from the A2AR, β1AR, β2AR, bovine rhodopsin, CXCR4, D3R, H1, M2, M3 and squid rhodopsin GPCR sub-families; 31 were G_s-coupled, 6 were $G_{q/11}$-, 7 were $G_{i/o}$-, and 15 were G_t-coupled. The second experiment included more recently determined structures from the same receptor sub-families, with 92 G_s-coupled, 11 $G_{q/11}$-, 13 $G_{i/o}$-, and 17 G_t-coupled. These structures were not filtered so as not to bias or limit the analysis in any way, and to capture the complete landscape of structure determined G protein coupling specificity across these 10 sub-families. Refer to Appendix B5 for a full list of structures, including PDB codes and reference to their publication. Of course additional receptors have been resolved in complex with G proteins not included in this body work, and could be further explored in future analyses.

PC Variance

The first two principal components accounted for 45% and 61%, of the variation in the extracellular and intracellular landmark coordinates, respectively, of the 59 receptors resolved in complex with a G protein, as shown in Table 5.2.1. The first principal component (PC1) accounted for 26% (extracellular) and 43% (intracellular) of the variation.

Table 5.2.1. Principal Component Variance of the Structures of Receptor-G Protein Complexes. The variance explained by the 14 principal components of 59 receptor-G protein structures, based on both the extra- and intracellular landmark coordinates. Variance is expressed as a percentage, alongside the total cumulative variance.

	Extracellular		Intracellular	
	%Variance	%Cumulative	%Variance	%Cumulative
PC1	25.843	25.843	43.427	43.427
PC2	19.101	44.944	17.179	60.606
PC3	17.496	62.44	10.376	70.981
PC4	9.236	71.676	7.704	78.685
PC5	6.533	78.209	4.98	83.666
PC6	4.487	82.696	3.833	87.499
PC7	4.015	86.711	2.992	90.49
PC8	3.603	90.314	2.384	92.875
PC9	2.654	92.968	1.894	94.768
PC10	2.271	95.24	1.665	96.434
PC11	1.843	97.083	1.263	97.697
PC12	1.291	98.374	1.013	98.71
PC13	0.873	99.246	0.744	99.455
PC14	0.754	100	0.545	100

PC1 and PC2 Scree Plots

Scree plots for the first two principal components, showing the location, direction and magnitude of variations are shown in Figures 5.2.1 and 5.2.2. For the extracellular landmarks, PC1 showed the greatest variation in the positions of transmembrane helices 1, 2, and 5, with some variation in the other helices. PC2 showed the greatest variation in helices 1, 3, and 7, and some variation in helices 4 and 5. For the intracellular landmarks, PC1 showed the greatest variation in helices 3, 5 and 6, with some variation in helix 1, 2, and 7. PC2 showed the greatest variation in helices 5 and 6, and some variation in helix 7.

Overall, these Scree plots have again highlighted variation in transmembrane helices 5 and 6, and is especially strong in the intracellular morphospace. Given that these helices form an essential mechanism in the classical activation of GPCRs, opening an intracellular cleft by conformational change which facilitates G protein coupling, these results are therefore consistent with this paradigm. In addition, variation was also highlighted in helices 1, 3, and 7 in particular, which likely represents the conformational diversity between these receptor-G protein complexes. Indeed, just as the structural heterogeneity was shown to be significantly affected by the class of ligand bound to receptor, as explored in Chapter 4, it may well be that this principle remains consistent with the link between receptor structures and their specificity of G protein coupling.

Figure 5.2.1. Scree Plots for both PC1 and PC2 of the Extracellular Receptor-G Protein Complex Landmarks. Scree plots for PC1 (a) and PC2 (b) are shown for the 7 extracellular landmarks of the 59 receptor structures in complex with a G protein.

Figure 5.2.2. Scree Plots for both PC1 and PC2 of the Intracellular Receptor-G Protein Complex Landmarks. Scree plots for PC1 (a) and PC2 (b) are shown for the 7 intracellular landmarks of the 59 receptor structures in complex with a G protein.

PC1 vs PC2 Morphospace

Scatterplots comparing PC1 and PC2 show a spread of data across both the extracellular (Figure 5.2.3a) and intracellular (Figure 5.2.3b) principal components, with evidence of some clustering and general variation between both PC1 and PC2. The intracellular landmark coordinates (b) revealed several smaller clusters around the 0.1-0.2 PC1, -0.00 PC2 morphospace, but no obviously distinct clusters were observed. The extracellular landmarks (a) revealed a cluster of points around the -0.10-0.15 PC1, -0.05 -0.1 PC2 morphospace, and the remaining data generally distributed between PC1 and PC2. Importantly, it should be noted that the PC axes for the intracellular data (b) were greater than the range of the extracellular

PC axes (a), reflecting a greater degree of variance in the intracellular landmarks, additionally highlighted by the greater cumulative percentage variation in Table 5.2.1.

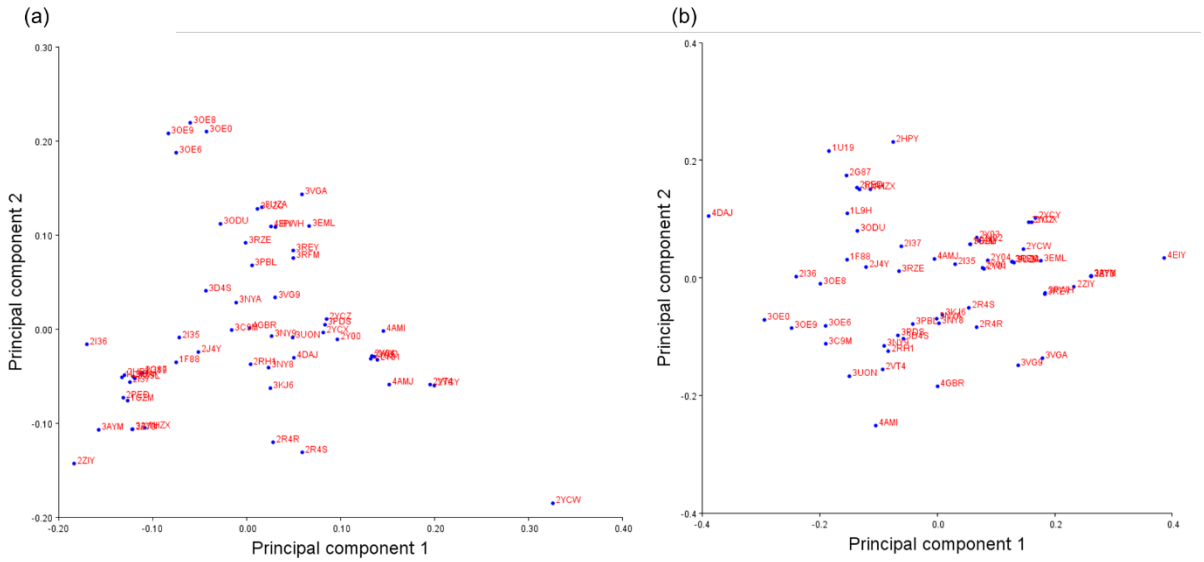

Figure 5.2.3. Morphospace of the Structures of Receptor-G Protein Complexes. Position in morphospace of 59 receptor-G protein complexes along PC1 and PC2, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Each structure is labelled with its PDB code.

For the receptors resolved in complex with a G protein then, it was hypothesised that the structural determinants of G protein coupling specificity might be detected with geometric morphometric analyses. As was suggested in Chapter 4, within the conformational landscape captured by shape morphology, the mapping of significant structural variation to ligand bound or activation states is consistent with the concept of the structure-based signalling of GPCRs. These results indicate that geometric morphometrics is also able to map these subtle conformational differences to G protein-bound states, linking the structural determination of biologically relevant states between both the extracellular and intracellular side of the cell membrane, from ligand binding to conformational change to G protein coupling.

As shown in Figure 5.2.4, both the extra- and intracellular morphospaces reveal fairly distinct clusters of categorised structures, with some overlap, though statistical testing revealed these differences to be very significant. By categorisation of receptor sub-family, the adenosine A2A receptor structures are blue circle, β1-adrenergic are blue square, β2-adrenergic are blue triangle, bovine rhodopsin are red circle, the CXCR4 receptor are yellow circle, the dopamine receptor D3 are yellow square, histamine H1 receptor are green triangle, M2 muscarinic are yellow diamond, M3 muscarinic are green circle, and squid rhodopsin are green square. In addition to this, by categorisation of the G protein sub-types these structures were resolved in complex with, G_s is blue, $G_{q/11}$ is green, $G_{i/o}$ is yellow, and G_t is red.

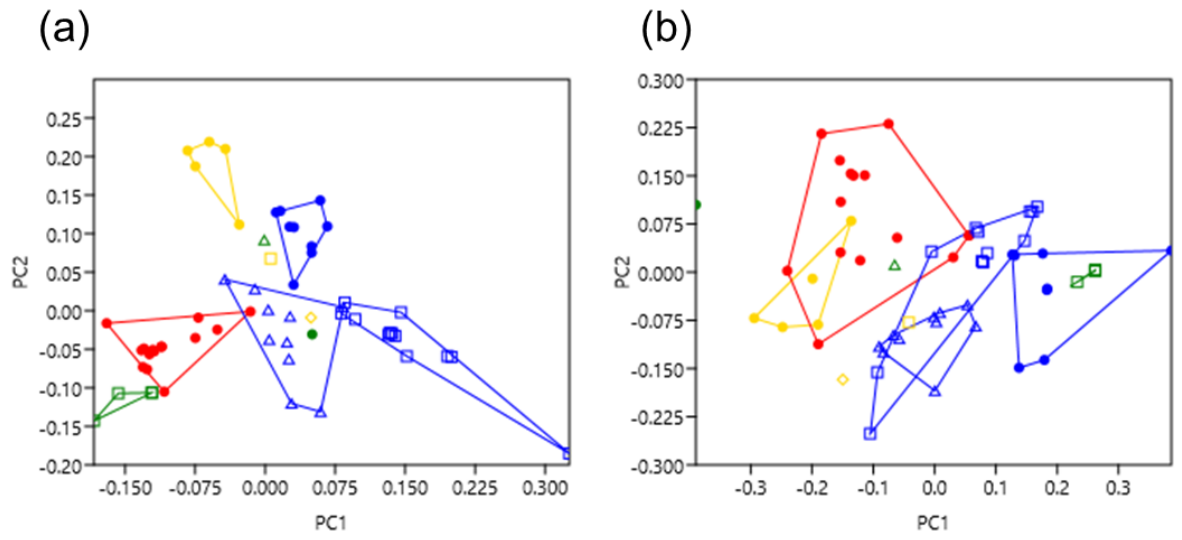

Figure 5.2.4. Morphospace of the Structures of Receptor-G Protein Complexes (Pre-2013). Position in morphospace of 59 receptor-G protein complexes, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Convex hulls delimit the structures as categorised, indicating group affiliation. A2AR is blue circle, β1AR is blue square, β2AR is blue triangle, bovine rhodopsin is red circle, CXCR4 is yellow circle, D3R is yellow square, H1 is green triangle, M2 is yellow diamond, M3 is green circle, and squid rhodopsin is green square. In addition, receptors coupled to G_s are blue, $G_{q/11}$ are green, $G_{i/o}$ are yellow, and G_t are red. These are the same structures explored by Kruse *et al.*, 2012, shown in Figure 5.2.

PERMANOVA (extracellular F = 20.64, p = 0.0001; intracellular F = 11.16, p = 0.0001) and ANOSIM (extracellular R = 0.529, p = 0.0001; intracellular R = 0.4649, p = 0.0001) showed that there were very significant differences between most categorised groups, summarised in Table 5.2.2. For the extracellular morphospace, both PERMANOVA and ANOSIM showed significant differences between all but one comparison of G protein-coupled structures; PERMANOVA did not find a significant difference between $G_{q/11}$- and G_t- coupled receptors. For the intracellular morphospace, PERMANOVA revealed significant variance between G_s- compared to $G_{i/o}$-, and G_s- compared to G_t- coupled structures. $G_{q/11}$- compared to both G_s- and $G_{i/o}$- were not significant. In contrast, ANOSIM showed all comparisons were significant, indicating a strong degree of clustering between receptors in complex with each sub-type of G protein. Overall, both PERMANOVA and ANOSIM for both the extra- and intracellular morphospaces were considered very significant with p values of 0.0001 across the dataset.

Table 5.2.2. Statistical Analyses of the Structures of Receptor-G Protein Complexes (Pre-2013). Pairwise comparisons of PC1 and PC2 scores of both the extra- and intracellular landmarks of the receptor-G protein complexes, in terms of their structure determination in complex with a particular G protein sub-type. Structures were categorised as either being in complex with G_s, $G_{i/o}$, $G_{q/11}$ or G_t, based on information gained from the *GPCRdb*. PERMANOVA *p* values are above the diagonal, and ANOSIM *p* values are below the diagonal. The Bonferroni sequential corrected significant *p* values are indicated in bold.

Extracellular					Intracellular				
	G_s	$G_{i/o}$	$G_{q/11}$	G_t		G_s	$G_{i/o}$	$G_{q/11}$	G_t
G_s		**0.0001**	**0.0001**	**0.0001**	G_s		**0.0001**	0.5406	**0.0001**
$G_{i/o}$	**0.0027**		**0.0045**	**0.0001**	$G_{i/o}$	**0.0001**		0.0271	**0.0007**
$G_{q/11}$	**0.001**	**0.0041**		0.5565	$G_{q/11}$	**0.0087**	**0.0045**		**0.0138**
G_t	**0.0001**	**0.0001**	**0.0027**		G_t	**0.0001**	**0.0036**	**0.0002**	

This geometric morphometric analysis of receptor-G protein complex structure was subsequently repeated, to build upon Figure 5.2.4, and the work of Kruse *et al.*, 2012, encompassing more recently resolved structures in addition to those solved pre-2013. This expanded the analysed dataset from 59 samples to 133, taken from the exact same receptor sub-families to maintain consistency across experiments. It was therefore hypothesised that either a greater degree of variation would be observed, reflecting an even more diverse conformational landscape across G protein-bound states, or, the observations revealed in this experiment might remain consistent with the pre-2013 data, highlighting the strength and reliability of geometric morphometrics. As shown in Figure 5.2.5 and Table 5.2.3, the latter hypothesis was supported, indicating a general consistency across experiments despite more than doubling the dataset, further highlighting the multifaceted nature of this approach.

Just as before, the structures categorised in Figure 5.2.5 are as follows: by categorisation of receptor sub-family, the adenosine A2A receptor structures are blue circle, β1-adrenergic are blue square, β2-adrenergic are blue triangle, bovine rhodopsin are red circle, the CXCR4 receptor are yellow circle, the dopamine receptor D3 are yellow square, histamine H1 receptor are green triangle, M2 muscarinic are yellow diamond, M3 muscarinic are green circle, and squid rhodopsin are green square. In addition to this, by categorisation of the G protein sub-types these structures were resolved in complex with, G_s is blue, $G_{q/11}$ is green, $G_{i/o}$ is yellow, and G_t is red.

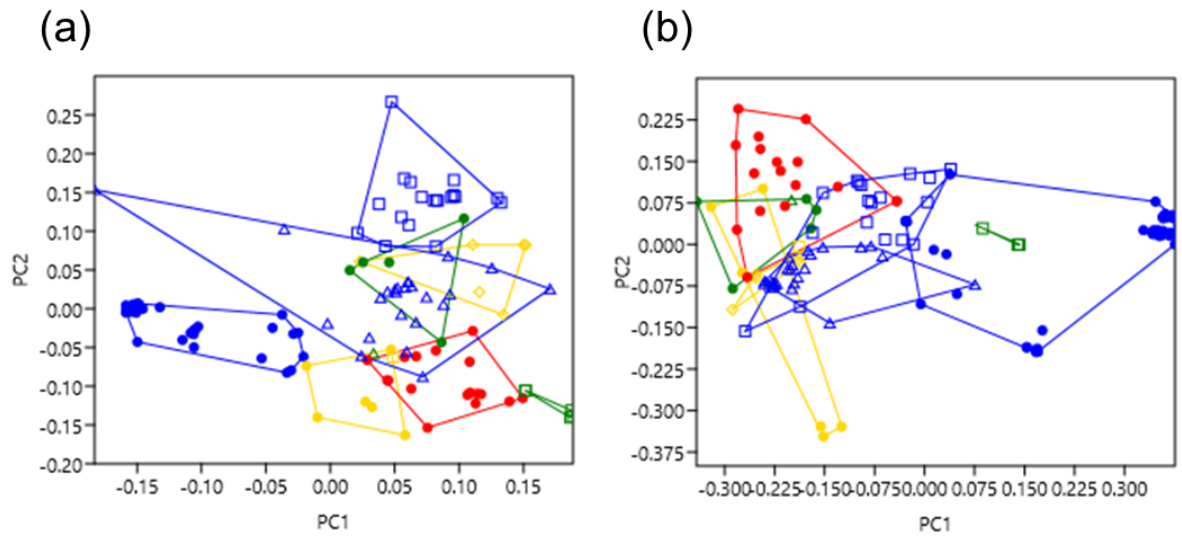

Figure 5.2.5. Morphospace of the Structures of Receptor-G Protein Complexes (Updated). Position in morphospace of 133 receptor-G protein complexes, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Convex hulls delimit the structures as categorised, indicating group affiliation. A2AR is blue circle, β1AR is blue square, β2AR is blue triangle, bovine rhodopsin is red circle, CXCR4 is yellow circle, D3R is yellow square, H1 is green triangle, M2 is yellow diamond, M3 is green circle, and squid rhodopsin is green square. In addition, receptors coupled to G_s are blue, $G_{q/11}$ are green, $G_{i/o}$ are yellow, and G_t are red. This encompasses more recently solved structures, building on Figure 5.2.4 and the work of Kruse *et al.*, 2012.

PERMANOVA (extracellular F = 17.03, p = 0.0001; intracellular F = 15.38, p = 0.0001) and ANOSIM (extracellular R = 0.2549, p = 0.0001; intracellular R = 0.1347, p = 0.0005) showed that there were very significant differences between most categorised groups, summarised in Table 5.2.3. For the extracellular morphospace, PERMANOVA showed significant differences between all but two comparisons of G protein-coupled structures; PERMANOVA did not find a significant difference between $G_{q/11}$- compared to G_t-, and $G_{i/o}$-coupled receptors, which remains partially consistent with the pre-2013 structures in Table 5.2.2. ANOSIM revealed significant differences between comparisons aside from $G_{q/11}$- and $G_{i/o}$-coupled structures which differs from the pre-2013 analysis. For the intracellular morphospace, both PERMANOVA and ANOSIM revealed significant differences between all groups, aside from G_s- compared to $G_{q/11}$-. Overall, both PERMANOVA and ANOSIM for both the extra- and intracellular morphospaces were again considered very significant across the dataset, remaining consistent with the pre-2013 analysis, despite using over twice as many G protein-coupled structures.

Table 5.2.3. Statistical Analyses of the Structures of Receptor-G Protein Complexes (Updated). Pairwise comparisons of PC1 and PC2 scores of both the extra- and intracellular landmarks of the receptor-G protein complexes, in terms of their structure determination in complex with a particular G protein sub-type. Structures were categorised as either being in complex with G_s, $G_{i/o}$, $G_{q/11}$ or G_t, based on information gained from the *GPCRdb*. PERMANOVA *p* values are above the diagonal, and ANOSIM *p* values are below the diagonal. The Bonferroni sequential corrected significant *p* values are indicated in bold.

Extracellular					Intracellular				
	G_s	$G_{i/o}$	$G_{q/11}$	G_t		G_s	$G_{i/o}$	$G_{q/11}$	G_t
G_s		**0.0004**	**0.0001**	**0.0001**	G_s		**0.0001**	0.0536	**0.0001**
$G_{i/o}$	**0.0004**		0.4603	**0.0158**	$G_{i/o}$	**0.0001**		**0.0021**	**0.0001**
$G_{q/11}$	**0.0002**	0.1152		0.0636	$G_{q/11}$	0.5793	**0.0005**		**0.0018**
G_t	**0.0001**	**0.0001**	**0.0002**		G_t	**0.0001**	**0.0001**	**0.0015**	

5.3 – The Structural Consequences of Thermostabilising Mutations

One of the major challenges of working with membrane proteins is their lack of stability, especially when solubilised away from the native membrane (Errasti-Murugarren, Bartoccioni and Palacín, 2021); it is common practice to retain protein samples on ice to prevent their degradation during expression, purification, and experimentation (Wiseman *et al.*, 2020). This can be particularly problematic for GPCRs during the processes required to determine their structure and pharmacological characterisation, and thus strategies to overcome this limiting step can be employed. One such example is thermostabilisation by systematic or random mutagenesis approaches to improve receptor solubility and stability at higher temperatures, for example (Vaidehi, Grisshammer and Tate, 2016). The compromise of this benefit is that the target protein becomes more rigid, exhibits a global decrease in mobility, but results in a more stable enthalpy. Despite these potential drawbacks, thermostabilisation has been shown to not affect receptor pharmacology in some cases, and has even been instrumental in the creation of stabilised receptor (StaR®) constructs by Sosei Heptares, which have directly led to 25 pre-clinical and 10 clinical candidates (Robertson *et al.*, 2011).

However, despite these successes, thermostabilising mutagenesis may indeed affect the conformational dynamics of a GPCR, leading to an altered pharmacological profile or loss of signalling capabilities, as demonstrated with the neurotensin 1 receptor (Shibata *et al.*, 2009). Four point mutations were used to successfully increase the stability of the neurotensin 1 receptor which is notoriously less stable than the A2AR, β1AR and β2AR receptors, though two of the locations were found to be essential for ligand binding and activation. It was found that agonist affinity was comparable to wild-type, antagonist affinity was reduced, but importantly the thermostabilised neurotensin 1 receptor activated G proteins poorly (Shibata

et al., 2009). Suffice to say, it is imperative that thermostabilising mutations are fully understood and considered, whether significant structural changes occur and as a result, altered GPCR pharmacology; it may well be that the thermostabilisation of membrane proteins can be 'safe' to conduct, and its assessment is necessitated on a case-by-case basis.

Here, geometric morphometrics was used to broadly compare the differences between thermostabilised receptors and non-thermostabilised equivalents from the same sub-family. It was hypothesised that the receptors thermostabilised by systematic mutagenesis might differ in their global receptor morphology, when compared to non-mutated equivalents; this was not shown to be the case. At the time of sample selection, 124 receptor structures were included in the dataset, with 60 thermostabilised and 64 non-thermostabilised, from the adenosine A2A, β1-adrenergic and β2-adrenergic receptor sub-families of the family A GPCRs. These groups of receptors were selected as they have historically been well-studied, were involved in the foundational geometric morphometric analyses detailed in Chapter 4, and included ample numbers of both thermostabilised and non-thermostabilised receptors for comparison. These structures were not filtered so as not to bias or limit the analysis in any way, and to capture the complete landscape of structural variation between thermostabilised and non-thermostabilised GPCRs. Refer to Appendix B6 for a full list of structures, including PDB codes and reference to their publication.

PC Variance

The first two principal components accounted for 57% and 73%, of the variation in the extracellular and intracellular landmark coordinates, respectively, of the 124 thermostabilised and non-thermostabilised receptor structures, as shown in Table 5.3.1. The first principal component (PC1) accounted for 39% (extracellular) and 59% (intracellular) of the variation.

Table 5.3.1. Principal Component Variance of Thermostabilised and Non-Thermostabilised Receptor Structures. The variance explained by the 14 principal components of 124 thermostabilised and non-thermostabilised receptor structures, based on both the extra- and intracellular landmark coordinates. Variance is expressed as a percentage, alongside the total cumulative variance.

	Extracellular		Intracellular	
	%Variance	%Cumulative	%Variance	%Cumulative
PC1	38.727	38.727	58.746	58.746
PC2	17.735	56.462	13.816	72.562
PC3	10.156	66.618	10.511	83.072
PC4	9.564	76.182	3.284	86.357
PC5	4.378	80.56	3.269	89.625
PC6	3.804	84.364	2.498	92.124
PC7	3.631	87.996	1.756	93.88
PC8	3.045	91.041	1.661	95.541
PC9	2.193	93.234	1.195	96.736
PC10	2.135	95.37	1.038	97.774
PC11	1.662	97.032	0.883	98.657
PC12	1.462	98.494	0.533	99.19
PC13	0.929	99.422	0.48	99.67
PC14	0.578	100	0.33	100

PC1 and PC2 Scree Plots

Scree plots for the first two principal components, showing the location, direction and magnitude of variations are shown in Figures 5.3.1 and 5.3.2. For the extracellular landmarks, PC1 showed the greatest variation in the positions of transmembrane helices 4, 6 and 7, with some variation in helices 1 and 2. PC2 showed the some variation in helices 1 and 7, and very little variation otherwise. For the intracellular landmarks, PC1 showed the greatest variation in helices 3 and 6, with some variation in helix 1, 5, and 7. PC2 showed the greatest variation in helices 5 and 6, and some variation in helices 4 and 7.

Overall, PC1 for both the extra- and intracellular landmarks again highlighted variation in transmembrane helix 6, and some variation in most other helices across both PC1 and PC2. These therefore suggest that there is a general degree of variation which is not helix specific, when making comparisons between receptor structures which either have or have not been thermostabilised by mutagenesis approaches. Given that the morphological and statistical analyses for this comparison did not find any significant differences, it is likely that the variation observed here is a collective effect of being a relatively larger dataset with receptors from several sub-families, with a mixture of activation states and stabilisation with different classes of ligand. Even though this particular comparison is focussed upon thermostabilising mutations, it may well be that the highlighted variation in TM6 is underpinned by differences between inactive- and active-state structures. This observation could therefore be further

refined to focus on one or the other, limiting the effect of activation state on this analysis, for example. These results are also therefore consistent with the findings of Chapter 4, where activation state was found to be the most dominant aspect of variation, and more subtle conformational differences may be found in lower ranked principal components. In terms of thermostabilisation, the lower principal components were analysed but did not reveal any significant differences.

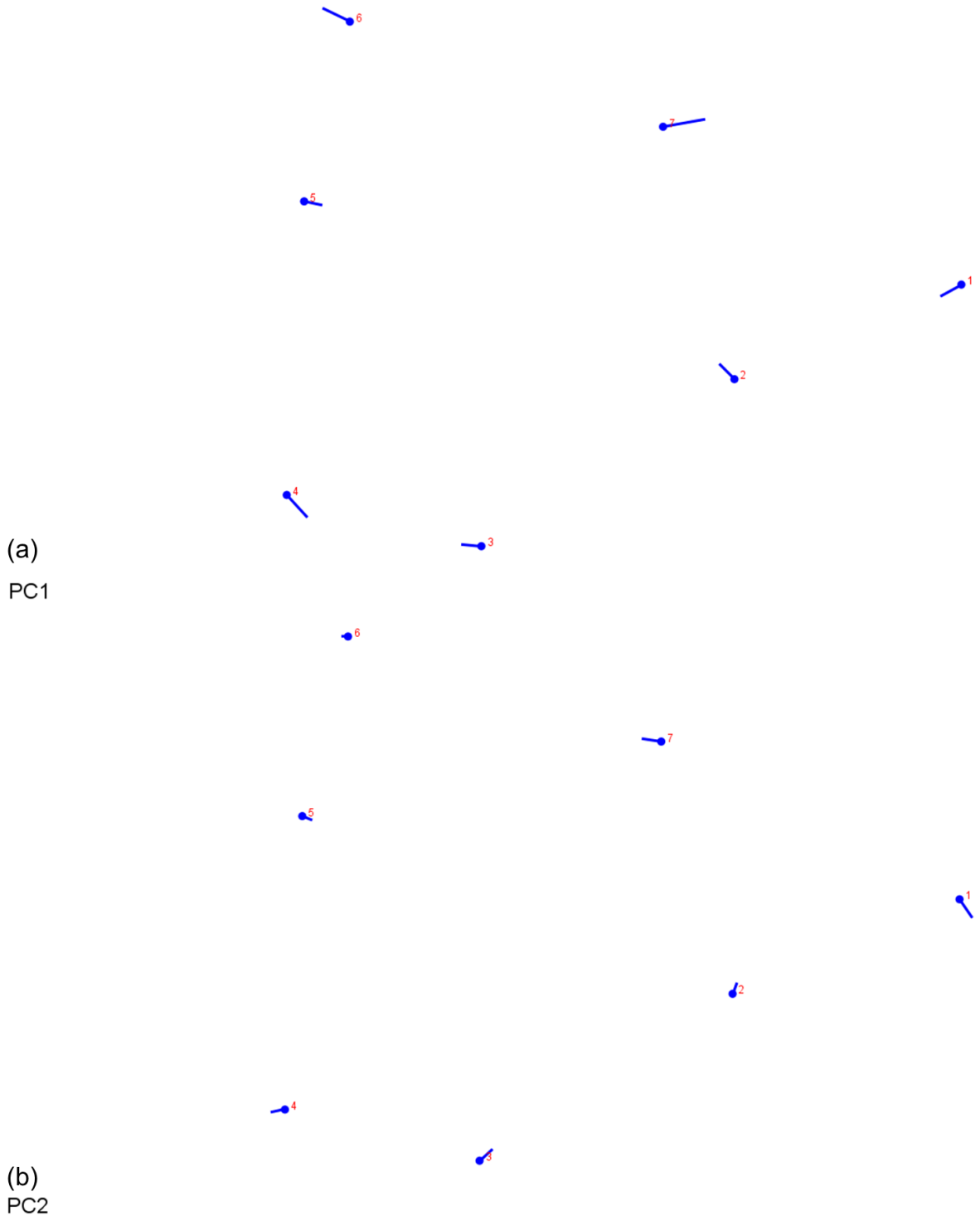

Figure 5.3.1. Scree Plots for both PC1 and PC2 of the Extracellular Thermostabilised and Non-Thermostabilised Receptor Landmarks. Scree plots for PC1 (a) and PC2 (b) are shown for the 7 extracellular landmarks of the 124 thermostabilised and non-thermostabilised receptor structures.

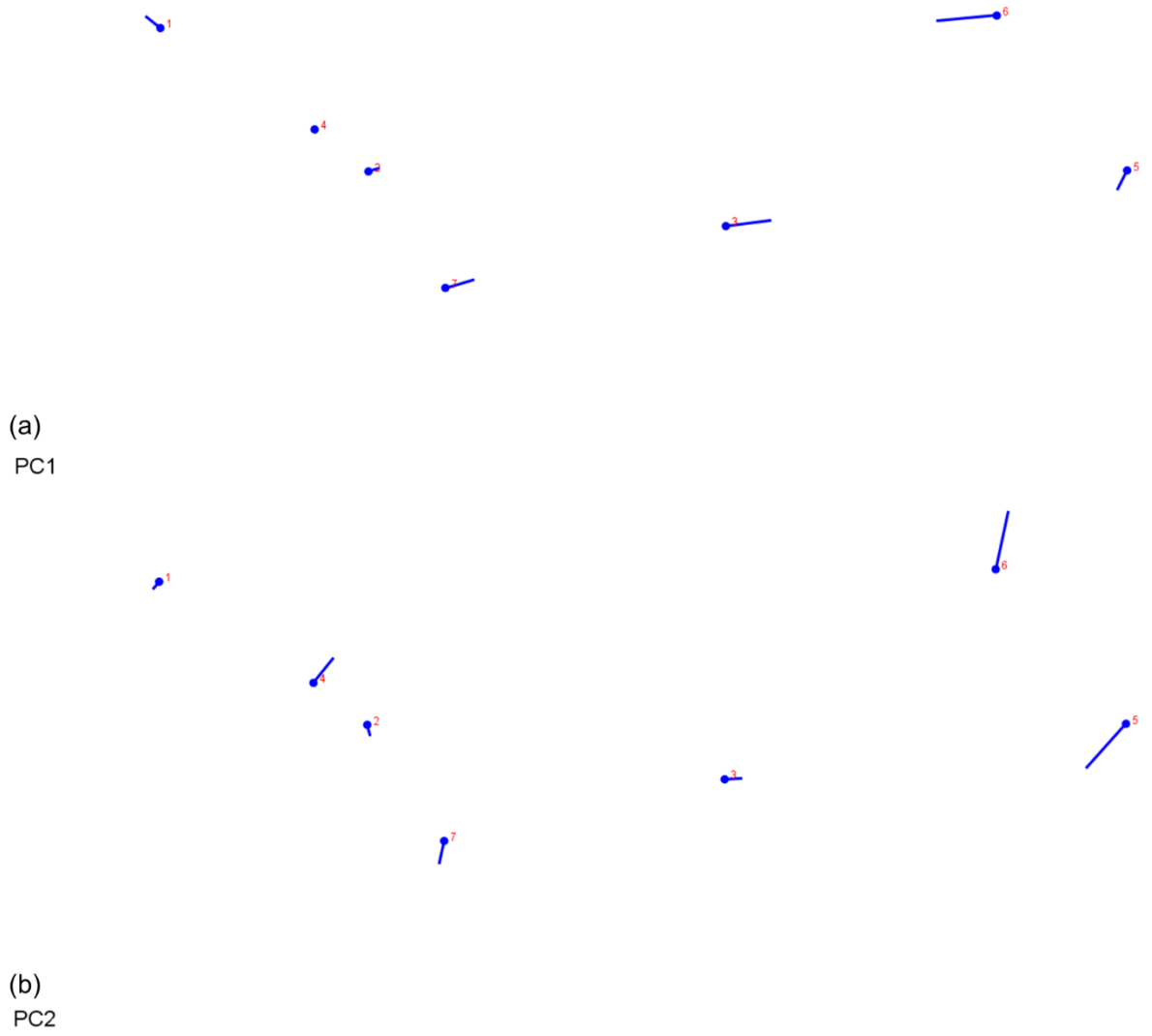

Figure 5.3.2. Scree Plots for both PC1 and PC2 of the Intracellular Thermostabilised and Non-Thermostabilised Receptor Landmarks. Scree plots for PC1 (a) and PC2 (b) are shown for the 7 intracellular landmarks of the 124 thermostabilised and non-thermostabilised receptor structures.

PC1 vs PC2 Morphospace

Scatterplots comparing PC1 and PC2 show a spread of data across both the extracellular (Figure 5.3.3a) and intracellular (Figure 5.3.3b) principal components, with evidence of some strong clustering and general variation. The intracellular landmark coordinates (b) revealed a main cluster of points around the -0.30 PC1, -0.00 PC2 morphospace, and the remainder of points varying generally towards the positive range of PC1, with a few more smaller clusters. The extracellular landmarks (a) revealed a main cluster of points around the 0.10-0.15 PC1, 0.05 PC2 morphospace, and the remaining data generally distributed between PC1 and PC2. Importantly, it should be noted that the PC axes for the extracellular data (a) were half the range of the intracellular PC axes (b), reflecting a greater degree of variance in the intracellular

landmarks, additionally highlighted by the greater cumulative percentage variation in Table 5.3.1.

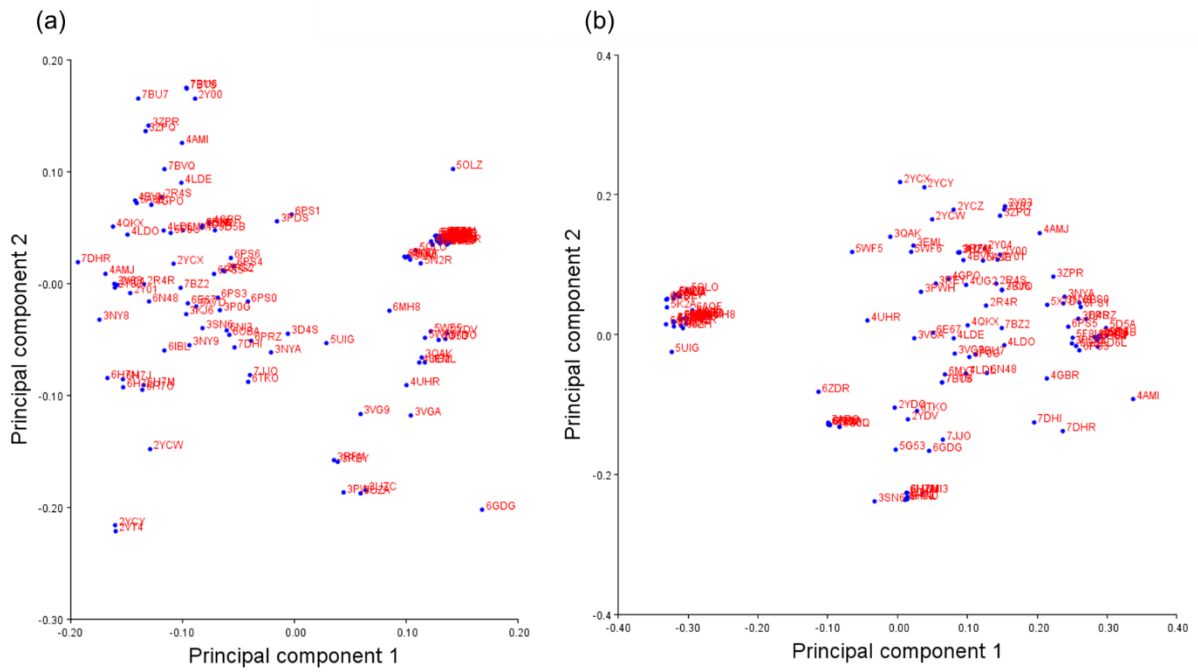

Figure 5.3.3. Morphospace of the Thermostabilised and Non-Thermostabilised Receptor Structures. Position in morphospace of 124 thermostabilised and non-thermostabilised receptor structures along PC1 and PC2, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Each structure is labelled with its PDB code.

For the thermostabilised and non-thermostabilised GPCR structures then, it was hypothesised that the presence of these mutations may contribute to a significant difference to global receptor morphology when compared to non-mutated equivalents. As shown in Figure 5.3.4, this was not the case, with both the extra- and intracellular morphospaces revealing a general overlap between categorised structures, and no significant differences between groups. The thermostabilised structures are red, and non-thermostabilised are blue, based on information gathered from the *GPCRdb*. In essence, while this analysis was not shown to be significant in terms of global morphology, mutations may well subtly affect the chemical compatibility needed for efficacious ligand binding and G protein coupling, especially with regards to the more subtle conformational landscapes in biased signalling. It is recommended that researchers should not become ignorant of the potential effects of thermostabilising mutations, and make these considerations during experimental design stages. Given that thermostabilising mutations are known to affect the pharmacology of certain receptors, it would be sensible to investigate their effects on a case-by-case basis in combination with other comparative techniques such as RMSD, Q score and functional assays, for example.

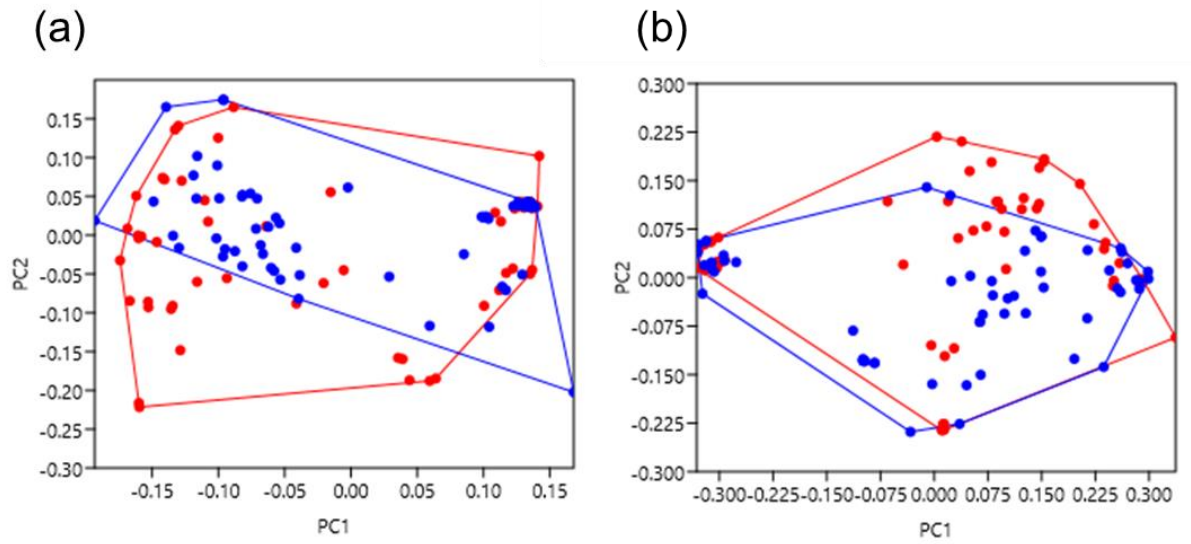

Figure 5.3.4. Categorised Morphospace of the Thermostabilised and Non-Thermostabilised Receptor Structures. Position in morphospace of 124 thermostabilised and non-thermostabilised receptor structures, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Convex hulls delimit the structures as categorised, indicating group affiliation. Thermostabilised is red, non-thermostabilised is blue.

PERMANOVA (extracellular F = 1.628, p = 0.1961; intracellular F = 2.591, p = 0.085) and ANOSIM (extracellular R = 0.02502, p = 0.0439; intracellular R = 0.03367, p = 0.0271) showed that there were no significant differences between the categorised groups, summarised in Table 5.3.2. For the extracellular morphospace, both PERMANOVA and ANOSIM showed no significant differences between the thermostabilised and non-thermostabilised receptor structures, with an extremely low R value indicating almost complete similarity between groups. For the intracellular morphospace, PERMANOVA was not significant, however, due to the groupings observed in Figure 5.3.4b, ANOSIM did indicate that this clustering was significant. Despite this, the extremely low R value again indicates almost complete similarity between the thermostabilised and non-thermostabilised receptor structures overall. The PERMANOVA F values for both morphospaces were both low, indicating a lack of variation between groups.

Table 5.3.2. Statistical Analyses of the Thermostabilised and Non-Thermostabilised Receptor Structures. Pairwise comparisons of PC1 and PC2 scores of both the extra- and intracellular landmarks of the thermostabilised and non-thermostabilised receptor structures, in terms of reported stabilising mutations. Structures were categorised as either thermostabilised (thermo) or non-thermostabilised (non-thermo), based on information gained from the *GPCRdb*. PERMANOVA *p* values are above the diagonal, and ANOSIM *p* values are below the diagonal. The Bonferroni sequential corrected significant *p* values are indicated in bold.

Extracellular			Intracellular		
	Thermo	Non-thermo		Thermo	Non-thermo
Thermo		0.1964	Thermo		0.0921
Non-thermo	0.0458		Non-thermo	**0.025**	

5.4 – The Structural Consequences of Fusion Proteins

In a similar manner to thermostabilising mutations, the use of fusion proteins can enhance the expression and purification of membrane proteins; fusion proteins can also aid in crystallisation and structure determination efforts, although not as frequently. Several strategies exist including fusion at either the N or C terminus, insertion into a helix or loop, or by terminal restraint, ultimately using this soluble fusion protein to form relatively large and stable scaffold for crystal contact (Liu and Li, 2022). The most commonly adopted method would be by terminal fusion, which was used in the seminal studies of the β2AR structure (Zou, Weis and Kobilka, 2012), and further refinement of the construct led to the determination of the β2AR-G_s complex (Rasmussen *et al.*, 2007), and the identification of carazolol which, as discussed in Chapter 4, highlights the concept of biased signalling (Cherezov *et al.*, 2007). In addition to terminal fusion, the insertion of the scaffold between helices has also been used in the crystallisation of GPCRs. T4 lysozyme was first inserted in between transmembrane helices 5 and 6 of β2AR, essentially replacing ICL3 – it was hypothesised that as ICL3 is highly flexible, this insertion would not significantly disturb the rest of the structure (Cherezov *et al.*, 2007). This approach was not only successful in its application to structure determination, but has been widely adopted to the resolution of many other GPCRs (Liu and Li, 2022).

One interesting example is the fusion insertion of a glycogen synthase from *Pyrococcus abysii* into ICL3 of the orexin receptors, as T4 lysozyme had failed (Yin *et al.*, 2015). The orexin, or hypocretin, receptors are responsive to cerebral neuropeptides which regulate sleep and are involved in dysregulation such as cataplexy and narcolepsy; indeed, the inhibition of the orexin receptors with the antagonist suvorexant, is an effective therapeutic for insomnia (Yin *et al.*, 2015). The use of fusion proteins then has been instrumental in the determination of GPCR structure, enhancing the solubility and stability required for these techniques. However, as

with the thermostabilising mutations, the inclusion of a fusion scaffold may have an impact on receptor structure or pharmacology. Although historically it has been argued that insertion into ICL3 shouldn't cause a significant effect, the recent understanding of structure-based biased signalling may prove otherwise; it may be that any structural effects of fusion insertion may be subtle enough that they do not impact major pathways, but instead may affect the more minor pathways involved in bias. If this is the case, the link between determined structures and the development of novel therapeutics should be carefully monitored.

As such, it was hypothesised that the insertion of GST into ICL3 of the orexin receptors, between transmembrane helices 5 and 6, might cause a subtle yet significant difference, given the involvement of these helices in conformational changes. Indeed, geometric morphometrics has thus far reliably revealed conformational heterogeneity under a variety of categorisations, consistently detecting variation in TM5 and TM6. At the time of sample selection, 22 orexin receptor structures were included in the dataset, with 9 GST-tagged and 13 untagged; the orexin receptors were selected as they were represented by a mixture of both tagged and untagged structures, and to gain further insight into their applications to structure-based drug discovery (Hellmann *et al.*, 2020). These structures were not filtered so as not to bias or limit the analysis in any way, and to capture any conformational differences between orexin receptors which contained the GST fusion protein, and those that did not. Refer to Appendix B7 for a full list of structures, including PDB codes and reference to their publication.

PC Variance

The first two principal components accounted for 74% and 88%, of the variation in the extracellular and intracellular landmark coordinates, respectively, of the 22 GST-tagged and untagged orexin receptor structures, as shown in Table 5.4.1. The first principal component (PC1) accounted for 57% (extracellular) and 67% (intracellular) of the variation.

Table 5.4.1. Principal Component Variance of GST-Tagged and Untagged Orexin Receptor Structures. The variance explained by the 14 principal components of 22 GST-tagged and untagged orexin receptor structures, based on both the extra- and intracellular landmark coordinates. Variance is expressed as a percentage, alongside the total cumulative variance.

	Extracellular		Intracellular	
	%Variance	%Cumulative	%Variance	%Cumulative
PC1	56.52	56.52	67.311	67.311
PC2	17.746	74.266	21.1	88.411
PC3	6.84	81.106	4.008	92.419
PC4	5.822	86.928	2.303	94.721
PC5	4.015	90.944	1.889	96.611
PC6	2.935	93.879	1.358	97.969
PC7	2.72	96.599	0.874	98.843
PC8	1.821	98.42	0.667	99.51
PC9	0.635	99.056	0.287	99.798
PC10	0.412	99.467	0.084	99.882
PC11	0.29	99.757	0.066	99.948
PC12	0.19	99.947	0.03	99.977
PC13	0.042	99.989	0.015	99.992
PC14	0.011	100	0.008	100

PC1 and PC2 Scree Plots

Scree plots for the first two principal components, showing the location, direction and magnitude of variations are shown in Figures 5.4.1 and 5.4.2. For the extracellular landmarks, PC1 showed the greatest variation in the positions of transmembrane helix 1, with some variation in helices 2, 4, 5, and 7. PC2 showed the greatest variation in helices 1 and 5, and some variation in helices 2, 3, 6, and 7. For the intracellular landmarks, PC1 showed the greatest variation in helix 6, with some variation in helices 1, 3, 5, and 7. PC2 showed the greatest variation in helix 6, and some variation in the remaining helices.

Overall, while the extracellular Scree plots did exhibit landmark variation, mainly in TM1, the intracellular results shown in Figure 5.4.2 are more interesting. Given that the GST fusion protein is inserted into ICL3 between the intracellular ends of transmembrane helices 5 and 6, one might expect any significant variation to be focussed at this location, and this is exactly shown to be the case when comparing the orexin receptors both with and without this GST fusion protein.

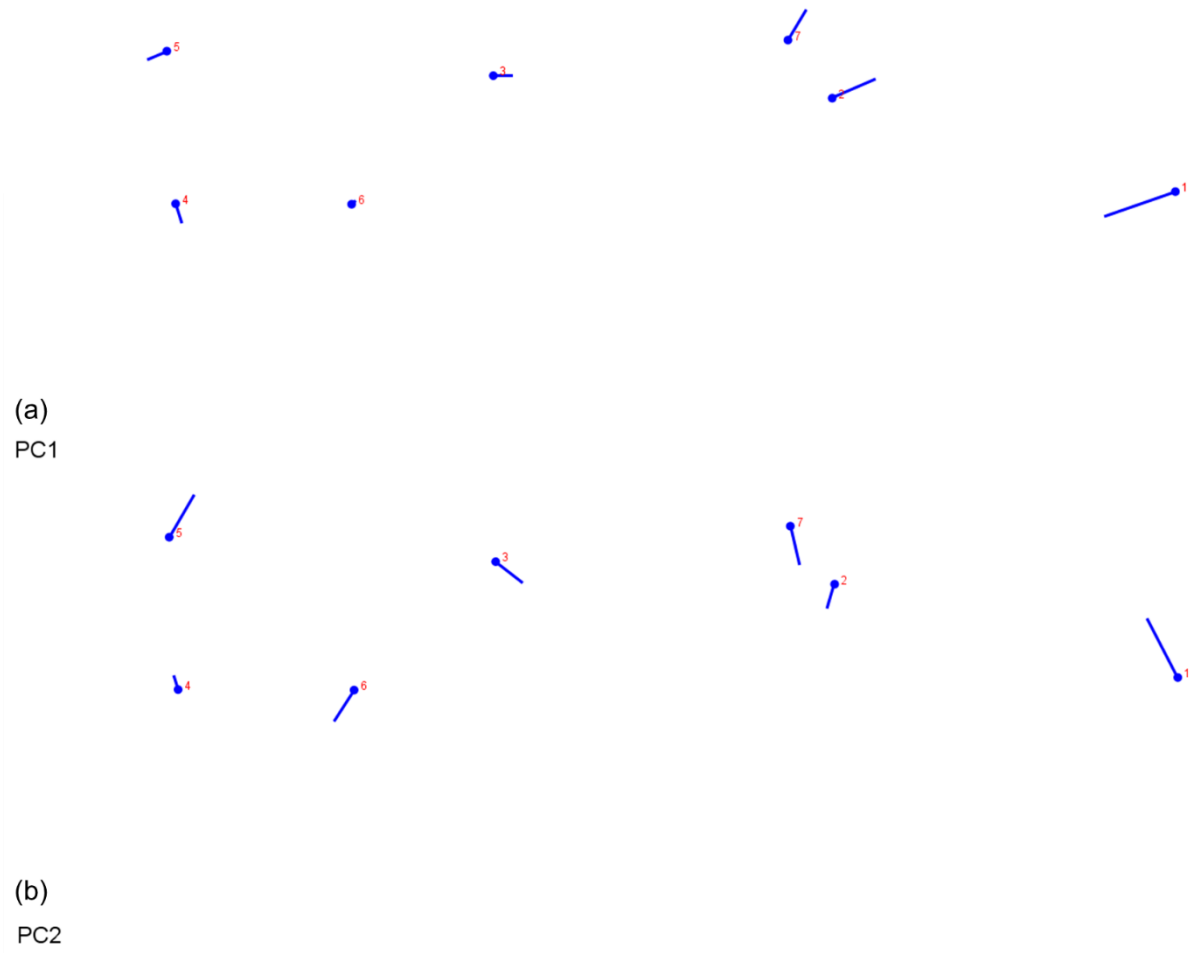

Figure 5.4.1. Scree Plots for both PC1 and PC2 of the Extracellular GST-Tagged and Untagged Orexin Receptor Landmarks. Scree plots for PC1 (a) and PC2 (b) are shown for the 7 extracellular landmarks of the 22 GST-tagged and untagged orexin receptor structures.

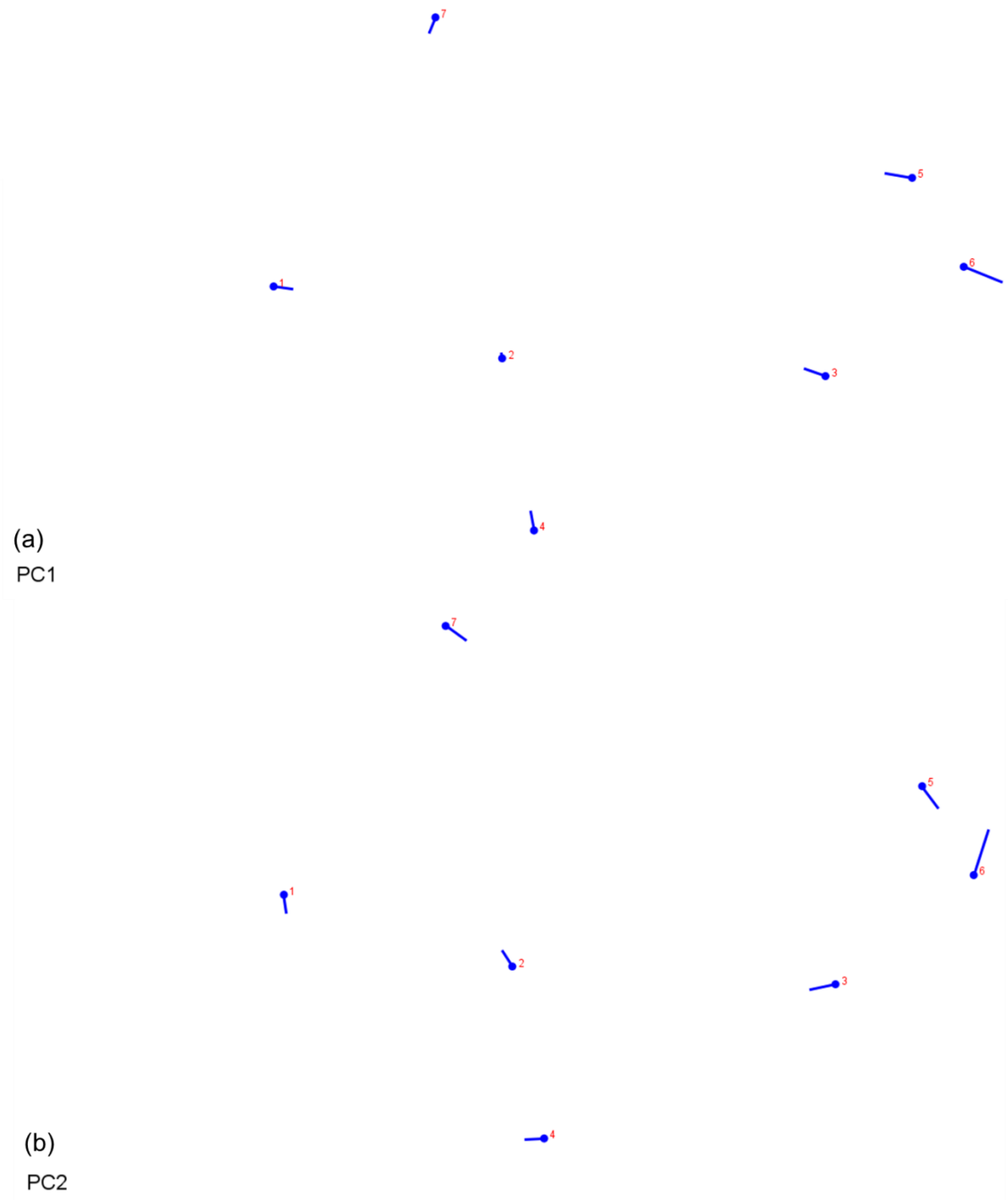

Figure 5.4.2. Scree Plots for both PC1 and PC2 of the Intracellular GST-Tagged and Untagged Orexin Receptor Landmarks. Scree plots for PC1 (a) and PC2 (b) are shown for the 7 intracellular landmarks of the 22 GST-tagged and untagged orexin receptor structures.

PC1 vs PC2 Morphospace

Scatterplots comparing PC1 and PC2 show a spread of data across both the extracellular (Figure 5.4.3a) and intracellular (Figure 5.4.3b) principal components, with evidence of both clustering and outliers. The intracellular landmark coordinates (b) revealed a distinct cluster of

points around the 0.25 PC1, -0.05 PC2 morphospace, with the majority of outliers varying generally between PC1 and PC2. Similarly, the extracellular landmarks (a) revealed a reasonable cluster of points around the -0.25 PC1, 0.00 PC2 morphospace, and remaining data also generally distributed between PC1 and PC2. Importantly, it should be noted that the PC1 axis for the intracellular data (b) was slightly more extended than the extracellular axis (a), reflecting a greater degree of variance for the intracellular cluster at 0.25, -0.05. This is also further supported by the greater cumulative percentage variation for intracellular landmarks, shown in Table 5.4.1.

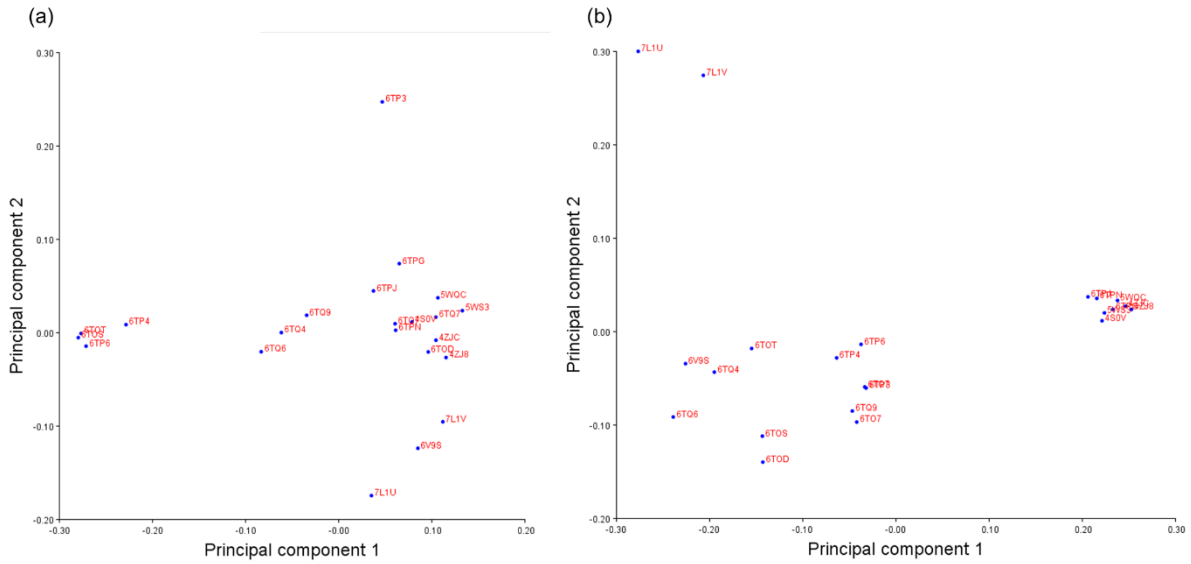

Figure 5.4.3. Morphospace of the GST-Tagged and Untagged Orexin Receptor Structures. Position in morphospace of 22 GST-tagged and untagged orexin receptor structures along PC1 and PC2, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Each structure is labelled with its PDB code.

For the orexin receptor structures then, it was hypothesised that the presence of the GST fusion protein inserted into ICL3 may contribute to a significant difference to the intracellular morphospace, when compared to the untagged orexin receptor structures. As shown in Figure 5.4.4, this was shown to be the case, with the extracellular morphospace (a) revealing a general overlap between categorised structures, and the intracellular morphospace (b) revealing a distinct cluster of orexin receptors with the GST fusion protein (red), at the positive terminus of PC1. PC1 therefore, as the dominant aspect of variation in this dataset, likely represents the structural effect of GST insertion into ICL3. The GST-fused structures are red, and untagged are blue, based on information gathered from the *GPCRdb*. Interestingly, Figure 5.4.4b reveals that one of the GST-fused orexin receptors does not cluster with the rest of its categorised group, and instead lies at the -0.225 PC1 area of the morphospace, within the untagged group (blue). This particular structure, 6V9S, was unique in this dataset as it was

resolved in complex with a new, subtype-selective orexin receptor antagonist called JH112, which differs from the other receptors resolved in complex with the typical non-selective orexin receptor antagonist, suvorexant (Hellmann *et al.*, 2020). One explanation for 6V9S's deviation from its expected cluster may be a difference in conformation stabilised by JH112, however, any other structural factors cannot be ruled out; further structural analyses of the effects of this selective antagonist should be explored. Either way, the inclusion of the GST fusion protein in ICL3 of these receptors does correlate with a significant difference to the untagged receptors in all but one structures, which differs from the previous understanding that ICL3 insertion should not impact on the rest of the structure.

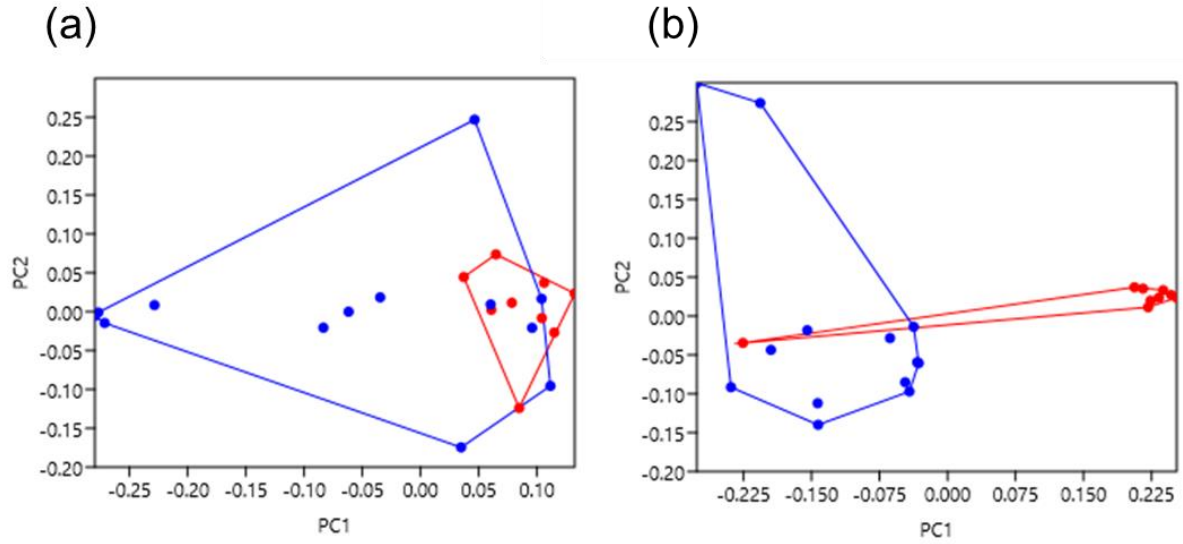

Figure 5.4.4. Categorised Morphospace of the GST-Tagged and Untagged Orexin Receptor Structures. Position in morphospace of 22 GST-tagged and untagged orexin receptor structures, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Convex hulls delimit the structures as categorised, indicating group affiliation. GST-tagged is red, untagged is blue.

PERMANOVA (extracellular F = 5.505, p = 0.0132; intracellular F = 19.36, p = 0.0001) and ANOSIM (extracellular R = 0.1054, p = 0.0852; intracellular R = 0.6166, p = 0.0001) showed that there were significant differences between the categorised groups, summarised in Table 5.4.2. For the extracellular morphospace, PERMANOVA showed a significant difference in variation between the GST-tagged and untagged orexin receptors, while ANOSIM was not significant. The lower F and R values for the extracellular morphospace are therefore consistent with the overlap between groups shown in Figure 5.4.4a. For the intracellular morphospace, both PERMANOVA and ANOSIM revealed significant differences between the GST-tagged and untagged orexin receptor structures. Relatively greater F and R values indicate a greater degree of variation, and greater dissimilarity between groups, as shown in Figure 5.4.4b.

Table 5.4.2. Statistical Analyses of the GST-Tagged and Untagged Orexin Receptor Structures. Pairwise comparisons of PC1 and PC2 scores of both the extra- and intracellular landmarks of the GST-tagged and untagged orexin receptor structures, in terms of reported inclusion of a GST fusion protein. Structures were categorised as either GST-tagged or untagged, based on information gained from the *GPCRdb*. PERMANOVA *p* values are above the diagonal, and ANOSIM *p* values are below the diagonal. The Bonferroni sequential corrected significant *p* values are indicated in bold.

Extracellular			Intracellular		
	GST	Untagged		GST	Untagged
GST		**0.0119**	GST		**0.0002**
Untagged	0.0781		Untagged	**0.0001**	

5.5 – Preliminary Comparison of AlphaFold and Experimentally Resolved Structures

As discussed throughout this thesis, proteins are essential for the processes underpinning biological life itself, and understanding of their structure/function relationships are key to elucidating the mechanisms involved in health, disease, and efficacious therapeutics. As such, a great deal of effort has led to the experimental determination of over 100,000 unique protein structures (wwPDB consortium, 2019). Despite this resounding success, this is only a fraction of known protein sequences, and experimental methods can take months to years of effort for each structure resolved. In order to overcome these limitations, and complement experimental efforts, computational techniques are needed to aid sequence-based structure prediction; many of these methods do not yield atomic accuracy, and are hampered by the difficulty of folding proteins realistically, *in silico* (Dill *et al.*, 2008).

The AlphaFold project, from Google's DeepMind AI, was hailed as a revolutionary advancement in the accurate prediction of protein structures, even in the absence of homology models (Jumper *et al.*, 2021). At its core, AlphaFold is a novel, machine learning technique which incorporates both protein sequence and biophysical data, enabling the prediction of near-experimental structures, and significantly outperforming alternative approaches in the Critical Assessment of Protein Structure Prediction (CASP14) network. In recent comparisons to RoseTTAFold and Modeller, AlphaFold and RoseTTAFold (the non-homology methods) both outperformed the homology-based Modeller program in predicting GPCR structures with no template, scored by RMSD (Lee, Su and Tseng, 2022).

Despite its resounding success in solving the 'protein folding problem', the AlphaFold models usually represent one single state, lacking the conformational heterogeneity observed in experimental structures. As highlighted in this thesis, this heterogeneity underpins the link between subtle conformational differences, and functional states; if the future of drug discovery necessitates biased or sub-type selective therapeutics, models which represent the

holistic dynamics of proteins are required. As such, several advancements have been proposed to expand the ensemble of AlphaFold's structural states, or to enable predictions which encompass bias towards the various states of activation transitions (Sala *et al.*, 2023). While these 'Fold' variants have great potential in overcoming this limitation in time, it has been highlighted that a current challenge exists in whether these predictive tools can accurately model conformational differences as a result of external factors such as mutations, lipid contacts and post-translational modifications, etc. (Nussinov *et al.*, 2022; Pak *et al.*, 2023).

While the progress made in predicting and folding protein structure is undeniable, models do not yet represent the conformational heterogeneity associated with experimentally resolved structures. The GPCR structures predicted by AlphaFold have been recently compared to their experimentally resolved equivalents, whereby it was revealed that the overall backbone was predicted fairly accurately (He *et al.*, 2023). However, several other aspects of GPCR structure were found to differ between the AlphaFold and experimentally resolved models including the transmembrane domain assembly, ligand binding domains, and the conformations of G protein binding domains (He *et al.*, 2023). Overall, the predicted models of GPCRs do not currently possess the resolution of structural information needed for functional mapping and application to structure-based novel drug discovery.

To complement these findings, a comparison between the experimentally resolved β2AR structures and the β2AR AlphaFold model using the geometric morphometric technique is shown in Figure 5.5. Refer to Appendix B8 for a full list of structures, including PDB codes and reference to their publication. While this is just a preliminary investigation, and no statistical testing could be performed against just one AlphaFold model, there are clear observable differences. The β2AR AlphaFold model (which is in an inactive conformation) shown in pink and indicated by the black arrows, clearly does not group with the inactive experimental structures shown in red. While the extracellular face (a) does exhibit general overlap between groups, the intracellular face (b) reveals a significant separation between the inactive (red) and active (green) or active with G_s (blue) groups; the AlphaFold model clearly does not share morphological similarity with the experimental groups, falling at the very negative end of PC2.

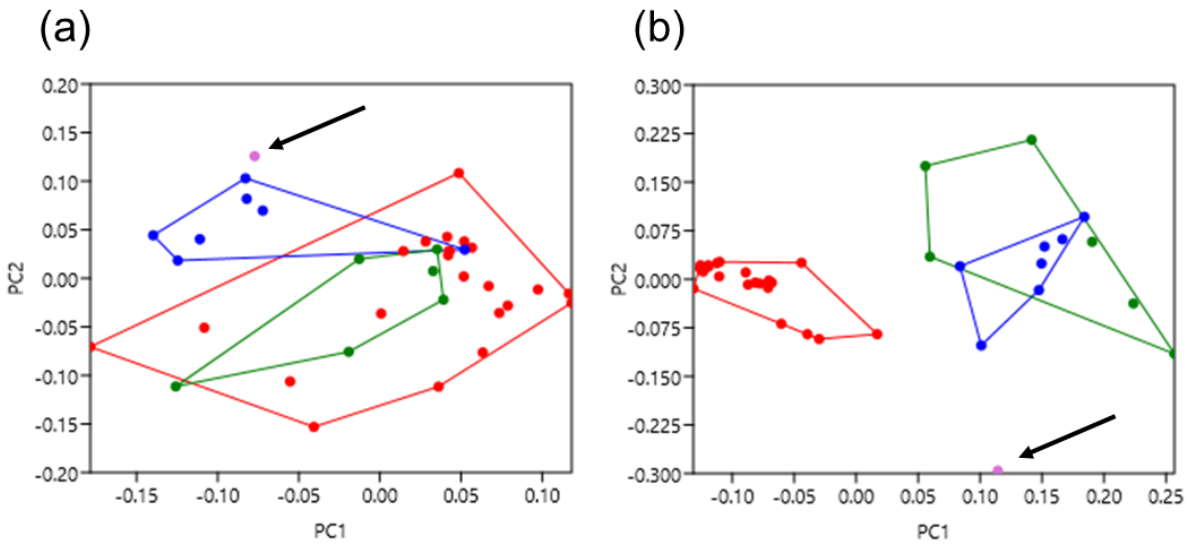

Figure 5.5. Morphospace of the AlphaFold and Experimentally Resolved Structures of the β2-Adrenergic Receptor. Position in morphospace of 37 β2-adrenergic receptor structures, based on the 7 extracellular landmarks (a) and 7 intracellular landmarks (b). Convex hulls delimit the structures as categorised, indicating group affiliation. Inactive is red, active is green, active with G_s is blue, the inactive AlphaFold structure is pink, indicated by the black arrow.

It is likely that with continuous enhancement over time, the predictive models folded *in silico* will improve in resolution and structural realism, and will become more and more comparable to those determined by experimental methods. In any case, these preliminary results suggest that geometric morphometrics and its rigorous statistical testing could also find use in predictive versus experimental comparisons of structure.

5.6 – Discussion

To reiterate, the foundational geometric morphometric analyses of GPCR structures presented in Chapter 4 were consistent in mapping statistically significant patterns in the conformational landscape to functionally relevant states. The aim of Chapter 5 therefore, was to further explore the multifaceted applicability of geometric morphometrics to various aspects of GPCR structure, and determine its usefulness beyond the scope of mapping activation state transitions. These novel experiments were also met with success, revealing both significant and non-significant morphological patterns in global structure, suggesting that this geometric morphometric technique has the potential for use across the wider membrane protein structural biology field.

It was reassuring to find that the concept of grouping receptors and their G protein coupling specificity by interhelical distances was also reflected in the principal component morphospaces, supported by statistically significant testing. These results therefore filled in some of the gaps left by previous research (Kruse *et al.*, 2012), and also expanded the analyses to the extracellular helical ends, as well as the existing focus on intracellular ends. Of course, the intracellular face of receptors makes direct contact with the heterotrimeric G proteins via the binding cleft opened as a result of ligand-activated conformational change, for example. However, G protein binding itself has also been shown to allosterically influence the receptor's orthosteric site of ligand binding (DeVree *et al.*, 2016), highlighting the importance of additionally analysing the extracellular receptor morphology; the results shown in section 5.2 revealed very significant differences between receptors coupled to different G proteins, for both the extra- and intracellular morphospace. Interestingly, this observation was also maintained with the addition of more recently determined structures, and a consequently larger dataset, reinforcing the capability of geometric morphometrics to reveal patterns in complex shape data.

Now, if the structure-based signalling of GPCRs is to be utilised in the development or discovery of novel therapeutics, the morphological groupings revealed by geometric morphometrics must be linked to functional outcomes, combining enhanced biophysical and biochemical assays with various G protein sub-types, and a consideration of biased signalling. Whilst this is still a work in progress, a recent study by Masuho *et al.*, 2023 undertook a systematic approach to quantitatively analyse the GEF activity of 124 GPCRs in an attempt to evaluate their G protein coupling specificity and promiscuity (Masuho *et al.*, 2023). They were able to establish an ordered classification of receptors, ranking each by their coupling preferences to the G_s, $G_{q/11}$, $G_{i/o}$ and $G_{12/13}$ sub-types. It was found that 73% of the 124 analysed GPCRs were able to activate more than one Gα sub-type, which suggests that the majority have both primary and secondary coupling specificities (Masuho *et al.*, 2023), which is consistent with the concept of the major and minor pathways of biased signalling.

In addition to this, Masuho *et al*. utilised a machine learning approach to accurately predict this coupling specificity, based on the amplitude and activation-rate of the GEF enzymatic activity of the receptors, highlighting the usefulness of machine learning in elucidating the structure/function relationship of proteins. In an attempt to determine the structural underpinnings of these observations, it was found that, at least for family A receptors, select residues from all seven transmembrane helices and all three intracellular loops contribute to G protein binding. TMs 5 and 6, and ICLs 2 and 3 contribute the majority of these residues, and it was theorised overall, that receptor Gα selectivity depends on the 3-dimensional space determined by the differential organisation of these residues (Masuho *et al.*, 2023). Overall,

this study is an excellent example of the functional investigations that geometric morphometrics should be paired with, and is a strong step towards building a holistic and multidisciplinary model of GPCR structure and function.

Next, in addition to elucidating the structural aspects underpinning G protein coupling specificity, another aim of Chapter 5 was to explore the structural consequences of protein modifications, with thermostabilising mutagenesis and the use of fusion insertions as examples. Of course, given the challenges of the expression and purification of membrane proteins in general, strategies to improve the solubility or stability of protein samples are sometimes a necessary compromise (Wiseman *et al.*, 2020). A particularly salient topic in drug discovery is that of more realistic experimental conditions, and the rise of organ on a chip technology (Singh *et al.*, 2022). The more realistic cell culture and pharmacology become, reflecting a more realistic human physiology, the greater the chances therapeutics have at clinical trial. Suffice to say, though protein modifications have been instrumental in their experimentation, the more modifications made to human proteins, the less realistic they become compared to wild-type. It is therefore recommended that researchers should be aware of the structural and functional effects of these modifications, and indeed embed these decisions into early experimental design stages. There are two obvious examples of how geometric morphometrics could aid in this endeavour, as a comparative sense-checking tool.

Firstly, upon determination of protein structure by X-ray crystallography or cryo-EM for example, the resulting model's quality is examined. While there are several quantitative factors which correlate with model quality such as resolution and Ramachandran distribution, there is no single measure which sufficiently describes a model's quality, and the determining factors of high or low limits (Domagalski *et al.*, 2014). While models do have the potential to be further refined by constantly evolving software, the PDB depository itself varies quite considerably in the quality and validation of its structures (Brzezinski *et al.*, 2020). The geometric morphometric method could be used as a sense-checking tool as one component of an enhanced quality and validation process for submitted structures. For example, if a particular receptor has been resolved in an inactive state, does it morphologically group with the other inactive state crystals of its sub-family? If a receptor has its structure determined with a novel structure-based biased agonist, does it group morphologically as expected? Again, if a receptor is resolved in complex with a particular G protein sub-type, does it group sensibly with the other experimentally resolved receptor-G protein complexes? If the answers are yes, there is an added credibility in observing receptor morphology which is consistent with its structural characteristics, however, if the answer is no, further exploration and explanation might be required.

Secondly, as well as sense-checking models at the end stages of structure determination, geometric morphometrics could find use in the very beginning stages of experimental design. For example, if the determination of a receptor's structure necessitates modification, such as thermostabilisation or the use of a fusion protein, the potential structural or functional effects of modifications could be modelled *in silico* prior to *in vitro* experiments. This might entail a sequence-based folding step, followed by geometric morphometric analysis of the modified receptor, and comparisons of morphology against non-modified equivalents. The technology to achieve this approach is not yet advanced enough (He *et al.*, 2023), requiring high-resolution, accurate sequence-based folding which can incorporate realistic, 3-dimensional effects of mutagenesis, as well as a refined process of landmark selection and morphometric analysis. Inevitably, as computational technology advances, the results of these *in silico* predictive comparisons might be of benefit prior to *in vitro* efforts.

To conclude then, GPCRs are highly attractive drugs targets due to their wide involvement in pathophysiological disease states, however, only approximately 1% of drugs successfully progress through clinical trials (Santos *et al.*, 2017). In order to develop safe, efficacious, novel therapeutics which act on these receptors, a comprehensive understanding of their structure/function relationship is needed to form a multidisciplinary framework from which drug discovery will have a greater chance of success. It is crucial to explore the pleiotropic relationship between ligand, receptor, and heterotrimeric G protein, as it is now clear that ligands are able to selectively bias receptors towards certain signalling pathways, and receptors themselves exhibit preference in their G protein coupling specificity. The diverse heterogeneity of conformational states observed in GPCR crystal structure snapshots correlates with their pleiotropic mechanisms of action, and an integration of structure with function must therefore encompass the complete dynamic range of these proteins. Geometric morphometrics is a highly promising technique which facilitates the quantitative decoding of the structural aspects of GPCRs which contribute to their functional plasticity. With ongoing refinement, and combination with enhanced functional assays such as those detailed in Chapter 3, geometric morphometrics has the potential to significantly contribute to the multidisciplinary elucidation of structure-based signalling, within the GPCR superfamily and indeed membrane proteins in general. This body of work therefore gives a reliable indication that analyses of global receptor morphology reveal both dominant and more subtle conformational variations in structure, and is widely applicable to the range of structural aspects which underpin functional outcomes.

Chapter 6: Future Directions and Conclusions

Research undertaken with GPCRs has faced significant challenges due to the dynamic nature of this superfamily of membrane proteins, and while a technological boom over the last few decades has certainly translated into an exponential growth in the capabilities to study GPCRs, significant challenges remain. While the structure/function relationship of proteins is a well-known paradigm, it is now clear that that of GPCRs is much more nuanced than originally thought. Some ligands are able to induce signalling in which receptors are selectively biased towards specific pathways or functional outcomes, which correlates with a diverse heterogeneous conformational landscape of activation transition and intermediate states. Moreover, the majority of GPCRs are now thought to signal through more than one G protein sub-type, preferentially coupling with ranked selectivity, and can additionally be allosterically modulated by a range of interactions, including those with the RAMPs. Given that the ultimate application of GPCR research is to discover and develop safe, effective, and specific therapeutic options for the plethora of diseases in which they are implicated, a holistic understanding of GPCRs is needed which encompasses all of these discussed points. This will be achieved with a comprehensive, multidisciplinary combination of biophysical and biochemical functional assays, determination and analyses of structure, and complementary computational approaches.

The first aim of this thesis was to develop a FRET-based biophysical assay for the calcitonin family of GPCRs and the RAMPs, to further elucidate their interaction dynamics, and to serve as a novel drug screening platform for these family B1 GPCRs. The understanding of the structural and pharmacological consequences of RAMP interactions have been well-characterised since their discovery in the 1990s, and as Kotliar *et al*., 2023 have highlighted, there is now a necessity for novel assays which include both receptors and RAMPs. As the inception of this FRET assay began in 2017, the fact that there is still a call for novel assays in 2023 further reinforces the need for and justification of this work. The results of Chapter 3 show that this FRET assay for the calcitonin receptors and RAMPs was indeed successfully established, using CFP and YFP as a FRET pair, along with the additional determination that the constructs retained their expected pharmacology in terms of cAMP responses to CGRP (Hay *et al.*, 2018).

Unfortunately, this project was significantly affected by the COVID-19 pandemic, and the established assay could not be used for its intended purposes, though it does form a novel basis from which future work will build. In the short term, the future directions of this project likely involve an investigation into the dynamics of the expression and interactions of receptors and RAMPs, and if receptor-associated RAMPs are able to outcompete each other,

contributing to an allosteric modulation of the functional plasticity of select GPCRs. In addition to this, the established FRET assay can be tested and implemented as a biophysical drug screening assay, specifically for the identification of compounds or therapeutics which promote or prevent receptor-RAMP interactions, and consequently link this to their functional outcomes. Furthermore, the expression and purification of mammalian membrane proteins from the photoreceptor cells of *Drosophila melanogaster* is a promising but underutilised tool, providing several advantages over more traditional approaches. The preliminary results for the expression of a fluorescently-tagged CGRP receptor shown in Chapter 3 are certainly favourable, however, this project was also ended prematurely by the COVID-19 pandemic. Future work would include generating the *Drosophila* strain which can be driven to express both receptor and RAMP, and a subsequent exploration of techniques to extract and purify these complexes, for direct application to the aforementioned biophysical and biochemical analyses, to gain insight into their structural and functional dynamics.

The second aim of this thesis was to gain further insight into the structure-based signalling of GPCRs, through the novel use of geometric morphometrics and principal component analysis, and contribute to the elucidation of their complete dynamic plasticity of structure and function. Technological advancements in X-ray crystallography and cryo-EM have led to breakthroughs in the determination of protein structure, offering valuable insights into their mechanisms of action, and structural characteristics when resolved in different states. The traditional methods of comparing resolved structures are certainly useful but do possess inherent limitations, for example, RMSD fails to capture more subtle conformational differences and Q score fails to capture larger global differences; geometric morphometrics can overcome these limitations, and facilitate a more nuanced approach. The case study results of Chapter 4 consistently and reliably demonstrate that the classification of GPCR structures is possible, by their global differences in shape morphology, and rigorous statistical testing for variation and dissimilarity between categorised groups. In addition, the results of Chapter 5 suggest that the geometric morphometric technique is able to reveal differences in receptor morphology caused by a variety of structural aspects, and likely has wide applicability to membrane proteins in general. Overall, this novel use of geometric morphometrics with GPCRs has been a resounding success, with much future potential.

With regards to the future work for this project, the refinement of intelligent landmark selection with machine learning and sequence alignment would enable a comprehensive investigation into landmark selection, and how best to represent protein shape morphology. As the geometric morphometric method is currently manually laborious, it would also massively benefit from some degree of automation for both coordinate collection from PDBs, and the morphometric analysis itself. It is currently challenging to quickly compare all principal

components and their morphospaces, and thus user friendliness should be a focal point of future work, to entice the wider structural biology community to engage with this approach. It would be ideal to develop a complete, user friendly software package which facilitates the whole process in one easy to use program, from choice of receptors and PDBs, intelligent landmark selection, the geometric morphometric and principal component analyses, visualisation of morphospaces, to group categorisation and statistical testing. This is particularly important for the comprehensive comparisons of lower ranked principal components which may represent the more subtle shape variations associated with biased signalling, for example. Furthermore, additional future work may also involve the application of geometric morphometrics to other relevant targets such as gated ion channels, non-GPCR receptors, and enzymes, which all form the majority of druggable proteins (Tiefenauer and Demarche, 2012).

Overall then, towards longer-term future directions, the geometric morphometric technique will be well-placed for use as a sense- and quality-checking tool for determined structures, and might even find involvement in the Protein Data Bank itself during their structure quality assessment, as well as services such as the *GPCRdb* and *mpstruc* databases. This links back to the need for an automated, user friendly experience where every component of this technique is found in an integrated package. In addition to this, the morphospace patterns revealed by morphometric analyses will need to be combined with functional assay data to definitively correlate structural underpinnings with functional outcomes. Using the receptor-RAMP interactions discussed in this thesis as an example, the biophysical FRET assay may be used to assess receptor complex dynamics or effects of drugs in live cells, and subsequently combine with a morphological analysis of receptor-RAMP shape changes, to link structure and function. This of course depends upon enhanced, temporally relevant assays and accurate, high-resolution structure determination which, as discussed, still face inherent limitations. It may well be that the computational sequence-based folding algorithms such as AlphaFold could supplement this approach when their predictive capabilities reach a consistent, near-experimental quality.

In summary, the novel application of geometric morphometrics to GPCR structures has enabled a quantitative analysis of their conformational landscape, revealing structural groupings and common shape morphologies which correlate with functional states. Importantly, the statistically affirmed results presented in this thesis indicate that the geometric morphometric method is effective in the exploration of states of bound ligand, states of activation transitions, and states of coupled G proteins. Taken in combination with the diverse conformational heterogeneity observed in structures, these results build towards a comprehensive understanding of the structure-based signalling of GPCRs. To truly correlate

this with signalling pathways, pharmacological profiles, and physiological responses in both healthy and disease states, the structural data from geometric morphometrics must be combined with enhanced biophysical and biochemical assays, such as the FRET assay established in Chapter 3, to build a holistic model of the complete dynamic range of GPCR function, and their structural determinants. Ultimately, building this multidisciplinary model of understanding will inform the drug discovery process for GPCRs, enabling unprecedented discovery and development of safe, effective therapeutics, for which in part, this thesis makes a contribution.

List of References

Abosamak, N.R. and Shahin, M.H. (2024) 'Beta2 Receptor Agonists and Antagonists', in *StatPearls*. Treasure Island (FL): StatPearls Publishing. Available at: http://www.ncbi.nlm.nih.gov/books/NBK559069/ (Accessed: 19 March 2024).

Adams, D.C., Rohlf, F.J. and Slice, D.E. (2004) 'Geometric morphometrics: Ten years of progress following the "revolution"', *Italian Journal of Zoology*, 71(1), pp. 5–16. Available at: https://doi.org/10.1080/11250000409356545.

Ahmad, M., Hirz, M., Pichler, H. and Schwab, H. (2014) 'Protein expression in Pichia pastoris: recent achievements and perspectives for heterologous protein production', *Applied Microbiology and Biotechnology*, 98(12), pp. 5301–5317. Available at: https://doi.org/10.1007/s00253-014-5732-5.

Ahn, S., Shenoy, S.K., Luttrell, L.M. and Lefkowitz, R.J. (2020) 'SnapShot: β-Arrestin Functions', *Cell*, 182(5), pp. 1362-1362.e1. Available at: https://doi.org/10.1016/j.cell.2020.07.034.

Aikawa, S., Ishii, M., Yanagisawa, M., Sakakibara, Y. and Sakurai, T. (2008) 'Effect of neuropeptide B on feeding behavior is influenced by endogenous corticotropin-releasing factor activities', *Regulatory Peptides*, 151(1), pp. 147–152. Available at: https://doi.org/10.1016/j.regpep.2008.08.001.

Aiyar, N., Disa, J., Ao, Z., Xu, D., Surya, A., Pillarisetti, K., Parameswaran, N., Gupta, S.K., Douglas, S.A. and Nambi, P. (2002) 'Molecular cloning and pharmacological characterization of bovine calcitonin receptor-like receptor from bovine aortic endothelial cells', *Biochemical Pharmacology*, 63(11), pp. 1949–1959. Available at: https://doi.org/10.1016/S0006-2952(02)00990-5.

Aiyar, N., Rand, K., Elshourbagy, N.A., Zeng, Z., Adamou, J.E., Bergsma, D.J. and Li, Y. (1996) 'A cDNA encoding the calcitonin gene-related peptide type 1 receptor', *The Journal of Biological Chemistry*, 271(19), pp. 11325–11329. Available at: https://doi.org/10.1074/jbc.271.19.11325.

Aksoydan, B. and Durdagi, S. (2022) 'Virtual drug repurposing study for the CGRPR identifies pentagastrin and leuprorelin as putative candidates', *Journal of Molecular Graphics & Modelling*, 116, p. 108254. Available at: https://doi.org/10.1016/j.jmgm.2022.108254.

Alexandrov, A.I., Mileni, M., Chien, E.Y.T., Hanson, M.A. and Stevens, R.C. (2008) 'Microscale Fluorescent Thermal Stability Assay for Membrane Proteins', *Structure*, 16(3), pp. 351–359. Available at: https://doi.org/10.1016/j.str.2008.02.004.

Almeida, J.G., Preto, A.J., Koukos, P.I., Bonvin, A.M.J.J. and Moreira, I.S. (2017) 'Membrane proteins structures: A review on computational modeling tools', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1859(10), pp. 2021–2039. Available at: https://doi.org/10.1016/j.bbamem.2017.07.008.

Almén, M.S., Nordström, K.J., Fredriksson, R. and Schiöth, H.B. (2009) 'Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin', *BMC Biology*, 7(1), p. 50. Available at: https://doi.org/10.1186/1741-7007-7-50.

Anandan, A. and Vrielink, A. (2016) 'Detergents in Membrane Protein Purification and Crystallisation', in I. Moraes (ed.) *The Next Generation in Membrane Protein Structure*

Determination. Cham: Springer International Publishing (Advances in Experimental Medicine and Biology), pp. 13–28. Available at: https://doi.org/10.1007/978-3-319-35072-1_2.

Andrade, D., Serra, R., Svensjö, E., Lima, A.P.C., Ramos Junior, E.S., Fortes, F.S., Morandini, A.C.F., Morandi, V., Soeiro, M. de N., Tanowitz, H.B. and Scharfstein, J. (2012) 'Trypanosoma cruzi invades host cells through the activation of endothelin and bradykinin receptors: a converging pathway leading to chagasic vasculopathy', *British Journal of Pharmacology*, 165(5), pp. 1333–1347. Available at: https://doi.org/10.1111/j.1476-5381.2011.01609.x.

Andréll, J. and Tate, C.G. (2013) 'Overexpression of membrane proteins in mammalian cells for structural studies', *Molecular Membrane Biology*, 30(1), pp. 52–63. Available at: https://doi.org/10.3109/09687688.2012.703703.

Arachea, B.T., Sun, Z., Potente, N., Malik, R., Isailovic, D. and Viola, R.E. (2012) 'Detergent selection for enhanced extraction of membrane proteins', *Protein Expression and Purification*, 86(1), pp. 12–20. Available at: https://doi.org/10.1016/j.pep.2012.08.016.

Araya, T., Matsuba, Y., Suzuki, H., Doura, T., Nuemket, N., Nango, E., Yamamoto, M., Im, D., Asada, H., Kiyonaka, S. and Iwata, S. (2024) 'Crystal structure reveals the binding mode and selectivity of a photoswitchable ligand for the adenosine A2A receptor', *Biochemical and Biophysical Research Communications*, 695, p. 149393. Available at: https://doi.org/10.1016/j.bbrc.2023.149393.

Arévalo-Martín, Á., García-Ovejero, D., Gómez, O., Rubio-Araiz, A., Navarro-Galve, B., Guaza, C., Molina-Holgado, E. and Molina-Holgado, F. (2008) 'CB2 cannabinoid receptors as an emerging target for demyelinating diseases: from neuroimmune interactions to cell replacement strategies', *British Journal of Pharmacology*, 153(2), pp. 216–225. Available at: https://doi.org/10.1038/sj.bjp.0707466.

Arora, P., Cuevas, B.D., Russo, A., Johnson, G.L. and Trejo, J. (2008) 'Persistent transactivation of EGFR and ErbB2/HER2 by protease-activated receptor-1 promotes breast carcinoma cell invasion', *Oncogene*, 27(32), pp. 4434–4445. Available at: https://doi.org/10.1038/onc.2008.84.

Asada, H., Uemura, T., Yurugi-Kobayashi, T., Shiroishi, M., Shimamura, T., Tsujimoto, H., Ito, K., Sugawara, T., Nakane, T., Nomura, N., Murata, T., Haga, T., Iwata, S. and Kobayashi, T. (2011) 'Evaluation of the Pichia pastoris expression system for the production of GPCRs for structural analysis', *Microbial Cell Factories*, 10(1), p. 24. Available at: https://doi.org/10.1186/1475-2859-10-24.

Attrill, H., Harding, P.J., Smith, E., Ross, S. and Watts, A. (2009) 'Improved yield of a ligand-binding GPCR expressed in E. coli for structural studies', *Protein Expression and Purification*, 64(1), pp. 32–38. Available at: https://doi.org/10.1016/j.pep.2008.10.001.

Avgoustou, P., Jailani, A.B.A., Zirimwabagabo, J.-O., Tozer, M.J., Gibson, K.R., Glossop, P.A., Mills, J.E.J., Porter, R.A., Blaney, P., Bungay, P.J., Wang, N., Shaw, A.P., Bigos, K.J.A., Holmes, J.L., Warrington, J.I., Skerry, T.M., Harrity, J.P.A. and Richards, G.O. (2020) 'Discovery of a First-in-Class Potent Small Molecule Antagonist against the Adrenomedullin-2 Receptor', *ACS Pharmacology & Translational Science*, 3(4), pp. 706–719. Available at: https://doi.org/10.1021/acsptsci.0c00032.

Awad, J.A., Johnson, R.A., Jakobs, K.H. and Schultz, G. (1983) 'Interactions of forskolin and adenylate cyclase. Effects on substrate kinetics and protection against inactivation by heat and N-ethylmaleimide.', *Journal of Biological Chemistry*, 258(5), pp. 2960–2965. Available at: https://doi.org/10.1016/S0021-9258(18)32814-X.

Ayers, K.L. and Thérond, P.P. (2010) 'Evaluating Smoothened as a G-protein-coupled receptor for Hedgehog signalling', *Trends in Cell Biology*, 20(5), pp. 287–298. Available at: https://doi.org/10.1016/j.tcb.2010.02.002.

Ayub, H., Clare, M., Broadbent, L., Simms, J., Goddard, A.D., Rothnie, A.J. and Bill, R.M. (2022) 'Membrane Protein Production in the Yeast P. pastoris', *Methods in Molecular Biology (Clifton, N.J.)*, 2507, pp. 187–199. Available at: https://doi.org/10.1007/978-1-0716-2368-8_10.

Bailey, R.J. and Hay, D.L. (2006) 'Pharmacology of the human CGRP1 receptor in Cos 7 cells', *Peptides*, 27(6), pp. 1367–1375. Available at: https://doi.org/10.1016/j.peptides.2005.11.014.

Bailey, S., Harris, M., Barkan, K., Winfield, I., Harper, M.T., Simms, J., Ladds, G., Wheatley, M. and Poyner, D. (2019) 'Interactions between RAMP2 and CRF receptors: The effect of receptor subtypes, splice variants and cell context', *Biochimica Et Biophysica Acta. Biomembranes*, 1861(5), pp. 997–1003. Available at: https://doi.org/10.1016/j.bbamem.2019.02.008.

Bajar, B.T., Wang, E.S., Zhang, S., Lin, M.Z. and Chu, J. (2016) 'A Guide to Fluorescent Protein FRET Pairs', *Sensors (Basel, Switzerland)*, 16(9), p. 1488. Available at: https://doi.org/10.3390/s16091488.

Baker, L.A., Folkers, G.E., Sinnige, T., Houben, K., Kaplan, M., van der Cruijsen, E.A.W. and Baldus, M. (2015) 'Magic-angle-spinning solid-state NMR of membrane proteins', *Methods in Enzymology*, 557, pp. 307–328. Available at: https://doi.org/10.1016/bs.mie.2014.12.023.

Ballesteros, J.A., Jensen, A.D., Liapakis, G., Rasmussen, S.G.F., Shi, L., Gether, U. and Javitch, J.A. (2001) 'Activation of the β2-Adrenergic Receptor Involves Disruption of an Ionic Lock between the Cytoplasmic Ends of Transmembrane Segments 3 and 6*', *Journal of Biological Chemistry*, 276(31), pp. 29171–29177. Available at: https://doi.org/10.1074/jbc.M103747200.

Ballesteros, J.A. and Weinstein, H. (1995) '[19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors', in S.C. Sealfon (ed.) *Methods in Neurosciences*. Academic Press (Receptor Molecular Biology), pp. 366–428. Available at: https://doi.org/10.1016/S1043-9471(05)80049-7.

Baneres, J.-L., Martin, A., Hullot, P., Girard, J.-P., Rossi, J.-C. and Parello, J. (2003) 'Structure-based Analysis of GPCR Function: Conformational Adaptation of both Agonist and Receptor upon Leukotriene B4 Binding to Recombinant BLT1', *Journal of Molecular Biology*, 329(4), pp. 801–814. Available at: https://doi.org/10.1016/S0022-2836(03)00438-8.

Barbash, S., Lorenzen, E., Persson, T., Huber, T. and Sakmar, T.P. (2017) 'GPCRs globally coevolved with receptor activity-modifying proteins, RAMPs', *Proceedings of the National Academy of Sciences of the United States of America*, 114(45), pp. 12015–12020. Available at: https://doi.org/10.1073/pnas.1713074114.

Barducci, A., Bussi, G. and Parrinello, M. (2008) 'Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method', *Physical Review Letters*, 100(2), p. 020603. Available at: https://doi.org/10.1103/PhysRevLett.100.020603.

Barisione, G., Baroffio, M., Crimi, E. and Brusasco, V. (2010) 'Beta-Adrenergic Agonists', *Pharmaceuticals*, 3(4), pp. 1016–1044. Available at: https://doi.org/10.3390/ph3041016.

Bartels, P., Hirsch, P.E., Svanbäck, R. and Eklöv, P. (2012) 'Water Transparency Drives Intra-Population Divergence in Eurasian Perch (Perca fluviatilis)', *PLOS ONE*, 7(8), p. e43641. Available at: https://doi.org/10.1371/journal.pone.0043641.

Bayburt, T.H. and Sligar, S.G. (2003) 'Self-assembly of single integral membrane proteins into soluble nanoscale phospholipid bilayers', *Protein Science*, 12(11), pp. 2476–2481. Available at: https://doi.org/10.1110/ps.03267503.

Becker, T., Bhushan, S., Jarasch, A., Armache, J.-P., Funes, S., Jossinet, F., Gumbart, J., Mielke, T., Berninghausen, O., Schulten, K., Westhof, E., Gilmore, R., Mandon, E. and Beckmann, R. (2009) 'Structure of Monomeric Yeast and Mammalian Sec61 Complexes Interacting with the Translating Ribosome', *Science (New York, N.Y.)*, 326(5958), pp. 1369–1373. Available at: https://doi.org/10.1126/science.1178535.

Beisser, P.S., Lavreysen, H., Bruggeman, C.A. and Vink, C. (2008) 'Chemokines and Chemokine Receptors Encoded by Cytomegaloviruses', in T.E. Shenk and M.F. Stinski (eds) *Human Cytomegalovirus*. Berlin, Heidelberg: Springer (Current Topics in Microbiology and Immunology), pp. 221–242. Available at: https://doi.org/10.1007/978-3-540-77349-8_13.

Bell, I.M., Gallicchio, S.N., Wood, M.R., Quigley, A.G., Stump, C.A., Zartman, C.B., Fay, J.F., Li, C.-C., Lynch, J.J., Moore, E.L., Mosser, S.D., Prueksaritanont, T., Regan, C.P., Roller, S., Salvatore, C.A., Kane, S.A., Vacca, J.P. and Selnick, H.G. (2010) 'Discovery of MK-3207: A Highly Potent, Orally Bioavailable CGRP Receptor Antagonist', *ACS Medicinal Chemistry Letters*, 1(1), pp. 24–29. Available at: https://doi.org/10.1021/ml900016y.

Benovic, J.L., DeBlasi, A., Stone, W.C., Caron, M.G. and Lefkowitz, R.J. (1989) 'β-Adrenergic Receptor Kinase: Primary Structure Delineates a Multigene Family', *Science*, 246(4927), pp. 235–240. Available at: https://doi.org/10.1126/science.2552582.

Berg, K.A., Harvey, J.A., Spampinato, U. and Clarke, W.P. (2005) 'Physiological relevance of constitutive activity of 5-HT2A and 5-HT2C receptors', *Trends in Pharmacological Sciences*, 26(12), pp. 625–630. Available at: https://doi.org/10.1016/j.tips.2005.10.008.

Bernaudat, F., Frelet-Barrand, A., Pochon, N., Dementin, S., Hivin, P., Boutigny, S., Rioux, J.-B., Salvi, D., Seigneurin-Berny, D., Richaud, P., Joyard, J., Pignol, D., Sabaty, M., Desnos, T., Pebay-Peyroula, E., Darrouzet, E., Vernet, T. and Rolland, N. (2011) 'Heterologous Expression of Membrane Proteins: Choosing the Appropriate Host', *PLOS ONE*, 6(12), p. e29191. Available at: https://doi.org/10.1371/journal.pone.0029191.

Berridge, M.J. (2016) 'The Inositol Trisphosphate/Calcium Signaling Pathway in Health and Disease', *Physiological Reviews*, 96(4), pp. 1261–1296. Available at: https://doi.org/10.1152/physrev.00006.2016.

Bertin, B., Freissmuth, M., Breyer, R.M., Schütz, W., Strosberg, A.D. and Marullo, S. (1992) 'Functional expression of the human serotonin 5-HT1A receptor in Escherichia coli. Ligand binding properties and interaction with recombinant G protein alpha-subunits.', *Journal of Biological Chemistry*, 267(12), pp. 8200–8206. Available at: https://doi.org/10.1016/S0021-9258(18)42427-1.

Bhate, M.P., Wylie, B.J., Tian, L. and McDermott, A.E. (2010) 'Conformational Dynamics in the Selectivity Filter of KcsA in Response to Potassium Ion Concentration', *Journal of molecular biology*, 401(2), pp. 155–166. Available at: https://doi.org/10.1016/j.jmb.2010.06.031.

Bhola, N.E. and Grandis, J.R. (2008) 'Crosstalk between G-protein-coupled receptors and epidermal growth factor receptor in cancer', *Frontiers in Bioscience-Landmark*, 13(5), pp. 1857–1865. Available at: https://doi.org/10.2741/2805.

Bill, R.M. and Hedfalk, K. (2021) 'Aquaporins – Expression, purification and characterization', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1863(9), p. 183650. Available at: https://doi.org/10.1016/j.bbamem.2021.183650.

Bill, R.M., Henderson, P.J.F., Iwata, S., Kunji, E.R.S., Michel, H., Neutze, R., Newstead, S., Poolman, B., Tate, C.G. and Vogel, H. (2011) 'Overcoming barriers to membrane protein structure determination', *Nature Biotechnology*, 29(4), pp. 335–340. Available at: https://doi.org/10.1038/nbt.1833.

Binkley, N., Bolognese, M., Sidorowicz-Bialynicka, A., Vally, T., Trout, R., Miller, C., Buben, C.E., Gilligan, J.P., Krause, D.S., and Oral Calcitonin in Postmenopausal Osteoporosis (ORACAL) Investigators (2012) 'A phase 3 trial of the efficacy and safety of oral recombinant calcitonin: the Oral Calcitonin in Postmenopausal Osteoporosis (ORACAL) trial', *Journal of Bone and Mineral Research: The Official Journal of the American Society for Bone and Mineral Research*, 27(8), pp. 1821–1829. Available at: https://doi.org/10.1002/jbmr.1602.

Birch, J., Axford, D., Foadi, J., Meyer, A., Eckhardt, A., Thielmann, Y. and Moraes, I. (2018) 'The fine art of integral membrane protein crystallisation', *Methods (San Diego, Calif.)*, 147, pp. 150–162. Available at: https://doi.org/10.1016/j.ymeth.2018.05.014.

Birch, J., Cheruvara, H., Gamage, N., Harrison, P.J., Lithgo, R. and Quigley, A. (2020) 'Changes in Membrane Protein Structural Biology', *Biology*, 9(11), p. 401. Available at: https://doi.org/10.3390/biology9110401.

Bischof, J., Maeda, R.K., Hediger, M., Karch, F. and Basler, K. (2007) 'An optimized transgenesis system for *Drosophila* using germ-line-specific φC31 integrases', *Proceedings of the National Academy of Sciences*, 104(9), pp. 3312–3317. Available at: https://doi.org/10.1073/pnas.0611511104.

Blin, N., Yun, J. and Wess, J. (1995) 'Mapping of single amino acid residues required for selective activation of Gq/11 by the m3 muscarinic acetylcholine receptor', *The Journal of Biological Chemistry*, 270(30), pp. 17741–17748. Available at: https://doi.org/10.1074/jbc.270.30.17741.

Bockaert, J. and Pin, J.P. (1999) 'Molecular tinkering of G protein-coupled receptors: an evolutionary success', *The EMBO journal*, 18(7), pp. 1723–1729. Available at: https://doi.org/10.1093/emboj/18.7.1723.

Böer, E., Steinborn, G., Kunze, G. and Gellissen, G. (2007) 'Yeast expression platforms', *Applied Microbiology and Biotechnology*, 77(3), pp. 513–523. Available at: https://doi.org/10.1007/s00253-007-1209-0.

Boguth, C.A., Singh, P., Huang, C. and Tesmer, J.J.G. (2010) 'Molecular basis for activation of G protein-coupled receptor kinases', *The EMBO journal*, 29(19), pp. 3249–3259. Available at: https://doi.org/10.1038/emboj.2010.206.

Boivineau, J., Haffke, M. and Jaakola, V.-P. (2020) 'Membrane Protein Expression in Insect Cells Using the Baculovirus Expression Vector System', *Methods in Molecular Biology (Clifton, N.J.)*, 2127, pp. 63–80. Available at: https://doi.org/10.1007/978-1-0716-0373-4_5.

Bomberger, J.M., Parameswaran, N., Hall, C.S., Aiyar, N. and Spielman, W.S. (2005) 'Novel function for receptor activity-modifying proteins (RAMPs) in post-endocytic receptor trafficking', *The Journal of Biological Chemistry*, 280(10), pp. 9297–9307. Available at: https://doi.org/10.1074/jbc.M413786200.

Bomberger, J.M., Spielman, W.S., Hall, C.S., Weinman, E.J. and Parameswaran, N. (2005) 'Receptor activity-modifying protein (RAMP) isoform-specific regulation of adrenomedullin receptor trafficking by NHERF-1', *The Journal of Biological Chemistry*, 280(25), pp. 23926–23935. Available at: https://doi.org/10.1074/jbc.M501751200.

Bookstein, F.L. (1984) 'A statistical method for biological shape comparisons', *Journal of Theoretical Biology*, 107(3), pp. 475–520. Available at: https://doi.org/10.1016/s0022-5193(84)80104-6.

Bookstein, F.L. (1997) *Morphometric Tools for Landmark Data*, *Morphometric Tools for Landmark Data*. Available at: https://ui.adsabs.harvard.edu/abs/1997mtld.book.....B (Accessed: 19 March 2024).

Bortolato, A., Doré, A.S., Hollenstein, K., Tehan, B.G., Mason, J.S. and Marshall, F.H. (2014) 'Structure of Class B GPCRs: new horizons for drug discovery', *British Journal of Pharmacology*, 171(13), pp. 3132–3145. Available at: https://doi.org/10.1111/bph.12689.

Bouschet, T. and Henley, J.M. (2005) 'Calcium as an extracellular signalling molecule: perspectives on the Calcium Sensing Receptor in the brain', *Comptes rendus biologies*, 328(8), pp. 691–700. Available at: https://doi.org/10.1016/j.crvi.2004.10.006.

Bouschet, T., Martin, S. and Henley, J.M. (2005) 'Receptor-activity-modifying proteins are required for forward trafficking of the calcium-sensing receptor to the plasma membrane', *Journal of cell science*, 118(Pt 20), pp. 4709–4720. Available at: https://doi.org/10.1242/jcs.02598.

Bouschet, T., Martin, S. and Henley, J.M. (2008) 'Regulation of calcium-sensing-receptor trafficking and cell-surface expression by GPCRs and RAMPs', *Trends in pharmacological sciences*, 29(12), pp. 633–639. Available at: https://doi.org/10.1016/j.tips.2008.09.002.

Boutari, C. and Mantzoros, C.S. (2022) 'A 2022 update on the epidemiology of obesity and a call to action: as its twin COVID-19 pandemic appears to be receding, the obesity and dysmetabolism pandemic continues to rage on', *Metabolism: Clinical and Experimental*, 133, p. 155217. Available at: https://doi.org/10.1016/j.metabol.2022.155217.

Bouvier, M., Hausdorff, W.P., De Blasi, A., O'Dowd, B.F., Kobilka, B.K., Caron, M.G. and Lefkowitz, R.J. (1988) 'Removal of phosphorylation sites from the beta 2-adrenergic receptor delays onset of agonist-promoted desensitization', *Nature*, 333(6171), pp. 370–373. Available at: https://doi.org/10.1038/333370a0.

Bower, R.L., Yule, L., Rees, T.A., Deganutti, G., Hendrikse, E.R., Harris, P.W.R., Kowalczyk, R., Ridgway, Z., Wong, A.G., Swierkula, K., Raleigh, D.P., Pioszak, A.A., Brimble, M.A., Reynolds, C.A., Walker, C.S. and Hay, D.L. (2018) 'Molecular Signature for Receptor Engagement in the Metabolic Peptide Hormone Amylin', *ACS Pharmacology & Translational Science*, 1(1), pp. 32–49. Available at: https://doi.org/10.1021/acsptsci.8b00002.

Bracey, M.H., Cravatt, B.F. and Stevens, R.C. (2004) 'Structural commonalities among integral membrane enzymes', *FEBS letters*, 567(2–3), pp. 159–165. Available at: https://doi.org/10.1016/j.febslet.2004.04.084.

Bretherton-Watt, D., Ghatei, M.A., Jamal, H., Gilbey, S.G., Jones, P.M. and Bloom, S.R. (1992) 'The physiology of calcitonin gene-related peptide in the islet compared with that of islet amyloid polypeptide (amylin)', *Annals of the New York Academy of Sciences*, 657, pp. 299–312. Available at: https://doi.org/10.1111/j.1749-6632.1992.tb22777.x.

Broecker, J., Eger, B.T. and Ernst, O.P. (2017) 'Crystallogenesis of Membrane Proteins Mediated by Polymer-Bounded Lipid Nanodiscs', *Structure*, 25(2), pp. 384–392. Available at: https://doi.org/10.1016/j.str.2016.12.004.

Brown, E.M. (2013) 'Role of the calcium-sensing receptor in extracellular calcium homeostasis', *Best Practice & Research. Clinical Endocrinology & Metabolism*, 27(3), pp. 333–343. Available at: https://doi.org/10.1016/j.beem.2013.02.006.

Brown, N.A., Schrevens, S., van Dijck, P. and Goldman, G.H. (2018) 'Fungal G-protein-coupled receptors: mediators of pathogenesis and targets for disease control', *Nature Microbiology*, 3(4), pp. 402–414. Available at: https://doi.org/10.1038/s41564-018-0127-5.

Brzezinski, D., Dauter, Z., Minor, W. and Jaskolski, M. (2020) 'On the evolution of the quality of macromolecular models in the PDB', *The Febs Journal*, 287(13), pp. 2685–2698. Available at: https://doi.org/10.1111/febs.15314.

Bühlmann, N., Leuthäuser, K., Muff, R., Fischer, J.A. and Born, W. (1999) 'A receptor activity modifying protein (RAMP)2-dependent adrenomedullin receptor is a calcitonin gene-related peptide receptor when coexpressed with human RAMP1', *Endocrinology*, 140(6), pp. 2883–2890. Available at: https://doi.org/10.1210/endo.140.6.6783.

Bunnett, N.W. (2006) 'Protease-Activated Receptors: How Proteases Signal to Cells to Cause Inflammation and Pain', *Seminars in Thrombosis and Hemostasis*, 32(S 1), pp. 39–48. Available at: https://doi.org/10.1055/s-2006-939553.

Busson, D. and Pret, A.-M. (2007) 'GAL4/UAS targeted gene expression for studying Drosophila Hedgehog signaling', *Methods in Molecular Biology (Clifton, N.J.)*, 397, pp. 161–201. Available at: https://doi.org/10.1007/978-1-59745-516-9_13.

Byrne, B. (2015) 'Pichia pastoris as an expression host for membrane protein structural biology', *Current Opinion in Structural Biology*, 32, pp. 9–17. Available at: https://doi.org/10.1016/j.sbi.2015.01.005.

Caffrey, M. (2015) 'A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes', *Acta Crystallographica. Section F, Structural Biology Communications*, 71(Pt 1), pp. 3–18. Available at: https://doi.org/10.1107/S2053230X14026843.

Calebiro, D., Nikolaev, V.O., Persani, L. and Lohse, M.J. (2010) 'Signaling by internalized G-protein-coupled receptors', *Trends in Pharmacological Sciences*, 31(5), pp. 221–228. Available at: https://doi.org/10.1016/j.tips.2010.02.002.

Campos, M.M. and Calixto, J.B. (2000) 'Neurokinin mediation of edema and inflammation', *Neuropeptides*, 34(5), pp. 314–322. Available at: https://doi.org/10.1054/npep.2000.0823.

Cao, J., Belousoff, M.J., Liang, Y.-L., Johnson, R.M., Josephs, T.M., Fletcher, M.M., Christopoulos, A., Hay, D.L., Danev, R., Wootten, D. and Sexton, P.M. (2022) 'A structural basis for amylin receptor phenotype', *Science (New York, N.Y.)*, 375(6587), p. eabm9609. Available at: https://doi.org/10.1126/science.abm9609.

Cao, Y. and Prescott, S.M. (2002) 'Many actions of cyclooxygenase-2 in cellular dynamics and in cancer', *Journal of Cellular Physiology*, 190(3), pp. 279–286. Available at: https://doi.org/10.1002/jcp.10068.

Carlesso, A., Delgado, R., Ruiz Isant, O., Uwangue, O., Valli, D., Bill, R.M. and Hedfalk, K. (2022) 'Yeast as a tool for membrane protein production and structure determination', *FEMS yeast research*, 22(1), p. foac047. Available at: https://doi.org/10.1093/femsyr/foac047.

Carpenter, B., Nehmé, R., Warne, T., Leslie, A.G.W. and Tate, C.G. (2016) 'Structure of the adenosine A(2A) receptor bound to an engineered G protein', *Nature*, 536(7614), pp. 104–107. Available at: https://doi.org/10.1038/nature18966.

Carpenter, E.P., Beis, K., Cameron, A.D. and Iwata, S. (2008) 'Overcoming the challenges of membrane protein crystallography', *Current Opinion in Structural Biology*, 18(5), pp. 581–586. Available at: https://doi.org/10.1016/j.sbi.2008.07.001.

Carugo, O. and Pongor, S. (2001) 'A normalized root-mean-spuare distance for comparing protein three-dimensional structures', *Protein Science : A Publication of the Protein Society*, 10(7), pp. 1470–1473.

Cary, B.P., Zhang, X., Cao, J., Johnson, R.M., Piper, S.J., Gerrard, E.J., Wootten, D. and Sexton, P.M. (2023) 'New Insights into the Structure and Function of Class B1 GPCRs', *Endocrine Reviews*, 44(3), pp. 492–517. Available at: https://doi.org/10.1210/endrev/bnac033.

Casares, D., Escribá, P.V. and Rosselló, C.A. (2019) 'Membrane Lipid Composition: Effect on Membrane and Organelle Structure, Function and Compartmentalization and Therapeutic Avenues', *International Journal of Molecular Sciences*, 20(9), p. 2167. Available at: https://doi.org/10.3390/ijms20092167.

Caulfield, M.P. and Birdsall, N.J.M. (1998) 'International Union of Pharmacology. XVII. Classification of Muscarinic Acetylcholine Receptors', *Pharmacological Reviews*, 50(2), pp. 279–290.

Cegla, J., Jones, B.J., Gardiner, J.V., Hodson, D.J., Marjot, T., McGlone, E.R., Tan, T.M. and Bloom, S.R. (2017) 'RAMP2 Influences Glucagon Receptor Pharmacology via Trafficking and Signaling', *Endocrinology*, 158(8), pp. 2680–2693. Available at: https://doi.org/10.1210/en.2016-1755.

Chakraborty, R., Xu, B., Bhullar, R.P. and Chelikani, P. (2015) 'Chapter Twelve - Expression of G Protein-Coupled Receptors in Mammalian Cells', in A.K. Shukla (ed.) *Methods in Enzymology*. Academic Press (Membrane Proteins—Production and Functional Characterization), pp. 267–281. Available at: https://doi.org/10.1016/bs.mie.2014.12.013.

Changeux, J.-P. (2018) 'The nicotinic acetylcholine receptor: a typical "allosteric machine"', *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 373(1749), p. 20170174. Available at: https://doi.org/10.1098/rstb.2017.0174.

Charlton, J. (2016) *Solubilisation and characterisation of G-protein-coupled receptors using styrene maleic acid polymer*. d_ph. University of Birmingham. Available at: https://etheses.bham.ac.uk/id/eprint/6525/ (Accessed: 17 January 2023).

Che, T., Dwivedi-Agnihotri, H., Shukla, A.K. and Roth, B.L. (2021) 'Biased ligands at opioid receptors: Current status and future directions', *Science signaling*, 14(677), p. eaav0320. Available at: https://doi.org/10.1126/scisignal.aav0320.

Chen, C.A. and Manning, D.R. (2001) 'Regulation of G proteins by covalent modification', *Oncogene*, 20(13), pp. 1643–1652. Available at: https://doi.org/10.1038/sj.onc.1204185.

Chen, G., Xu, J., Inoue, A., Schmidt, M.F., Bai, C., Lu, Q., Gmeiner, P., Liu, Z. and Du, Y. (2022) 'Activation and allosteric regulation of the orphan GPR88-Gi1 signaling complex', *Nature Communications*, 13(1), p. 2375. Available at: https://doi.org/10.1038/s41467-022-30081-5.

Chen, H., Shaffer, P.L., Huang, X. and Rose, P.E. (2013) 'Rapid screening of membrane protein expression in transiently transfected insect cells', *Protein Expression and Purification*, 88(1), pp. 134–142. Available at: https://doi.org/10.1016/j.pep.2012.12.003.

Chen, J. and Siu, S.W.I. (2020) 'Machine Learning Approaches for Quality Assessment of Protein Structures', *Biomolecules*, 10(4), p. 626. Available at: https://doi.org/10.3390/biom10040626.

Chen, Q., Perry, N.A., Vishnivetskiy, S.A., Berndt, S., Gilbert, N.C., Zhuo, Y., Singh, P.K., Tholen, J., Ohi, M.D., Gurevich, E.V., Brautigam, C.A., Klug, C.S., Gurevich, V.V. and Iverson, T.M. (2017) 'Structural basis of arrestin-3 activation and signaling', *Nature Communications*, 8(1), p. 1427. Available at: https://doi.org/10.1038/s41467-017-01218-8.

Chen, X., Lou, G., Meng, Z. and Huang, W. (2011) 'TGR5: A Novel Target for Weight Maintenance and Glucose Metabolism', *Journal of Diabetes Research*, 2011, p. e853501. Available at: https://doi.org/10.1155/2011/853501.

Chen, X., Wang, L., Cui, Q., Ding, Z., Han, L., Kou, Y., Zhang, W., Wang, H., Jia, X., Dai, M., Shi, Z., Li, Y., Li, X. and Geng, Y. (2021) 'Structural insights into the activation of human calcium-sensing receptor', *eLife*, 10, p. e68578. Available at: https://doi.org/10.7554/eLife.68578.

Chen, Z., Singer, W.D., Sternweis, P.C. and Sprang, S.R. (2005) 'Structure of the p115RhoGEF rgRGS domain–Gα13/i1 chimera complex suggests convergent evolution of a GTPase activator', *Nature Structural & Molecular Biology*, 12(2), pp. 191–197. Available at: https://doi.org/10.1038/nsmb888.

Cheng, A., van Hoek, A.N., Yeager, M., Verkman, A.S. and Mitra, A.K. (1997) 'Three-dimensional organization of a human water channel', *Nature*, 387(6633), pp. 627–630. Available at: https://doi.org/10.1038/42517.

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. and Stevens, R.C. (2007) 'High Resolution Crystal Structure of an Engineered Human β2-Adrenergic G protein-Coupled Receptor', *Science (New York, N.Y.)*, 318(5854), p. 1258. Available at: https://doi.org/10.1126/science.1150577.

Chhikara, B.S. and Parang, K. (2023) 'Global Cancer Statistics 2022: the trends projection analysis', *Chemical Biology Letters*, 10(1), pp. 451–451. Available at: https://scholar.google.com/scholar?q=urn:nbn:sciencein.cbl.2023.v10.451.

Choe, H.-W., Kim, Y.J., Park, J.H., Morizumi, T., Pai, E.F., Krauß, N., Hofmann, K.P., Scheerer, P. and Ernst, O.P. (2011) 'Crystal structure of metarhodopsin II', *Nature*, 471(7340), pp. 651–655. Available at: https://doi.org/10.1038/nature09789.

Choi, B.-K., Bobrowicz, P., Davidson, R.C., Hamilton, S.R., Kung, D.H., Li, H., Miele, R.G., Nett, J.H., Wildt, S. and Gerngross, T.U. (2003) 'Use of combinatorial genetic libraries to

humanize N-linked glycosylation in the yeast Pichia pastoris', *Proceedings of the National Academy of Sciences*, 100(9), pp. 5022–5027. Available at: https://doi.org/10.1073/pnas.0931263100.

Choksi, T., Hay, D.L., Legon, S., Poyner, D.R., Hagner, S., Bloom, S.R. and Smith, D.M. (2002) 'Comparison of the expression of calcitonin receptor-like receptor (CRLR) and receptor activity modifying proteins (RAMPs) with CGRP and adrenomedullin binding in cell lines', *British Journal of Pharmacology*, 136(5), pp. 784–792. Available at: https://doi.org/10.1038/sj.bjp.0704761.

Chrencik, J.E., Roth, C.B., Terakado, M., Kurata, H., Omi, R., Kihara, Y., Warshaviak, D., Nakade, S., Asmar-Rovira, G., Mileni, M., Mizuno, H., Griffith, M.T., Rodgers, C., Han, G.W., Velasquez, J., Chun, J., Stevens, R.C. and Hanson, M.A. (2015) 'Crystal Structure of Antagonist Bound Human Lysophosphatidic Acid Receptor 1', *Cell*, 161(7), pp. 1633–1643. Available at: https://doi.org/10.1016/j.cell.2015.06.002.

Christopher, J.A., Aves, S.J., Bennett, K.A., Doré, A.S., Errey, J.C., Jazayeri, A., Marshall, F.H., Okrasa, K., Serrano-Vega, M.J., Tehan, B.G., Wiggin, G.R. and Congreve, M. (2015) 'Fragment and Structure-Based Drug Discovery for a Class C GPCR: Discovery of the mGlu5 Negative Allosteric Modulator HTL14242 (3-Chloro-5-[6-(5-fluoropyridin-2-yl)pyrimidin-4-yl]benzonitrile)', *Journal of Medicinal Chemistry*, 58(16), pp. 6653–6664. Available at: https://doi.org/10.1021/acs.jmedchem.5b00892.

Christopher, J.A., Orgován, Z., Congreve, M., Doré, A.S., Errey, J.C., Marshall, F.H., Mason, J.S., Okrasa, K., Rucktooa, P., Serrano-Vega, M.J., Ferenczy, G.G. and Keserű, G.M. (2019) 'Structure-Based Optimization Strategies for G Protein-Coupled Receptor (GPCR) Allosteric Modulators: A Case Study from Analyses of New Metabotropic Glutamate Receptor 5 (mGlu5) X-ray Structures', *Journal of Medicinal Chemistry*, 62(1), pp. 207–222. Available at: https://doi.org/10.1021/acs.jmedchem.7b01722.

Christopoulos, A., Christopoulos, G., Morfis, M., Udawela, M., Laburthe, M., Couvineau, A., Kuwasako, K., Tilakaratne, N. and Sexton, P.M. (2003) 'Novel receptor partners and function of receptor activity-modifying proteins', *The Journal of Biological Chemistry*, 278(5), pp. 3293–3297. Available at: https://doi.org/10.1074/jbc.C200629200.

Chung, K.Y., Kim, T.H., Manglik, A., Alvares, R., Kobilka, B.K. and Prosser, R.S. (2012) 'Role of Detergents in Conformational Exchange of a G Protein-coupled Receptor', *The Journal of Biological Chemistry*, 287(43), pp. 36305–36311. Available at: https://doi.org/10.1074/jbc.M112.406371.

Clark, L.J., Krieger, J., White, A.D., Bondarenko, V., Lei, S., Fang, F., Lee, J.Y., Doruker, P., Böttke, T., Jean-Alphonse, F., Tang, P., Gardella, T.J., Xiao, K., Sutkeviciute, I., Coin, I., Bahar, I. and Vilardaga, J.-P. (2020) 'Allosteric interactions in the parathyroid hormone GPCR–arrestin complex formation', *Nature chemical biology*, 16(10), pp. 1096–1104. Available at: https://doi.org/10.1038/s41589-020-0567-0.

Cohn, J.N. and Tognoni, G. (2001) 'A Randomized Trial of the Angiotensin-Receptor Blocker Valsartan in Chronic Heart Failure', *New England Journal of Medicine*, 345(23), pp. 1667–1675. Available at: https://doi.org/10.1056/NEJMoa010713.

Cong, Z., Liang, Y.-L., Zhou, Q., Darbalaei, S., Zhao, F., Feng, W., Zhao, L., Xu, H.E., Yang, D. and Wang, M.-W. (2022) 'Structural perspective of class B1 GPCR signaling', *Trends in Pharmacological Sciences*, 43(4), pp. 321–334. Available at: https://doi.org/10.1016/j.tips.2022.01.002.

Congreve, M., de Graaf, C., Swain, N.A. and Tate, C.G. (2020) 'Impact of GPCR Structures on Drug Discovery', *Cell*, 181(1), pp. 81–91. Available at: https://doi.org/10.1016/j.cell.2020.03.003.

Cooper, G.M. (2000) 'Transport of Small Molecules', *The Cell: A Molecular Approach. 2nd edition* [Preprint]. Available at: https://www.ncbi.nlm.nih.gov/books/NBK9847/ (Accessed: 1 March 2023).

Cottrell, G.S., Padilla, B., Pikios, S., Roosterman, D., Steinhoff, M., Grady, E.F. and Bunnett, N.W. (2007) 'Post-endocytic Sorting of Calcitonin Receptor-like Receptor and Receptor Activity-modifying Protein 1*', *Journal of Biological Chemistry*, 282(16), pp. 12260–12271. Available at: https://doi.org/10.1074/jbc.M606338200.

Coura, J.R. and Borges-Pereira, J. (2012) 'Chagas disease: What is known and what should be improved: a systemic review', *Revista da Sociedade Brasileira de Medicina Tropical*, 45, pp. 286–296. Available at: https://doi.org/10.1590/S0037-86822012000300002.

Couvineau, A. and Laburthe, M. (2012) 'The family B1 GPCR: structural aspects and interaction with accessory proteins', *Current Drug Targets*, 13(1), pp. 103–115. Available at: https://doi.org/10.2174/138945012798868434.

Crilly, S.E. and Puthenveedu, M.A. (2021) 'Compartmentalized GPCR Signaling from Intracellular Membranes', *The Journal of Membrane Biology*, 254(3), pp. 259–271. Available at: https://doi.org/10.1007/s00232-020-00158-7.

Crine, S.L. and Acharya, K.R. (2022) 'Molecular basis of C-mannosylation - a structural perspective', *The FEBS journal*, 289(24), pp. 7670–7687. Available at: https://doi.org/10.1111/febs.16265.

Croop, R., Madonia, J., Conway, C., Thiry, A., Forshaw, M., Murphy, A., Jensen, C., Dubowchik, G., Coric, V. and Lipton, R. (2021) 'Intranasal Zavegepant is Effective and Well Tolerated for the Acute Treatment of Migraine: A Phase 2/3 Dose-Ranging Clinical trial (4976)', *Neurology*, 96(15 Supplement). Available at: https://n.neurology.org/content/96/15_Supplement/4976 (Accessed: 2 June 2023).

Cross, T.A., Ekanayake, V., Paulino, J. and Wright, A. (2014) 'Solid state NMR: The essential technology for helical membrane protein structural characterization', *Journal of magnetic resonance (San Diego, Calif.: 1997)*, 239, pp. 100–109. Available at: https://doi.org/10.1016/j.jmr.2013.12.006.

Cuesta, A.M., Albiñana, V., Gallardo-Vara, E., Recio-Poveda, L., de Rojas-P, I., de Las Heras, K.V.G., Aguirre, D.T. and Botella, L.M. (2019) 'The β2-adrenergic receptor antagonist ICI-118,551 blocks the constitutively activated HIF signalling in hemangioblastomas from von Hippel-Lindau disease', *Scientific Reports*, 9(1), p. 10062. Available at: https://doi.org/10.1038/s41598-019-46448-6.

Danev, R., Belousoff, M., Liang, Y.-L., Zhang, X., Eisenstein, F., Wootten, D. and Sexton, P.M. (2021) 'Routine sub-2.5 Å cryo-EM structure determination of GPCRs', *Nature Communications*, 12(1), p. 4333. Available at: https://doi.org/10.1038/s41467-021-24650-3.

Dawaliby, R., Trubbia, C., Delporte, C., Masureel, M., Van Antwerpen, P., Kobilka, B.K. and Govaerts, C. (2016) 'Allosteric regulation of G protein–coupled receptor activity by phospholipids', *Nature Chemical Biology*, 12(1), pp. 35–39. Available at: https://doi.org/10.1038/nchembio.1960.

De Ambrogi, M., Volpe, S. and Tamanini, C. (2003) 'Ghrelin: central and peripheral effects of a novel peptydil hormone', *Medical science monitor*, 9(9), pp. RA217-24.

Deganutti, G., Atanasio, S., Rujan, R.-M., Sexton, P.M., Wootten, D. and Reynolds, C.A. (2021) 'Exploring Ligand Binding to Calcitonin Gene-Related Peptide Receptors', *Frontiers in Molecular Biosciences*, 8, p. 720561. Available at: https://doi.org/10.3389/fmolb.2021.720561.

Delaitre, C., Boisbrun, M., Lecat, S. and Dupuis, F. (2021) 'Targeting the Angiotensin II Type 1 Receptor in Cerebrovascular Diseases: Biased Signaling Raises New Hopes', *International Journal of Molecular Sciences*, 22(13), p. 6738. Available at: https://doi.org/10.3390/ijms22136738.

Delmar, J.A., Bolla, J.R., Su, C.-C. and Yu, E.W. (2015) 'Crystallization of Membrane Proteins by Vapor Diffusion', *Methods in enzymology*, 557, pp. 363–392. Available at: https://doi.org/10.1016/bs.mie.2014.12.018.

Delmas, P., Parpaite, T. and Coste, B. (2022) 'PIEZO channels and newcomers in the mammalian mechanosensitive ion channel family', *Neuron*, 110(17), pp. 2713–2727. Available at: https://doi.org/10.1016/j.neuron.2022.07.001.

Depping, P., Román Lara, M.M., Kesidis, A., Bill, R.M., Rothnie, A.J., Browning, D.F. and Goddard, A.D. (2022) 'Heterologous Expression of Membrane Proteins in E. coli', *Methods in Molecular Biology (Clifton, N.J.)*, 2507, pp. 59–78. Available at: https://doi.org/10.1007/978-1-0716-2368-8_4.

Deshpande, A.D., Harris-Hayes, M. and Schootman, M. (2008) 'Epidemiology of Diabetes and Diabetes-Related Complications', *Physical Therapy*, 88(11), pp. 1254–1264. Available at: https://doi.org/10.2522/ptj.20080020.

Deshpande, I., Liang, J., Hedeen, D., Roberts, K.J., Zhang, Y., Ha, B., Latorraca, N.R., Faust, B., Dror, R.O., Beachy, P.A., Myers, B.R. and Manglik, A. (2019) 'Smoothened stimulation by membrane sterols drives Hedgehog pathway activity', *Nature*, 571(7764), pp. 284–288. Available at: https://doi.org/10.1038/s41586-019-1355-4.

DeVree, B.T., Mahoney, J.P., Vélez-Ruiz, G.A., Rasmussen, S.G.F., Kuszak, A.J., Edwald, E., Fung, J.-J., Manglik, A., Masureel, M., Du, Y., Matt, R.A., Pardon, E., Steyaert, J., Kobilka, B.K. and Sunahara, R.K. (2016) 'Allosteric coupling from G protein to the agonist binding pocket in GPCRs', *Nature*, 535(7610), pp. 182–186. Available at: https://doi.org/10.1038/nature18324.

Dill, K.A., Ozkan, S.B., Shell, M.S. and Weikl, T.R. (2008) 'The protein folding problem', *Annual Review of Biophysics*, 37, pp. 289–316. Available at: https://doi.org/10.1146/annurev.biophys.37.092707.153558.

Dilworth, M.V., Piel, M.S., Bettaney, K.E., Ma, P., Luo, J., Sharples, D., Poyner, D.R., Gross, S.R., Moncoq, K., Henderson, P.J.F., Miroux, B. and Bill, R.M. (2018) 'Microbial expression systems for membrane proteins', *Methods*, 147, pp. 3–39. Available at: https://doi.org/10.1016/j.ymeth.2018.04.009.

Ding, Y., Fujimoto, L.M., Yao, Y., Plano, G.V. and Marassi, F.M. (2015) 'Influence of the lipid membrane environment on structure and activity of the outer membrane protein Ail from Yersinia pestis', *Biochimica et biophysica acta*, 1848(2), pp. 712–720. Available at: https://doi.org/10.1016/j.bbamem.2014.11.021.

Dirnberger, B., Korona, D., Popovic, R., Deery, M.J., Barber, H., Russell, S. and Lilley, K.S. (2023) 'Enrichment of Membrane Proteins for Downstream Analysis Using Styrene Maleic Acid Lipid Particles (SMALPs) Extraction', *Bio-protocol*, 13(15), p. e4728. Available at: https://doi.org/10.21769/BioProtoc.4728.

Dobson, L., Reményi, I. and Tusnády, G.E. (2015) 'The human transmembrane proteome', *Biology Direct*, 10, p. 31. Available at: https://doi.org/10.1186/s13062-015-0061-x.

Domagalski, M.J., Zheng, H., Zimmerman, M.D., Dauter, Z., Wlodawer, A. and Minor, W. (2014) 'The Quality and Validation of Structures from Structural Genomics', *Methods in molecular biology (Clifton, N.J.)*, 1091, pp. 297–314. Available at: https://doi.org/10.1007/978-1-62703-691-7_21.

Doré, A.S., Okrasa, K., Patel, J.C., Serrano-Vega, M., Bennett, K., Cooke, R.M., Errey, J.C., Jazayeri, A., Khan, S., Tehan, B., Weir, M., Wiggin, G.R. and Marshall, F.H. (2014) 'Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain', *Nature*, 511(7511), pp. 557–562. Available at: https://doi.org/10.1038/nature13396.

Doré, A.S., Robertson, N., Errey, J.C., Ng, I., Hollenstein, K., Tehan, B., Hurrell, E., Bennett, K., Congreve, M., Magnani, F., Tate, C.G., Weir, M. and Marshall, F.H. (2011) 'Structure of the Adenosine A2A Receptor in Complex with ZM241385 and the Xanthines XAC and Caffeine', *Structure*, 19(9), pp. 1283–1293. Available at: https://doi.org/10.1016/j.str.2011.06.014.

Dörr, J.M., Koorengevel, M.C., Schäfer, M., Prokofyev, A.V., Scheidelaar, S., van der Cruijsen, E.A.W., Dafforn, T.R., Baldus, M. and Killian, J.A. (2014) 'Detergent-free isolation, characterization, and functional reconstitution of a tetrameric K+ channel: The power of native nanodiscs', *Proceedings of the National Academy of Sciences of the United States of America*, 111(52), pp. 18607–18612. Available at: https://doi.org/10.1073/pnas.1416205112.

Druey, K.M., Blumer, K.J., Kang, V.H. and Kehrl, J.H. (1996) 'Inhibition of G-protein-mediated MAP kinase activation by a new mammalian gene family', *Nature*, 379(6567), pp. 742–746. Available at: https://doi.org/10.1038/379742a0.

Dryden, I.L. and Mardia, K.V. (2016) *Statistical Shape Analysis: With Applications in R*. John Wiley & Sons.

Dubochet, J., Adrian, M., Chang, J.-J., Homo, J.-C., Lepault, J., McDowall, A.W. and Schultz, P. (1988) 'Cryo-electron microscopy of vitrified specimens', *Quarterly Reviews of Biophysics*, 21(2), pp. 129–228. Available at: https://doi.org/10.1017/S0033583500004297.

Duffy, J.B. (2002) 'GAL4 system in drosophila: A fly geneticist's swiss army knife', *genesis*, 34(1–2), pp. 1–15. Available at: https://doi.org/10.1002/gene.10150.

Dunwiddie, T.V. and Masino, S.A. (2001) 'The role and regulation of adenosine in the central nervous system', *Annual Review of Neuroscience*, 24, pp. 31–55. Available at: https://doi.org/10.1146/annurev.neuro.24.1.31.

Dutta, S.K., Yao, Y. and Marassi, F.M. (2017) 'Structural Insights into the Yersinia pestis Outer Membrane Protein Ail in Lipid Bilayers', *The journal of physical chemistry. B*, 121(32), pp. 7561–7570. Available at: https://doi.org/10.1021/acs.jpcb.7b03941.

Edvinsson, L., Haanes, K.A., Warfvinge, K. and Krause, D.N. (2018) 'CGRP as the target of new migraine therapies - successful translation from bench to clinic', *Nature Reviews. Neurology*, 14(6), pp. 338–350. Available at: https://doi.org/10.1038/s41582-018-0003-1.

Eglen, R.M. and Reisine, T. (2009) 'New Insights into GPCR Function: Implications for HTS', in W.R. Leifert (ed.) *G Protein-Coupled Receptors in Drug Discovery*. Totowa, NJ: Humana Press (Methods in Molecular Biology), pp. 1–13. Available at: https://doi.org/10.1007/978-1-60327-317-6_1.

Ehses, J.A., Lee, S.S.T., Pederson, R.A. and McIntosh, C.H.S. (2001) 'A New Pathway for Glucose-dependent Insulinotropic Polypeptide (GIP) Receptor Signaling: EVIDENCE FOR THE INVOLVEMENT OF PHOSPHOLIPASE A2 IN GIP-STIMULATED INSULIN SECRETION *', *Journal of Biological Chemistry*, 276(26), pp. 23667–23673. Available at: https://doi.org/10.1074/jbc.M103023200.

Elste, A.P. and Petersen, I. (2010) 'Expression of proteinase-activated receptor 1-4 (PAR 1-4) in human cancer', *Journal of Molecular Histology*, 41(2), pp. 89–99. Available at: https://doi.org/10.1007/s10735-010-9274-6.

Emmerstorfer-Augustin, A., Wriessnegger, T., Hirz, M., Zellnig, G. and Pichler, H. (2019) 'Membrane Protein Production in Yeast: Modification of Yeast Membranes for Human Membrane Protein Production', in B. Gasser and D. Mattanovich (eds) *Recombinant Protein Production in Yeast*. New York, NY: Springer (Methods in Molecular Biology), pp. 265–285. Available at: https://doi.org/10.1007/978-1-4939-9024-5_12.

Engelman, D.M., Steitz, T.A. and Goldman, A. (1986) 'Identifying Nonpolar Transbilayer Helices in Amino Acid Sequences of Membrane Proteins', *Annual Review of Biophysics and Biophysical Chemistry*, 15(1), pp. 321–353. Available at: https://doi.org/10.1146/annurev.bb.15.060186.001541.

Eroglu, Ç., Cronet, P., Panneels, V., Beaufils, P. and Sinning, I. (2002) 'Functional reconstitution of purified metabotropic glutamate receptor expressed in the fly eye', *EMBO reports*, 3(5), pp. 491–496. Available at: https://doi.org/10.1093/embo-reports/kvf088.

Errasti-Murugarren, E., Bartoccioni, P. and Palacín, M. (2021) 'Membrane Protein Stabilization Strategies for Structural and Functional Studies', *Membranes*, 11(2), p. 155. Available at: https://doi.org/10.3390/membranes11020155.

Exton, J.H. (1996) 'Regulation of Phosphoinositide Phospholipases by Hormones, Neurotransmitters, and Other Agonists Linked to G Proteins', *Annual Review of Pharmacology and Toxicology*, 36(1), pp. 481–509. Available at: https://doi.org/10.1146/annurev.pa.36.040196.002405.

Fagerberg, L., Jonasson, K., von Heijne, G., Uhlén, M. and Berglund, L. (2010) 'Prediction of the human membrane proteome', *PROTEOMICS*, 10(6), pp. 1141–1149. Available at: https://doi.org/10.1002/pmic.200900258.

Faham, S. and Bowie, J.U. (2002) 'Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure11Edited by D. Rees', *Journal of Molecular Biology*, 316(1), pp. 1–6. Available at: https://doi.org/10.1006/jmbi.2001.5295.

Fairhead, M. and Howarth, M. (2015) 'Site-specific biotinylation of purified proteins using BirA', *Methods in molecular biology (Clifton, N.J.)*, 1266, pp. 171–184. Available at: https://doi.org/10.1007/978-1-4939-2272-7_12.

Farrens, D.L., Altenbach, C., Yang, K., Hubbell, W.L. and Khorana, H.G. (1996) 'Requirement of Rigid-Body Motion of Transmembrane Helices for Light Activation of Rhodopsin', *Science*, 274(5288), pp. 768–770. Available at: https://doi.org/10.1126/science.274.5288.768.

Fatima, N., Patel, S.N. and Hussain, T. (2021) 'Angiotensin II Type 2 Receptor', *Hypertension (Dallas, Tex.: 1979)*, 77(6), pp. 1845–1856. Available at: https://doi.org/10.1161/HYPERTENSIONAHA.120.11941.

Feng, X.-T., Leng, J., Xie, Z., Li, S.-L., Zhao, W. and Tang, Q.-L. (2012) 'GPR40: A therapeutic target for mediating insulin secretion (Review)', *International Journal of Molecular Medicine*, 30(6), pp. 1261–1266. Available at: https://doi.org/10.3892/ijmm.2012.1142.

Ferguson, S.S., Downey, W.E., Colapietro, A.M., Barak, L.S., Ménard, L. and Caron, M.G. (1996) 'Role of beta-arrestin in mediating agonist-promoted G protein-coupled receptor internalization', *Science (New York, N.Y.)*, 271(5247), pp. 363–366. Available at: https://doi.org/10.1126/science.271.5247.363.

Ferrero, M.E. (2011) 'Purinoceptors in inflammation: potential as anti-inflammatory therapeutic targets', *Frontiers in Bioscience*, 16(1), p. 2172. Available at: https://doi.org/10.2741/3846.

Findlay, D.M. and Sexton, P.M. (2004) 'Calcitonin', *Growth Factors (Chur, Switzerland)*, 22(4), pp. 217–224. Available at: https://doi.org/10.1080/08977190410001728033.

Fischgräbe, J., Götte, M., Michels, K., Kiesel, L. and Wülfing, P. (2010) 'Targeting endothelin A receptor enhances anti-proliferative and anti-invasive effects of the HER2 antibody trastuzumab in HER2-overexpressing breast cancer cells', *International Journal of Cancer*, 127(3), pp. 696–706. Available at: https://doi.org/10.1002/ijc.25076.

Flahaut, M., Rossier, B.C. and Firsov, D. (2002) 'Respective roles of calcitonin receptor-like receptor (CRLR) and receptor activity-modifying proteins (RAMP) in cell surface expression of CRLR/RAMP heterodimeric receptors', *The Journal of Biological Chemistry*, 277(17), pp. 14731–14737. Available at: https://doi.org/10.1074/jbc.M112084200.

Flühmann, B., Lauber, M., Lichtensteiger, W., Fischer, J.A. and Born, W. (1997) 'Tissue-specific mRNA expression of a calcitonin receptor-like receptor during fetal and postnatal development', *Brain Research*, 774(1–2), pp. 184–192. Available at: https://doi.org/10.1016/s0006-8993(97)81702-7.

Flühmann, B., Muff, R., Hunziker, W., Fischer, J.A. and Born, W. (1995) 'A human orphan calcitonin receptor-like structure', *Biochemical and Biophysical Research Communications*, 206(1), pp. 341–347. Available at: https://doi.org/10.1006/bbrc.1995.1047.

Ford, C.E., Skiba, N.P., Bae, H., Daaka, Y., Reuveny, E., Shekter, L.R., Rosal, R., Weng, G., Yang, C.-S., Iyengar, R., Miller, R.J., Jan, L.Y., Lefkowitz, R.J. and Hamm, H.E. (1998) 'Molecular Basis for Interactions of G Protein βγ Subunits with Effectors', *Science*, 280(5367), pp. 1271–1274. Available at: https://doi.org/10.1126/science.280.5367.1271.

Forster, Th. (1946) 'Energiewanderung und Fluoreszenz', *Naturwissenschaften*, 33(6), pp. 166–175. Available at: https://doi.org/10.1007/BF00585226.

Franco, R., Lillo, A., Navarro, G. and Reyes-Resina, I. (2022) 'The adenosine A2A receptor is a therapeutic target in neurological, heart and oncogenic diseases', *Expert Opinion on Therapeutic Targets*, 26(9), pp. 791–800. Available at: https://doi.org/10.1080/14728222.2022.2136570.

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., Garoff, H., Löw, C., Cheng, Y. and Nordlund, P. (2016) 'A saposin-lipoprotein nanoparticle system for membrane proteins', *Nature Methods*, 13(4), pp. 345–351. Available at: https://doi.org/10.1038/nmeth.3801.

Fredriksson, R., Lagerström, M.C., Lundin, L.-G. and Schiöth, H.B. (2003) 'The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogon Groups, and Fingerprints', *Molecular Pharmacology*, 63(6), pp. 1256–1272. Available at: https://doi.org/10.1124/mol.63.6.1256.

Fredriksson, R. and Schiöth, H.B. (2005) 'The repertoire of G-protein-coupled receptors in fully sequenced genomes', *Molecular Pharmacology*, 67(5), pp. 1414–1425. Available at: https://doi.org/10.1124/mol.104.009001.

Fujimoto, W.Y., Ensinck, J.W., Merchant, F.W., Williams, R.H., Smith, P.H. and Johnson, D.G. (1978) 'Stimulation by Gastric Inhibitory Polypeptide of Insulin and Glucagon Secretion by Rat Islet Cultures', *Proceedings of the Society for Experimental Biology and Medicine*, 157(1), pp. 89–93. Available at: https://doi.org/10.3181/00379727-157-39997.

Fujiyoshi, Y., Mitsuoka, K., de Groot, B.L., Philippsen, A., Grubmüller, H., Agre, P. and Engel, A. (2002) 'Structure and Function of Water Channels', *Current Opinion in Structural Biology*, 12(4), pp. 509–515. Available at: https://doi.org/10.1016/S0959-440X(02)00355-X.

Gadsby, D.C., Vergani, P. and Csanády, L. (2006) 'The ABC protein turned chloride channel whose failure causes cystic fibrosis', *Nature*, 440(7083), pp. 477–483. Available at: https://doi.org/10.1038/nature04712.

Garavito, R.M. and Mulichak, A.M. (2003) 'The structure of mammalian cyclooxygenases', *Annual Review of Biophysics and Biomolecular Structure*, 32, pp. 183–206. Available at: https://doi.org/10.1146/annurev.biophys.32.110601.141906.

García-Nafría, J. and Tate, C.G. (2021) 'Structure determination of GPCRs: cryo-EM compared with X-ray crystallography', *Biochemical Society Transactions*, 49(5), pp. 2345–2355. Available at: https://doi.org/10.1042/BST20210431.

Gautier, A., Mott, H.R., Bostock, M.J., Kirkpatrick, J.P. and Nietlispach, D. (2010) 'Structure determination of the seven-helical transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy', *Nature structural & molecular biology*, 17(6), pp. 768–774. Available at: https://doi.org/10.1038/nsmb.1807.

Geisse, S. and Fux, C. (2009) 'Chapter 15 Recombinant Protein Production by Transient Gene Transfer into Mammalian Cells', in R.R. Burgess and M.P. Deutscher (eds) *Methods in Enzymology*. Academic Press (Guide to Protein Purification, 2nd Edition), pp. 223–238. Available at: https://doi.org/10.1016/S0076-6879(09)63015-9.

Gelmini, S., Mangoni, M., Serio, M., Romagnani, P. and Lazzeri, E. (2008) 'The critical role of SDF-1/CXCR4 axis in cancer and cancer stem cells metastasis', *Journal of Endocrinological Investigation*, 31(9), pp. 809–819. Available at: https://doi.org/10.1007/BF03349262.

Geng, Y., Bush, M., Mosyak, L., Wang, F. and Fan, Q.R. (2013) 'Structural mechanism of ligand activation in human GABAB receptor', *Nature*, 504(7479), pp. 254–259. Available at: https://doi.org/10.1038/nature12725.

Geng, Y., Mosyak, L., Kurinov, I., Zuo, H., Sturchler, E., Cheng, T.C., Subramanyam, P., Brown, A.P., Brennan, S.C., Mun, H.-C., Bush, M., Chen, Y., Nguyen, T.X., Cao, B., Chang, D.D., Quick, M., Conigrave, A.D., Colecraft, H.M., McDonald, P. and Fan, Q.R. (2016) 'Structural mechanism of ligand activation in human calcium-sensing receptor', *eLife*, 5, p. e13662. Available at: https://doi.org/10.7554/eLife.13662.

George, A.M. and Jones, P.M. (2012) 'Perspectives on the structure–function of ABC transporters: The Switch and Constant Contact Models', *Progress in Biophysics and Molecular Biology*, 109(3), pp. 95–107. Available at: https://doi.org/10.1016/j.pbiomolbio.2012.06.003.

Gest, H. (2004) 'The discovery of microorganisms by Robert Hooke and Antoni van Leeuwenhoek, Fellows of The Royal Society', *Notes and Records of the Royal Society of London*, 58(2), pp. 187–201. Available at: https://doi.org/10.1098/rsnr.2004.0055.

Gesty-Palmer, D., Chen, M., Reiter, E., Ahn, S., Nelson, C.D., Wang, S., Eckhardt, A.E., Cowan, C.L., Spurney, R.F., Luttrell, L.M. and Lefkowitz, R.J. (2006) 'Distinct beta-arrestin- and G protein-dependent pathways for parathyroid hormone receptor-stimulated ERK1/2 activation', *The Journal of Biological Chemistry*, 281(16), pp. 10856–10864. Available at: https://doi.org/10.1074/jbc.M513380200.

Gesty-Palmer, D., Flannery, P., Yuan, L., Corsino, L., Spurney, R., Lefkowitz, R.J. and Luttrell, L.M. (2009) 'A β-Arrestin–Biased Agonist of the Parathyroid Hormone Receptor (PTH1R) Promotes Bone Formation Independent of G Protein Activation', *Science Translational Medicine*, 1(1), pp. 1ra1-1ra1. Available at: https://doi.org/10.1126/scitranslmed.3000071.

Gether, U. and Kobilka, B.K. (1998) 'G Protein-coupled Receptors: II. MECHANISM OF AGONIST ACTIVATION *', *Journal of Biological Chemistry*, 273(29), pp. 17979–17982. Available at: https://doi.org/10.1074/jbc.273.29.17979.

Ghosh, E., Kumari, P., Jaiman, D. and Shukla, A.K. (2015) 'Methodological advances: the unsung heroes of the GPCR structural revolution', *Nature Reviews. Molecular Cell Biology*, 16(2), pp. 69–81. Available at: https://doi.org/10.1038/nrm3933.

Gilman, A.G. (1987) 'G Proteins: Transducers of Receptor-Generated Signals', *Annual Review of Biochemistry*, 56(1), pp. 615–649. Available at: https://doi.org/10.1146/annurev.bi.56.070187.003151.

Gingell, J., Simms, J., Barwell, J., Poyner, D.R., Watkins, H.A., Pioszak, A.A., Sexton, P.M. and Hay, D.L. (2016) 'An allosteric role for receptor activity-modifying proteins in defining GPCR pharmacology', *Cell Discovery*, 2(1), pp. 1–14. Available at: https://doi.org/10.1038/celldisc.2016.12.

Goddard, A.D., Dijkman, P.M., Adamson, R.J. and Watts, A. (2013) 'Chapter 18 - Lipid-Dependent GPCR Dimerization', in P.M. Conn (ed.) *Methods in Cell Biology*. Academic Press (Receptor-Receptor Interactions), pp. 341–357. Available at: https://doi.org/10.1016/B978-0-12-408143-7.00018-9.

Goddard, A.D. and Watts, A. (2012) 'Regulation of G protein-coupled receptors by palmitoylation and cholesterol', *BMC Biology*, 10(1), p. 27. Available at: https://doi.org/10.1186/1741-7007-10-27.

Goers, R., Thoma, J., Ritzmann, N., Di Silvestro, A., Alter, C., Gunkel-Grabole, G., Fotiadis, D., Müller, D.J. and Meier, W. (2018) 'Optimized reconstitution of membrane proteins into synthetic membranes', *Communications Chemistry*, 1(1), pp. 1–10. Available at: https://doi.org/10.1038/s42004-018-0037-8.

Gofman, Y., Haliloglu, T. and Ben-Tal, N. (2012) 'The Transmembrane Helix Tilt May Be Determined by the Balance between Precession Entropy and Lipid Perturbation', *Journal of Chemical Theory and Computation*, 8(8), pp. 2896–2904. Available at: https://doi.org/10.1021/ct300128x.

Goricanec, D., Stehle, R., Egloff, P., Grigoriu, S., Plückthun, A., Wagner, G. and Hagn, F. (2016) 'Conformational dynamics of a G-protein α subunit is tightly regulated by nucleotide binding', *Proceedings of the National Academy of Sciences of the United States of America*, 113(26), pp. E3629–E3638. Available at: https://doi.org/10.1073/pnas.1604125113.

Gorter, E. and Grendel, F. (1925) 'ON BIMOLECULAR LAYERS OF LIPOIDS ON THE CHROMOCYTES OF THE BLOOD', *Journal of Experimental Medicine*, 41(4), pp. 439–443. Available at: https://doi.org/10.1084/jem.41.4.439.

Gould, A. and Missailidis, S. (2011) 'Targeting the Hedgehog Pathway: The development of Cyclopamine and the Development of Anti-Cancer Drugs Targeting the Hedgehog Pathway', *Mini Reviews in Medicinal Chemistry*, 11(3), pp. 200–213.

Grandits, A.M. and Wieser, R. (2021) 'Gene expression changes contribute to stemness and therapy resistance of relapsed acute myeloid leukemia: roles of SOCS2, CALCRL, MTSS1, and KDM6A', *Experimental hematology*, 99, pp. 1–11. Available at: https://doi.org/10.1016/j.exphem.2021.05.004.

Grisshammer, R. (2006) 'Understanding recombinant expression of membrane proteins', *Current Opinion in Biotechnology*, 17(4), pp. 337–340. Available at: https://doi.org/10.1016/j.copbio.2006.06.001.

Grisshammer, R., Duckworth, R. and Henderson, R. (1993) 'Expression of a rat neurotensin receptor in Escherichia coli', *Biochemical Journal*, 295(2), pp. 571–576. Available at: https://doi.org/10.1042/bj2950571.

Gu, W., Yan, H., Winters, K.A., Komorowski, R., Vonderfecht, S., Atangan, L., Sivits, G., Hill, D., Yang, J., Bi, V., Shen, Y., Hu, S., Boone, T., Lindberg, R.A. and Véniant, M.M. (2009) 'Long-Term Inhibition of the Glucagon Receptor with a Monoclonal Antibody in Mice Causes Sustained Improvement in Glycemic Control, with Reversible α-Cell Hyperplasia and Hyperglucagonemia', *Journal of Pharmacology and Experimental Therapeutics*, 331(3), pp. 871–881. Available at: https://doi.org/10.1124/jpet.109.157685.

Guex, N. and Peitsch, M.C. (1997) 'SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling', *Electrophoresis*, 18(15), pp. 2714–2723. Available at: https://doi.org/10.1002/elps.1150181505.

Gurevich, V.V. and Gurevich, E.V. (2008) 'GPCR monomers and oligomers: it takes all kinds', *Trends in Neurosciences*, 31(2), pp. 74–81. Available at: https://doi.org/10.1016/j.tins.2007.11.007.

Gurevich, V.V. and Gurevich, E.V. (2014) 'Extensive shape shifting underlies functional versatility of arrestins', *Current Opinion in Cell Biology*, 27, pp. 1–9. Available at: https://doi.org/10.1016/j.ceb.2013.10.007.

Gurevich, V.V. and Gurevich, E.V. (2020) 'Biased GPCR signaling: Possible mechanisms and inherent limitations', *Pharmacology & Therapeutics*, 211, p. 107540. Available at: https://doi.org/10.1016/j.pharmthera.2020.107540.

Gusach, A., García-Nafría, J. and Tate, C.G. (2023) 'New insights into GPCR coupling and dimerisation from cryo-EM structures', *Current Opinion in Structural Biology*, 80, p. 102574. Available at: https://doi.org/10.1016/j.sbi.2023.102574.

Haase, T., Kamann, S. and Brüx, A. (2011) 'FRET as a Tool to Study G-protein Coupled Receptor Oligome- rization in HEK Cells.', p. 4.

Haga, K., Kruse, A.C., Asada, H., Yurugi-Kobayashi, T., Shiroishi, M., Zhang, C., Weis, W.I., Okada, T., Kobilka, B.K., Haga, T. and Kobayashi, T. (2012) 'Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist', *Nature*, 482(7386), pp. 547–551. Available at: https://doi.org/10.1038/nature10753.

Hagn, F., Etzkorn, M., Raschle, T. and Wagner, G. (2013) 'Optimized Phospholipid Bilayer Nanodiscs Facilitate High-Resolution Structure Determination of Membrane Proteins', *Journal of the American Chemical Society*, 135(5), pp. 1919–1925. Available at: https://doi.org/10.1021/ja310901f.

Hammer, O., Harper, D.A.T. and Ryan, P.D. (2001) 'PAST: Paleontological Statistics Software Package for Education and Data Analysis', *Palaeontologia Electronica*, 4(9). Available at: http://palaeo-electronica.org/2001_1/past/issue1_01.htm.

Han, T., Zuo, Z., Qu, M., Zhou, Y., Li, Q. and Wang, H. (2021) 'Comprehensive Analysis of Inflammatory Response–Related Genes, and Prognosis and Immune Infiltration in Patients With Low-Grade Glioma', *Frontiers in Pharmacology*, 12, p. 748993. Available at: https://doi.org/10.3389/fphar.2021.748993.

Hanson, M.A., Cherezov, V., Griffith, M.T., Roth, C.B., Jaakola, V.-P., Chien, E.Y.T., Velasquez, J., Kuhn, P. and Stevens, R.C. (2008) 'A Specific Cholesterol Binding Site Is Established by the 2.8 Å Structure of the Human β2-Adrenergic Receptor', *Structure*, 16(6), pp. 897–905. Available at: https://doi.org/10.1016/j.str.2008.05.001.

Hanson, M.A., Roth, C.B., Jo, E., Griffith, M.T., Scott, F.L., Reinhart, G., Desale, H., Clemons, B., Cahalan, S.M., Schuerer, S.C., Sanna, M.G., Han, G.W., Kuhn, P., Rosen, H. and Stevens, R.C. (2012) 'Crystal structure of a lipid G protein-coupled receptor', *Science (New York, N.Y.)*, 335(6070), pp. 851–855. Available at: https://doi.org/10.1126/science.1215904.

Harikumar, K.G., Simms, J., Christopoulos, G., Sexton, P.M. and Miller, L.J. (2009) 'THE MOLECULAR BASIS OF ASSOCIATION OF RECEPTOR ACTIVITY-MODIFYING PROTEIN 3 WITH THE FAMILY B G PROTEIN-COUPLED SECRETIN RECEPTOR', *Biochemistry*, 48(49), pp. 11773–11785. Available at: https://doi.org/10.1021/bi901326k.

Harmar, A.J. (2001) 'Family-B G-protein-coupled receptors', *Genome Biology*, 2(12), p. reviews3013.1-reviews3013.10.

Harmar, A.J., Fahrenkrug, J., Gozes, I., Laburthe, M., May, V., Pisegna, J.R., Vaudry, D., Vaudry, H., Waschek, J.A. and Said, S.I. (2012) 'Pharmacology and functions of receptors for vasoactive intestinal peptide and pituitary adenylate cyclase-activating polypeptide: IUPHAR Review 1', *British Journal of Pharmacology*, 166(1), pp. 4–17. Available at: https://doi.org/10.1111/j.1476-5381.2012.01871.x.

Has, C., Sivadas, P. and Das, S.L. (2022) 'Insights into Membrane Curvature Sensing and Membrane Remodeling by Intrinsically Disordered Proteins and Protein Regions', *The Journal of Membrane Biology*, 255(2–3), pp. 237–259. Available at: https://doi.org/10.1007/s00232-022-00237-x.

Haskó, G., Kuhel, D.G., Chen, J.-F., Schwarzschild, M.A., Deitch, E.A., Mabley, J.G., Marton, A. and Szabó, C. (2000) 'Adenosine inhibits IL-12 and TNF-α production via adenosine A2a receptor-dependent and independent mechanisms', *The FASEB Journal*, 14(13), pp. 2065–2074. Available at: https://doi.org/10.1096/fj.99-0508com.

Hattab, G., Warschawski, D.E., Moncoq, K. and Miroux, B. (2015) 'Escherichia coli as host for membrane protein structure determination: a global analysis', *Scientific Reports*, 5(1), p. 12097. Available at: https://doi.org/10.1038/srep12097.

Hauer, F., Gerle, C., Fischer, N., Oshima, A., Shinzawa-Itoh, K., Shimada, S., Yokoyama, K., Fujiyoshi, Y. and Stark, H. (2015) 'GraDeR: Membrane Protein Complex Preparation for Single-Particle Cryo-EM', *Structure (London, England: 1993)*, 23(9), pp. 1769–1775. Available at: https://doi.org/10.1016/j.str.2015.06.029.

Hauge-Evans, A.C., Richardson, C.C., Milne, H.M., Christie, M.R., Persaud, S.J. and Jones, P.M. (2006) 'A role for kisspeptin in islet function', *Diabetologia*, 49(9), pp. 2131–2135. Available at: https://doi.org/10.1007/s00125-006-0343-z.

Haupt, S. and Haupt, Y. (2006) 'Importance of p53 for cancer onset and therapy', *Anti-Cancer Drugs*, 17(7), p. 725. Available at: https://doi.org/10.1097/01.cad.0000217422.52208.fa.

Hauser, A.S., Attwood, M.M., Rask-Andersen, M., Schiöth, H.B. and Gloriam, D.E. (2017) 'Trends in GPCR drug discovery: new agents, targets and indications', *Nature reviews. Drug discovery*, 16(12), pp. 829–842. Available at: https://doi.org/10.1038/nrd.2017.178.

Hay, D.L., Chen, S., Lutz, T.A., Parkes, D.G. and Roth, J.D. (2015) 'Amylin: Pharmacology, Physiology, and Clinical Potential', *Pharmacological Reviews*, 67(3), pp. 564–600. Available at: https://doi.org/10.1124/pr.115.010629.

Hay, D.L., Garelja, M.L., Poyner, D.R. and Walker, C.S. (2018) 'Update on the pharmacology of calcitonin/CGRP family of peptides: IUPHAR Review 25', *British Journal of Pharmacology*, 175(1), pp. 3–17. Available at: https://doi.org/10.1111/bph.14075.

Hay, D.L. and Pioszak, A.A. (2016) 'RAMPs (Receptor-Activity Modifying Proteins): New Insights and Roles', *Annual review of pharmacology and toxicology*, 56, pp. 469–487. Available at: https://doi.org/10.1146/annurev-pharmtox-010715-103120.

He, X., You, C., Jiang, H., Jiang, Y., Xu, H.E. and Cheng, X. (2023) 'AlphaFold2 versus experimental structures: evaluation on G protein-coupled receptors', *Acta Pharmacologica Sinica*, 44(1), pp. 1–7. Available at: https://doi.org/10.1038/s41401-022-00938-y.

van Heel, M. and Schatz, M. (2005) 'Fourier shell correlation threshold criteria', *Journal of Structural Biology*, 151(3), pp. 250–262. Available at: https://doi.org/10.1016/j.jsb.2005.05.009.

von Heijne, G. (1992) 'Membrane protein structure prediction: Hydrophobicity analysis and the positive-inside rule', *Journal of Molecular Biology*, 225(2), pp. 487–494. Available at: https://doi.org/10.1016/0022-2836(92)90934-C.

von Heijne, G. (2007) 'The membrane protein universe: what's out there and why bother?', *Journal of Internal Medicine*, 261(6), pp. 543–557. Available at: https://doi.org/10.1111/j.1365-2796.2007.01792.x.

Heitzler, D., Durand, G., Gallay, N., Rizk, A., Ahn, S., Kim, J., Violin, J.D., Dupuy, L., Gauthier, C., Piketty, V., Crépieux, P., Poupon, A., Clément, F., Fages, F., Lefkowitz, R.J. and Reiter, E. (2012) 'Competing G protein-coupled receptor kinases balance G protein and β-arrestin signaling', *Molecular Systems Biology*, 8, p. 590. Available at: https://doi.org/10.1038/msb.2012.22.

Hellmann, J., Drabek, M., Yin, J., Gunera, J., Pröll, T., Kraus, F., Langmead, C.J., Hübner, H., Weikert, D., Kolb, P., Rosenbaum, D.M. and Gmeiner, P. (2020) 'Structure-based development of a subtype-selective orexin 1 receptor antagonist', *Proceedings of the National Academy of Sciences*, 117(30), pp. 18059–18067. Available at: https://doi.org/10.1073/pnas.2002704117.

Henderson, R. and Unwin, P.N.T. (1975) 'Three-dimensional model of purple membrane obtained by electron microscopy', *Nature*, 257(5521), pp. 28–32. Available at: https://doi.org/10.1038/257028a0.

Heng, B.C., Aubel, D. and Fussenegger, M. (2013) 'An overview of the diverse roles of G-protein coupled receptors (GPCRs) in the pathophysiology of various human diseases', *Biotechnology Advances*, 31(8), pp. 1676–1694. Available at: https://doi.org/10.1016/j.biotechadv.2013.08.017.

Hernán, M.A., Takkouche, B., Caamaño-Isorna, F. and Gestal-Otero, J.J. (2002) 'A meta-analysis of coffee drinking, cigarette smoking, and the risk of Parkinson's disease', *Annals of Neurology*, 52(3), pp. 276–284. Available at: https://doi.org/10.1002/ana.10277.

Héroux, M., Breton, B., Hogue, M. and Bouvier, M. (2007) 'Assembly and signaling of CRLR and RAMP1 complexes assessed by BRET', *Biochemistry*, 46(23), pp. 7022–7033. Available at: https://doi.org/10.1021/bi0622470.

Herzik, M.A., Fraser, J.S. and Lander, G.C. (2019) 'A Multi-model Approach to Assessing Local and Global Cryo-EM Map Quality', *Structure*, 27(2), pp. 344-358.e3. Available at: https://doi.org/10.1016/j.str.2018.10.003.

Hewitt, D.J., Aurora, S.K., Dodick, D.W., Goadsby, P.J., Ge, Y.J., Bachman, R., Taraborelli, D., Fan, X., Assaid, C., Lines, C. and Ho, T.W. (2011) 'Randomized controlled trial of the CGRP receptor antagonist MK-3207 in the acute treatment of migraine', *Cephalalgia: An International Journal of Headache*, 31(6), pp. 712–722. Available at: https://doi.org/10.1177/0333102411398399.

Hilairet, S., Bélanger, C., Bertrand, J., Laperrière, A., Foord, S.M. and Bouvier, M. (2001) 'Agonist-promoted internalization of a ternary complex between calcitonin receptor-like receptor, receptor activity-modifying protein 1 (RAMP1), and beta-arrestin', *The Journal of Biological Chemistry*, 276(45), pp. 42182–42190. Available at: https://doi.org/10.1074/jbc.M107323200.

Hilger, D., Kumar, K.K., Hu, H., Pedersen, M.F., O'Brien, E.S., Giehm, L., Jennings, C., Eskici, G., Inoue, A., Lerch, M., Mathiesen, J.M., Skiniotis, G. and Kobilka, B.K. (2020) 'Structural insights into differences in G protein activation by family A and family B GPCRs', *Science (New York, N.Y.)*, 369(6503), p. eaba3373. Available at: https://doi.org/10.1126/science.aba3373.

Hilger, D., Masureel, M. and Kobilka, B.K. (2018) 'Structure and dynamics of GPCR signaling complexes', *Nature structural & molecular biology*, 25(1), p. 4. Available at: https://doi.org/10.1038/s41594-017-0011-7.

Hiller, S., Garces, R.G., Malia, T.J., Orekhov, V.Y., Colombini, M. and Wagner, G. (2008) 'Solution structure of the integral human membrane protein VDAC-1 in detergent micelles', *Science (New York, N.Y.)*, 321(5893), pp. 1206–1210. Available at: https://doi.org/10.1126/science.1161302.

Hinson, J.P., Kapas, S. and Smith, D.M. (2000) 'Adrenomedullin, a multifunctional regulatory peptide', *Endocrine Reviews*, 21(2), pp. 138–167. Available at: https://doi.org/10.1210/edrv.21.2.0396.

Hirz, M., Richter, G., Leitner, E., Wriessnegger, T. and Pichler, H. (2013) 'A novel cholesterol-producing Pichia pastoris strain is an ideal host for functional expression of human Na,K-ATPase α3β1 isoform', *Applied Microbiology and Biotechnology*, 97(21), pp. 9465–9478. Available at: https://doi.org/10.1007/s00253-013-5156-7.

Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R. and Stüber, D. (1988) 'Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent', *Bio/Technology*, 6(11), pp. 1321–1325. Available at: https://doi.org/10.1038/nbt1188-1321.

Hollenberg, C.P. and Gellissen, G. (1997) 'Production of recombinant proteins by methylotrophic yeasts', *Current Opinion in Biotechnology*, 8(5), pp. 554–560. Available at: https://doi.org/10.1016/S0958-1669(97)80028-6.

Holst, J.J., Vilsbøll, T. and Deacon, C.F. (2009) 'The incretin system and its role in type 2 diabetes mellitus', *Molecular and Cellular Endocrinology*, 297(1), pp. 127–136. Available at: https://doi.org/10.1016/j.mce.2008.08.012.

Homan, K.T. and Tesmer, J.J.G. (2015) 'Molecular basis for small molecule inhibition of G protein-coupled receptor kinases', *ACS chemical biology*, 10(1), pp. 246–256. Available at: https://doi.org/10.1021/cb5003976.

Hondo, M., Ishii, M. and Sakurai, T. (2008) 'The NPB/NPW Neuropeptide System and Its Role in Regulating Energy Homeostasis,Pain, and Emotion', in O. Civelli and Q.-Y. Zhou (eds) *Orphan G Protein-Coupled Receptors and Novel Neuropeptides*. Berlin, Heidelberg: Springer (Results and Problems in Cell Differentiation), pp. 239–256. Available at: https://doi.org/10.1007/400_2007_056.

Hoppler, S. and Kavanagh, C.L. (2007) 'Wnt signalling: variety at the core', *Journal of Cell Science*, 120(3), pp. 385–393. Available at: https://doi.org/10.1242/jcs.03363.

Houthuijzen, J.M., Daenen, L.G.M., Roodhart, J.M.L., Oosterom, I., van Jaarsveld, M.T.M., Govaert, K.M., Smith, M.E., Sadatmand, S.J., Rosing, H., Kruse, F., Helms, B.J., van Rooijen, N., Beijnen, J.H., Haribabu, B., van de Lest, C.H.A. and Voest, E.E. (2014) 'Lysophospholipids secreted by splenic macrophages induce chemotherapy resistance via interference with the DNA damage response', *Nature Communications*, 5(1), p. 5275. Available at: https://doi.org/10.1038/ncomms6275.

Hu, Y. and Jin, C. (2022) 'Conformational dynamics in GPCR signaling by NMR', *Magnetic Resonance Letters*, 2(3), pp. 139–146. Available at: https://doi.org/10.1016/j.mrl.2022.06.006.

Huang, C.-C., Orban, T., Jastrzebska, B., Palczewski, K. and Tesmer, J.J.G. (2011) 'Activation of G protein-coupled receptor kinase 1 involves interactions between its N-terminal region and its kinase domain', *Biochemistry*, 50(11), pp. 1940–1949. Available at: https://doi.org/10.1021/bi101606e.

Huang, C.-Y., Olieric, V., Howe, N., Warshamanage, R., Weinert, T., Panepucci, E., Vogeley, L., Basu, S., Diederichs, K., Caffrey, M. and Wang, M. (2018) 'In situ serial crystallography for rapid de novo membrane protein structure determination', *Communications Biology*, 1, p. 124. Available at: https://doi.org/10.1038/s42003-018-0123-6.

Huang, S., Apasov, S., Koshiba, M. and Sitkovsky, M. (1997) 'Role of A2a Extracellular Adenosine Receptor-Mediated Signaling in Adenosine-Mediated Inhibition of T-Cell Activation and Expansion', *Blood*, 90(4), pp. 1600–1610. Available at: https://doi.org/10.1182/blood.V90.4.1600.

Huang, W., Manglik, A., Venkatakrishnan, A.J., Laeremans, T., Feinberg, E.N., Sanborn, A.L., Kato, H.E., Livingston, K.E., Thorsen, T.S., Kling, R.C., Granier, S., Gmeiner, P., Husbands, S.M., Traynor, J.R., Weis, W.I., Steyaert, J., Dror, R.O. and Kobilka, B.K. (2015) 'Structural insights into μ-opioid receptor activation', *Nature*, 524(7565), pp. 315–321. Available at: https://doi.org/10.1038/nature14886.

Husmann, K., Born, W., Fischer, J.A. and Muff, R. (2003) 'Three receptor-activity-modifying proteins define calcitonin gene-related peptide or adrenomedullin selectivity of the mouse calcitonin-like receptor in COS-7 cells', *Biochemical Pharmacology*, 66(11), pp. 2107–2115. Available at: https://doi.org/10.1016/j.bcp.2003.07.009.

Hussain, S.A. (2009) 'An Introduction to Fluorescence Resonance Energy Transfer (FRET)'. arXiv. Available at: https://doi.org/10.48550/arXiv.0908.1815.

Iost, I., Guillerez, J. and Dreyfus, M. (1992) 'Bacteriophage T7 RNA polymerase travels far ahead of ribosomes in vivo', *Journal of Bacteriology*, 174(2), pp. 619–622. Available at: https://doi.org/10.1128/jb.174.2.619-622.1992.

Ippolito, M. and Benovic, J.L. (2021) 'Biased agonism at β-adrenergic receptors', *Cellular signalling*, 80, p. 109905. Available at: https://doi.org/10.1016/j.cellsig.2020.109905.

Isberg, V., de Graaf, C., Bortolato, A., Cherezov, V., Katritch, V., Marshall, F.H., Mordalski, S., Pin, J.-P., Stevens, R.C., Vriend, G. and Gloriam, D.E. (2015) 'Generic GPCR Residue Numbers - Aligning Topology Maps Minding The Gaps', *Trends in pharmacological sciences*, 36(1), pp. 22–31. Available at: https://doi.org/10.1016/j.tips.2014.11.001.

Ishibashi, K., Hara, S. and Kondo, S. (2009) 'Aquaporin water channels in mammals', *Clinical and Experimental Nephrology*, 13(2), pp. 107–117. Available at: https://doi.org/10.1007/s10157-008-0118-6.

Ja, W.W., Carvalho, G.B., Madrigal, M., Roberts, R.W. and Benzer, S. (2009) 'The Drosophila G protein-coupled receptor, Methuselah, exhibits a promiscuous response to peptides', *Protein Science : A Publication of the Protein Society*, 18(11), pp. 2203–2208. Available at: https://doi.org/10.1002/pro.221.

Jaakola, V.-P., Griffith, M.T., Hanson, M.A., Cherezov, V., Chien, E.Y.T., Lane, J.R., IJzerman, A.P. and Stevens, R.C. (2008) 'The 2.6 Angstrom Crystal Structure of a Human A2A Adenosine Receptor Bound to an Antagonist', *Science*, 322(5905), pp. 1211–1217. Available at: https://doi.org/10.1126/science.1164772.

van Jaarsveld, M.T.M., Houthuijzen, J.M. and Voest, E.E. (2016) 'Molecular mechanisms of target recognition by lipid GPCRs: relevance for cancer', *Oncogene*, 35(31), pp. 4021–4035. Available at: https://doi.org/10.1038/onc.2015.467.

Jacob, A., Wu, R. and Wang, P. (2012) 'Regulation of RAMP expression in diseases', *Advances in Experimental Medicine and Biology*, 744, pp. 87–103. Available at: https://doi.org/10.1007/978-1-4614-2364-5_8.

Jacobs, P.P., Geysens, S., Vervecken, W., Contreras, R. and Callewaert, N. (2009) 'Engineering complex-type N-glycosylation in Pichia pastoris using GlycoSwitch technology', *Nature Protocols*, 4(1), pp. 58–70. Available at: https://doi.org/10.1038/nprot.2008.213.

Jacobson, K.A., Gao, Z.-G., Matricon, P., Eddy, M.T. and Carlsson, J. (2022) 'Adenosine A2A receptor antagonists: from caffeine to selective non-xanthines', *British Journal of Pharmacology*, 179(14), pp. 3496–3511. Available at: https://doi.org/10.1111/bph.15103.

Jacoby, E., Bouhelal, R., Gerspacher, M. and Seuwen, K. (2006) 'The 7 TM G-protein-coupled receptor target family', *ChemMedChem*, 1(8), pp. 761–782. Available at: https://doi.org/10.1002/cmdc.200600134.

Jailani, A.B.A., Bigos, K.J.A., Avgoustou, P., Egan, J.L., Hathway, R.A., Skerry, T.M. and Richards, G.O. (2022) 'Targeting the adrenomedullin-2 receptor for the discovery and development of novel anti-cancer agents', *Expert Opinion on Drug Discovery*, 17(8), pp. 839–848. Available at: https://doi.org/10.1080/17460441.2022.2090541.

Jamshad, M., Charlton, J., Lin, Y.-P., Routledge, S.J., Bawa, Z., Knowles, T.J., Overduin, M., Dekker, N., Dafforn, T.R., Bill, R.M., Poyner, D.R. and Wheatley, M. (2015) 'G-protein coupled receptor solubilization and purification for biophysical analysis and functional studies, in the total absence of detergent', *Bioscience Reports*, 35(2), p. e00188. Available at: https://doi.org/10.1042/BSR20140171.

Jiang, H., Galtes, D., Wang, J. and Rockman, H.A. (2022) 'G protein-coupled receptor signaling: transducers and effectors', *American Journal of Physiology-Cell Physiology*, 323(3), pp. C731–C748. Available at: https://doi.org/10.1152/ajpcell.00210.2022.

Jianlin Cheng, Tegge, A.N. and Baldi, P. (2008) 'Machine Learning Methods for Protein Structure Prediction', *IEEE Reviews in Biomedical Engineering*, 1, pp. 41–49. Available at: https://doi.org/10.1109/RBME.2008.2008239.

Jolliffe, I. (2002) *Principal Component Analysis*. 2nd edn. New York, NY: Springer (Springer Series in Statistics). Available at: https://doi.org/10.1007/b98835 (Accessed: 8 October 2023).

Josephs, T.M., Belousoff, M.J., Liang, Y.-L., Piper, S.J., Cao, J., Garama, D.J., Leach, K., Gregory, K.J., Christopoulos, A., Hay, D.L., Danev, R., Wootten, D. and Sexton, P.M. (2021) 'Structure and dynamics of the CGRP receptor in apo and peptide-bound forms', *Science (New York, N.Y.)*, 372(6538), p. eabf7258. Available at: https://doi.org/10.1126/science.abf7258.

Joubert, O., Nehmé, R., Bidet, M. and Mus-Veteau, I. (2010) 'Heterologous Expression of Human Membrane Receptors in the Yeast Saccharomyces cerevisiae', in I. Mus-Veteau (ed.) *Heterologous Expression of Membrane Proteins: Methods and Protocols*. Totowa, NJ: Humana Press (Methods in Molecular Biology™), pp. 87–103. Available at: https://doi.org/10.1007/978-1-60761-344-2_6.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P. and Hassabis, D. (2021) 'Highly accurate protein structure prediction with AlphaFold', *Nature*, 596(7873), pp. 583–589. Available at: https://doi.org/10.1038/s41586-021-03819-2.

Kaebisch, C., Schipper, D., Babczyk, P. and Tobiasch, E. (2015) 'The role of purinergic receptors in stem cell differentiation', *Computational and Structural Biotechnology Journal*, 13, pp. 75–84. Available at: https://doi.org/10.1016/j.csbj.2014.11.003.

Kamato, D., Thach, L., Bernard, R., Chan, V., Zheng, W., Kaur, H., Brimble, M., Osman, N. and Little, P.J. (2015) 'Structure, Function, Pharmacology, and Therapeutic Potential of the G Protein, Gα/q,11', *Frontiers in Cardiovascular Medicine*, 2. Available at: https://doi.org/10.3389/fcvm.2015.00014.

Katoh, Masuko and Katoh, Masaru (2007) 'WNT Signaling Pathway and Stem Cell Signaling Network', *Clinical Cancer Research*, 13(14), pp. 4042–4045. Available at: https://doi.org/10.1158/1078-0432.CCR-06-2316.

Kawabe, Y. and Schaap, P. (2022) 'Adenylate cyclase A amplification and functional diversification during Polyspondylium pallidum development', *EvoDevo*, 13, p. 18. Available at: https://doi.org/10.1186/s13227-022-00203-7.

Kelly, E., Conibear, A. and Henderson, G. (2023) 'Biased Agonism: Lessons from Studies of Opioid Receptor Agonists', *Annual Review of Pharmacology and Toxicology*, 63(Volume 63, 2023), pp. 491–515. Available at: https://doi.org/10.1146/annurev-pharmtox-052120-091058.

Kelly, L., Pieper, U., Eswar, N., Hays, F.A., Li, M., Roe-Zurz, Z., Kroetz, D.L., Giacomini, K.M., Stroud, R.M. and Sali, A. (2009) 'A survey of integral α-helical membrane proteins', *Journal of Structural and Functional Genomics*, 10(4), pp. 269–280. Available at: https://doi.org/10.1007/s10969-009-9069-8.

Kermani, A.A. (2021) 'A guide to membrane protein X-ray crystallography', *The FEBS Journal*, 288(20), pp. 5788–5804. Available at: https://doi.org/10.1111/febs.15676.

Kesidis, A., Depping, P., Lodé, A., Vaitsopoulou, A., Bill, R.M., Goddard, A.D. and Rothnie, A.J. (2020) 'Expression of eukaryotic membrane proteins in eukaryotic and prokaryotic hosts', *Methods*, 180, pp. 3–18. Available at: https://doi.org/10.1016/j.ymeth.2020.06.006.

Khan, M.A.B., Hashim, M.J., King, J.K., Govender, R.D., Mustafa, H. and Al Kaabi, J. (2020) 'Epidemiology of Type 2 Diabetes – Global Burden of Disease and Forecasted Trends', *Journal of Epidemiology and Global Health*, 10(1), pp. 107–111. Available at: https://doi.org/10.2991/jegh.k.191028.001.

Khanapure, S.P., Garvey, D.S., Janero, D.R. and Gordon Letts, L. (2007) 'Eicosanoids in Inflammation: Biosynthesis, Pharmacology, and Therapeutic Frontiers', *Current Topics in Medicinal Chemistry*, 7(3), pp. 311–340.

Kim, L.Y., Rice, W.J., Eng, E.T., Kopylov, M., Cheng, A., Raczkowski, A.M., Jordan, K.D., Bobe, D., Potter, C.S. and Carragher, B. (2018) 'Benchmarking cryo-EM Single Particle Analysis Workflow', *Frontiers in Molecular Biosciences*, 5. Available at: https://www.frontiersin.org/articles/10.3389/fmolb.2018.00050 (Accessed: 17 May 2023).

Kisfalvi, K., Rey, O., Young, S.H., Sinnett-Smith, J. and Rozengurt, E. (2007) 'Insulin Potentiates Ca2+ Signaling and Phosphatidylinositol 4,5-Bisphosphate Hydrolysis Induced by Gq Protein-Coupled Receptor Agonists through an mTOR-Dependent Pathway', *Endocrinology*, 148(7), pp. 3246–3257. Available at: https://doi.org/10.1210/en.2006-1711.

Kita, T. and Kitamura, K. (2022) 'Adrenomedullin Therapy in Moderate to Severe COVID-19', *Biomedicines*, 10(3), p. 533. Available at: https://doi.org/10.3390/biomedicines10030533.

Klammt, C., Schwarz, D., Löhr, F., Schneider, B., Dötsch, V. and Bernhard, F. (2006) 'Cell-free expression as an emerging technique for the large scale production of integral membrane protein', *The FEBS Journal*, 273(18), pp. 4141–4153. Available at: https://doi.org/10.1111/j.1742-4658.2006.05432.x.

Klein Herenbrink, C., Sykes, D.A., Donthamsetti, P., Canals, M., Coudrat, T., Shonberg, J., Scammells, P.J., Capuano, B., Sexton, P.M., Charlton, S.J., Javitch, J.A., Christopoulos, A. and Lane, J.R. (2016) 'The role of kinetic context in apparent biased agonism at GPCRs', *Nature Communications*, 7(1), p. 10842. Available at: https://doi.org/10.1038/ncomms10842.

Klein, K.R., Matson, B.C. and Caron, K.M. (2016) 'The expanding repertoire of receptor activity modifying protein (RAMP) function', *Critical Reviews in Biochemistry and Molecular Biology*, 51(1), pp. 65–71. Available at: https://doi.org/10.3109/10409238.2015.1128875.

Klingenberg, C.P. (2011) 'MorphoJ: an integrated software package for geometric morphometrics: COMPUTER PROGRAM NOTE', *Molecular Ecology Resources*, 11(2), pp. 353–357. Available at: https://doi.org/10.1111/j.1755-0998.2010.02924.x.

Knowles, T.J., Finka, R., Smith, C., Lin, Y.-P., Dafforn, T. and Overduin, M. (2009) 'Membrane Proteins Solubilized Intact in Lipid Containing Nanoparticles Bounded by Styrene Maleic Acid Copolymer', *Journal of the American Chemical Society*, 131(22), pp. 7484–7485. Available at: https://doi.org/10.1021/ja810046q.

Kobilka, B.K. (2007) 'G protein coupled receptor structure and activation', *Biochimica Et Biophysica Acta*, 1768(4), pp. 794–807. Available at: https://doi.org/10.1016/j.bbamem.2006.10.021.

Koehl, A., Hu, H., Feng, D., Sun, B., Zhang, Y., Robertson, M.J., Chu, M., Kobilka, T.S., Pardon, E., Steyaert, J., Tarrasch, J., Dutta, S., Fonseca, R., Weis, W.I., Mathiesen, J.M., Skiniotis, G. and Kobilka, B.K. (2019) 'Structural Insights into Metabotropic Glutamate Receptor Activation', *Nature*, 566(7742), pp. 79–84. Available at: https://doi.org/10.1038/s41586-019-0881-4.

Kojima, M. and Kangawa, K. (2005) 'Ghrelin: Structure and Function', *Physiological Reviews*, 85(2), pp. 495–522. Available at: https://doi.org/10.1152/physrev.00012.2004.

Kolakowski, L.F. (1994) 'GCRDb: a G-protein-coupled receptor database', *Receptors & Channels*, 2(1), pp. 1–7.

Kooistra, A.J., Mordalski, S., Pándy-Szekeres, G., Esguerra, M., Mamyrbekov, A., Munk, C., Keserű, G.M. and Gloriam, D.E. (2021) 'GPCRdb in 2021: integrating GPCR sequence, structure and function', *Nucleic Acids Research*, 49(D1), pp. D335–D343. Available at: https://doi.org/10.1093/nar/gkaa1080.

Kotliar, I.B., Lorenzen, E., Schwenk, J.M., Hay, D.L. and Sakmar, T.P. (2023) 'Elucidating the Interactome of G Protein-Coupled Receptors and Receptor Activity-Modifying Proteins', *Pharmacological Reviews*, 75(1), pp. 1–34. Available at: https://doi.org/10.1124/pharmrev.120.000180.

Kozielewicz, P., Turku, A. and Schulte, G. (2020) 'Molecular Pharmacology of Class F Receptor Activation', *Molecular Pharmacology*, 97(2), pp. 62–71. Available at: https://doi.org/10.1124/mol.119.117986.

Kreida, S. and Törnroth-Horsefield, S. (2015) 'Structural insights into aquaporin selectivity and regulation', *Current Opinion in Structural Biology*, 33, pp. 126–134. Available at: https://doi.org/10.1016/j.sbi.2015.08.004.

Krepkiy, D., Wong, K., Gawrisch, K. and Yeliseev, A. (2006) 'Bacterial expression of functional, biotinylated peripheral cannabinoid receptor CB2', *Protein Expression and Purification*, 49(1), pp. 60–70. Available at: https://doi.org/10.1016/j.pep.2006.03.002.

Krissinel, E. and Henrick, K. (2004) 'Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions', *Acta Crystallographica Section D: Biological Crystallography*, 60(12), pp. 2256–2268. Available at: https://doi.org/10.1107/S0907444904026460.

Kruse, A.C., Hu, J., Pan, A.C., Arlow, D.H., Rosenbaum, D.M., Rosemond, E., Green, H.F., Liu, T., Chae, P.S., Dror, R.O., Shaw, D.E., Weis, W.I., Wess, J. and Kobilka, B.K. (2012) 'Structure and dynamics of the M3 muscarinic acetylcholine receptor', *Nature*, 482(7386), pp. 552–556. Available at: https://doi.org/10.1038/nature10867.

Kruse, A.C., Kobilka, B.K., Gautam, D., Sexton, P.M., Christopoulos, A. and Wess, J. (2014) 'Muscarinic acetylcholine receptors: novel opportunities for drug development', *Nature Reviews. Drug Discovery*, 13(7), pp. 549–560. Available at: https://doi.org/10.1038/nrd4295.

Kruse, A.C., Ring, A.M., Manglik, A., Hu, J., Hu, K., Eitel, K., Hübner, H., Pardon, E., Valant, C., Sexton, P.M., Christopoulos, A., Felder, C.C., Gmeiner, P., Steyaert, J., Weis, W.I., Garcia, K.C., Wess, J. and Kobilka, B.K. (2013) 'Activation and allosteric modulation of a muscarinic acetylcholine receptor', *Nature*, 504(7478), pp. 101–106. Available at: https://doi.org/10.1038/nature12735.

Kruse, E., Uehlein, N. and Kaldenhoff, R. (2006) 'The aquaporins', *Genome Biology*, 7(2), p. 206. Available at: https://doi.org/10.1186/gb-2006-7-2-206.

Kruse, T., Hansen, J.L., Dahl, K., Schäffer, L., Sensfuss, U., Poulsen, C., Schlein, M., Hansen, A.M.K., Jeppesen, C.B., Dornonville de la Cour, C., Clausen, T.R., Johansson, E., Fulle, S., Skyggebjerg, R.B. and Raun, K. (2021) 'Development of Cagrilintide, a Long-Acting Amylin Analogue', *Journal of Medicinal Chemistry*, 64(15), pp. 11183–11194. Available at: https://doi.org/10.1021/acs.jmedchem.1c00565.

Kuang, Q., Purhonen, P. and Hebert, H. (2015) 'Structure of potassium channels', *Cellular and Molecular Life Sciences*, 72, pp. 3677–3693. Available at: https://doi.org/10.1007/s00018-015-1948-5.

Kufareva, I. and Abagyan, R. (2012) 'Methods of protein structure comparison', *Methods in molecular biology (Clifton, N.J.)*, 857, pp. 231–257. Available at: https://doi.org/10.1007/978-1-61779-588-6_10.

Kulbacka, J., Choromańska, A., Rossowska, J., Wężgowiec, J., Saczko, J. and Rols, M.-P. (2017) 'Cell Membrane Transport Mechanisms: Ion Channels and Electrical Properties of Cell Membranes', in J. Kulbacka and S. Satkauskas (eds) *Transport Across Natural and Modified Biological Membranes and its Implications in Physiology and Therapy*. Cham: Springer International Publishing (Advances in Anatomy, Embryology and Cell Biology), pp. 39–58. Available at: https://doi.org/10.1007/978-3-319-56895-9_3.

Kumar, J.P. and Ready, D.F. (1995) 'Rhodopsin plays an essential structural role in Drosophila photoreceptor development', *Development*, 121(12), pp. 4359–4370. Available at: https://doi.org/10.1242/dev.121.12.4359.

Kunishima, N., Shimada, Y., Tsuji, Y., Sato, T., Yamamoto, M., Kumasaka, T., Nakanishi, S., Jingami, H. and Morikawa, K. (2000) 'Structural basis of glutamate recognition by a dimeric metabotropic glutamate receptor', *Nature*, 407(6807), pp. 971–977. Available at: https://doi.org/10.1038/35039564.

Kuwasako, K., Cao, Y.-N., Chu, C.-P., Iwatsubo, S., Eto, T. and Kitamura, K. (2006) 'Functions of the cytoplasmic tails of the human receptor activity-modifying protein components of calcitonin gene-related peptide and adrenomedullin receptors', *The Journal of Biological Chemistry*, 281(11), pp. 7205–7213. Available at: https://doi.org/10.1074/jbc.M511147200.

Kuwasako, K., Shimekake, Y., Masuda, M., Nakahara, K., Yoshida, T., Kitaura, M., Kitamura, K., Eto, T. and Sakata, T. (2000) 'Visualization of the calcitonin receptor-like receptor and its receptor activity-modifying proteins during internalization and recycling', *The Journal of Biological Chemistry*, 275(38), pp. 29602–29609. Available at: https://doi.org/10.1074/jbc.M004534200.

Kwan, T.O.C., Reis, R., Siligardi, G., Hussain, R., Cheruvara, H. and Moraes, I. (2019) 'Selection of Biophysical Methods for Characterisation of Membrane Proteins', *International Journal of Molecular Sciences*, 20(10), p. 2605. Available at: https://doi.org/10.3390/ijms20102605.

La Fountaine, M.F., Hohn, A.N., Leahy, C.L., Weir, J.P. and Testa, A.J. (2022) 'Observations from a prospective small cohort study suggest that CGRP genes contribute to acute posttraumatic headache burden after concussion', *Frontiers in Neurology*, 13, p. 947524. Available at: https://doi.org/10.3389/fneur.2022.947524.

Laferrère, B., Hart, A.B. and Bowers, C.Y. (2006) 'Obese Subjects Respond to the Stimulatory Effect of the Ghrelin Agonist Growth Hormone-Releasing Peptide-2 on Food Intake', *Obesity*, 14(6), pp. 1056–1063. Available at: https://doi.org/10.1038/oby.2006.121.

Lagane, B., Gaibelet, G., Meilhoc, E., Masson, J.-M., Cézanne, L. and Lopez, A. (2000) 'Role of Sterols in Modulating the Human μ-Opioid Receptor Function in Saccharomyces cerevisiae *', *Journal of Biological Chemistry*, 275(43), pp. 33197–33200. Available at: https://doi.org/10.1074/jbc.C000576200.

Lagerström, M.C. and Schiöth, H.B. (2008) 'Structural diversity of G protein-coupled receptors and significance for drug discovery', *Nature Reviews Drug Discovery*, 7(4), pp. 339–357. Available at: https://doi.org/10.1038/nrd2518.

Lai, C.-H., Chen, A.T., Burns, A.B., Sriram, K., Luo, Y., Tang, X., Branciamore, S., O'Meally, D., Chang, S.-L., Huang, P.-H., Shyy, J.Y.-J., Chien, S., Rockne, R.C. and Chen, Z.B. (2021) 'RAMP2-AS1 Regulates Endothelial Homeostasis and Aging', *Frontiers in Cell and Developmental Biology*, 9, p. 635307. Available at: https://doi.org/10.3389/fcell.2021.635307.

Lam, D.W. and LeRoith, D. (2012) 'The worldwide diabetes epidemic', *Current Opinion in Endocrinology, Diabetes and Obesity*, 19(2), p. 93. Available at: https://doi.org/10.1097/MED.0b013e328350583a.

Lambright, D.G., Sondek, J., Bohm, A., Skiba, N.P., Hamm, H.E. and Sigler, P.B. (1996) 'The 2.0 Å crystal structure of a heterotrimeric G protein', *Nature*, 379(6563), pp. 311–319. Available at: https://doi.org/10.1038/379311a0.

Landau, E.M. and Rosenbusch, J.P. (1996) 'Lipidic cubic phases: A novel concept for the crystallization of membrane proteins', *Proceedings of the National Academy of Sciences*, 93(25), pp. 14532–14535. Available at: https://doi.org/10.1073/pnas.93.25.14532.

Langley, J.N. (1905) ‘On the reaction of cells and of nerve-endings to certain poisons, chiefly as regards the reaction of striated muscle to nicotine and to curari’, *The Journal of Physiology*, 33(4–5), pp. 374–413. Available at: https://doi.org/10.1113/jphysiol.1905.sp001128.

Larsen, A.T., Sonne, N., Andreassen, K.V., Karsdal, M.A. and Henriksen, K. (2020) ‘Dose Frequency Optimization of the Dual Amylin and Calcitonin Receptor Agonist KBP-088: Long-Lasting Improvement in Food Preference and Body Weight Loss’, *The Journal of Pharmacology and Experimental Therapeutics*, 373(2), pp. 269–278. Available at: https://doi.org/10.1124/jpet.119.263400.

Lau, D.C.W., Erichsen, L., Francisco, A.M., Satylganova, A., le Roux, C.W., McGowan, B., Pedersen, S.D., Pietiläinen, K.H., Rubino, D. and Batterham, R.L. (2021) ‘Once-weekly cagrilintide for weight management in people with overweight and obesity: a multicentre, randomised, double-blind, placebo-controlled and active-controlled, dose-finding phase 2 trial’, *Lancet (London, England)*, 398(10317), pp. 2160–2172. Available at: https://doi.org/10.1016/S0140-6736(21)01751-7.

Leach, K., Gregory, K.J., Kufareva, I., Khajehali, E., Cook, A.E., Abagyan, R., Conigrave, A.D., Sexton, P.M. and Christopoulos, A. (2016) ‘Towards a structural understanding of allosteric drugs at the human calcium-sensing receptor’, *Cell Research*, 26(5), pp. 574–592. Available at: https://doi.org/10.1038/cr.2016.36.

Leavesley, S.J. and Rich, T.C. (2016) ‘Overcoming Limitations of FRET Measurements’, *Cytometry. Part A : the journal of the International Society for Analytical Cytology*, 89(4), pp. 325–327. Available at: https://doi.org/10.1002/cyto.a.22851.

Lebon, G., Warne, T., Edwards, P.C., Bennett, K., Langmead, C.J., Leslie, A.G.W. and Tate, C.G. (2011) ‘Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation’, *Nature*, 474(7352), pp. 521–525. Available at: https://doi.org/10.1038/nature10136.

Lee, C., Su, B.-H. and Tseng, Y.J. (2022) ‘Comparative studies of AlphaFold, RoseTTAFold and Modeller: a case study involving the use of G-protein-coupled receptors’, *Briefings in Bioinformatics*, 23(5), p. bbac308. Available at: https://doi.org/10.1093/bib/bbac308.

van der Lee, M.M.C., Verkaar, F., Wat, J.W.Y., van Offenbeek, J., Timmerman, M., Voorneveld, L., van Lith, L.H.C.J. and Zaman, G.J.R. (2013) ‘β-Arrestin-biased signaling of PTH analogs of the type 1 parathyroid hormone receptor’, *Cellular Signalling*, 25(2), pp. 527–538. Available at: https://doi.org/10.1016/j.cellsig.2012.11.012.

Lee, S.C., Knowles, T.J., Postis, V.L.G., Jamshad, M., Parslow, R.A., Lin, Y., Goldman, A., Sridhar, P., Overduin, M., Muench, S.P. and Dafforn, T.R. (2016) ‘A method for detergent-free isolation of membrane proteins in their local lipid environment’, *Nature Protocols*, 11(7), pp. 1149–1162. Available at: https://doi.org/10.1038/nprot.2016.070.

Lee, S.-M., Booe, J.M. and Pioszak, A.A. (2015) ‘Structural insights into ligand recognition and selectivity for class A, B, and C GPCRs’, *European journal of pharmacology*, 763(0 0), pp. 196–205. Available at: https://doi.org/10.1016/j.ejphar.2015.05.013.

Lee, Y., Basith, S. and Choi, S. (2018) ‘Recent Advances in Structure-Based Drug Design Targeting Class A G Protein-Coupled Receptors Utilizing Crystal Structures and Computational Simulations’, *Journal of Medicinal Chemistry*, 61(1), pp. 1–46. Available at: https://doi.org/10.1021/acs.jmedchem.6b01453.

Lefkowitz, R.J. and Shenoy, S.K. (2005) 'Transduction of receptor signals by beta-arrestins', *Science (New York, N.Y.)*, 308(5721), pp. 512–517. Available at: https://doi.org/10.1126/science.1109237.

Lei, S., Clydesdale, L., Dai, A., Cai, X., Feng, Y., Yang, D., Liang, Y.-L., Koole, C., Zhao, P., Coudrat, T., Christopoulos, A., Wang, M.-W., Wootten, D. and Sexton, P.M. (2018) 'Two distinct domains of the glucagon-like peptide-1 receptor control peptide-mediated biased agonism', *The Journal of Biological Chemistry*, 293(24), pp. 9370–9387. Available at: https://doi.org/10.1074/jbc.RA118.003278.

Leman, J.K., Ulmschneider, M.B. and Gray, J.J. (2015) 'Computational modeling of membrane proteins', *Proteins*, 83(1), pp. 1–24. Available at: https://doi.org/10.1002/prot.24703.

Lenhart, P.M., Broselid, S., Barrick, C.J., Leeb-Lundberg, L.M.F. and Caron, K.M. (2013) 'G-protein Coupled Receptor 30 Interacts with Receptor Activity Modifying Protein 3 and Confers Sex-Dependent Cardioprotection', *Journal of molecular endocrinology*, 51(1), pp. 191–202. Available at: https://doi.org/10.1530/JME-13-0021.

Leurs, R., Smit, M.J., Alewijnse, A.E. and Timmerman, H. (1998) 'Agonist-independent regulation of constitutively active G-protein-coupled receptors', *Trends in Biochemical Sciences*, 23(11), pp. 418–422. Available at: https://doi.org/10.1016/S0968-0004(98)01287-0.

L. Hall, S.C., Tognoloni, C., Charlton, J., C. Bragginton, É., J. Rothnie, A., Sridhar, P., Wheatley, M., J. Knowles, T., Arnold, T., J. Edler, K. and R. Dafforn, T. (2018) 'An acid-compatible co-polymer for the solubilization of membranes and proteins into lipid bilayer-containing nanoparticles', *Nanoscale*, 10(22), pp. 10609–10619. Available at: https://doi.org/10.1039/C8NR01322E.

Li, M., Hays, F.A., Roe-Zurz, Z., Vuong, L., Kelly, L., Ho, C.-M., Robbins, R.M., Pieper, U., O'Connell, J.D., Miercke, L.J.W., Giacomini, K.M., Sali, A. and Stroud, R.M. (2009) 'Selecting Optimum Eukaryotic Integral Membrane Proteins for Structure Determination by Rapid Expression and Solubilization Screening', *Journal of Molecular Biology*, 385(3), pp. 820–830. Available at: https://doi.org/10.1016/j.jmb.2008.11.021.

Li, X., Dang, S., Yan, C., Gong, X., Wang, J. and Shi, Y. (2013) 'Structure of a presenilin family intramembrane aspartate protease', *Nature*, 493(7430), pp. 56–61. Available at: https://doi.org/10.1038/nature11801.

Liang, Y.-L., Belousoff, M.J., Fletcher, M.M., Zhang, X., Khoshouei, M., Deganutti, G., Koole, C., Furness, S.G.B., Miller, L.J., Hay, D.L., Christopoulos, A., Reynolds, C.A., Danev, R., Wootten, D. and Sexton, P.M. (2020) 'Structure and Dynamics of Adrenomedullin Receptors AM1 and AM2 Reveal Key Mechanisms in the Control of Receptor Phenotype by Receptor Activity-Modifying Proteins', *ACS Pharmacology & Translational Science*, 3(2), pp. 263–284. Available at: https://doi.org/10.1021/acsptsci.9b00080.

Liang, Y.-L., Belousoff, M.J., Zhao, P., Koole, C., Fletcher, M.M., Truong, T.T., Julita, V., Christopoulos, G., Xu, H.E., Zhang, Y., Khoshouei, M., Christopoulos, A., Danev, R., Sexton, P.M. and Wootten, D. (2020) 'Toward a Structural Understanding of Class B GPCR Peptide Binding and Activation', *Molecular Cell*, 77(3), pp. 656-668.e5. Available at: https://doi.org/10.1016/j.molcel.2020.01.012.

Liang, Y.-L., Khoshouei, M., Deganutti, G., Glukhova, A., Koole, C., Peat, T.S., Radjainia, M., Plitzko, J.M., Baumeister, W., Miller, L.J., Hay, D.L., Christopoulos, A., Reynolds, C.A., Wootten, D. and Sexton, P.M. (2018) 'Cryo-EM structure of the active, Gs-protein complexed,

human CGRP receptor', *Nature*, 561(7724), pp. 492–497. Available at: https://doi.org/10.1038/s41586-018-0535-y.

Liang, Y.-L., Khoshouei, M., Glukhova, A., Furness, S.G.B., Zhao, P., Clydesdale, L., Koole, C., Truong, T.T., Thal, D.M., Lei, S., Radjainia, M., Danev, R., Baumeister, W., Wang, M.-W., Miller, L.J., Christopoulos, A., Sexton, P.M. and Wootten, D. (2018) 'Phase-plate cryo-EM structure of a biased agonist-bound human GLP-1 receptor-Gs complex', *Nature*, 555(7694), pp. 121–125. Available at: https://doi.org/10.1038/nature25773.

Lim, J., Petersen, M., Bunz, M., Simon, C. and Schindler, M. (2022) 'Flow cytometry based-FRET: basics, novel developments and future perspectives', *Cellular and Molecular Life Sciences*, 79(4), p. 217. Available at: https://doi.org/10.1007/s00018-022-04232-2.

Lin, H.Y., Harris, T.L., Flannery, M.S., Aruffo, A., Kaji, E.H., Gorn, A., Kolakowski, L.F., Lodish, H.F. and Goldring, S.R. (1991) 'Expression cloning of an adenylate cyclase-coupled calcitonin receptor', *Science (New York, N.Y.)*, 254(5034), pp. 1022–1024. Available at: https://doi.org/10.1126/science.1658940.

Lin, X., Li, M., Wang, N., Wu, Y., Luo, Z., Guo, S., Han, G.-W., Li, S., Yue, Y., Wei, X., Xie, X., Chen, Y., Zhao, S., Wu, J., Lei, M. and Xu, F. (2020) 'Structural basis of ligand recognition and self-activation of orphan GPR52', *Nature*, 579(7797), pp. 152–157. Available at: https://doi.org/10.1038/s41586-020-2019-0.

Ling, S., Shi, P., Liu, S., Meng, X., Zhou, Y., Sun, W., Chang, S., Zhang, X., Zhang, L., Shi, C., Sun, D., Liu, L. and Tian, C. (2021) 'Structural mechanism of cooperative activation of the human calcium-sensing receptor by Ca2+ ions and L-tryptophan', *Cell Research*, 31(4), pp. 383–394. Available at: https://doi.org/10.1038/s41422-021-00474-0.

Lipowsky, R. (2014) 'Remodeling of membrane compartments: some consequences of membrane fluidity', *Biological Chemistry*, 395(3), pp. 253–274. Available at: https://doi.org/10.1515/hsz-2013-0244.

Liu, C., Chen, J., Gao, Y., Deng, B. and Liu, K. (2021) 'Endothelin receptor antagonists for pulmonary arterial hypertension', *The Cochrane Database of Systematic Reviews*, 3(3), p. CD004434. Available at: https://doi.org/10.1002/14651858.CD004434.pub6.

Liu, J., Conklin, B.R., Blin, N., Yun, J. and Wess, J. (1995) 'Identification of a receptor/G-protein contact site critical for signaling specificity and G-protein activation.', *Proceedings of the National Academy of Sciences of the United States of America*, 92(25), pp. 11642–11646.

Liu, S. and Li, W. (2022) 'Protein Fusion Strategies for Membrane Protein Stabilization and Crystal Structure Determination', *Crystals*, 12(8), p. 1041. Available at: https://doi.org/10.3390/cryst12081041.

Liu, X. (2019) 'ABC Family Transporters', in X. Liu and G. Pan (eds) *Drug Transporters in Drug Disposition, Effects and Toxicity*. Singapore: Springer (Advances in Experimental Medicine and Biology), pp. 13–100. Available at: https://doi.org/10.1007/978-981-13-7647-4_2.

Logez, C., Damian, M., Legros, C., Dupré, C., Guéry, M., Mary, S., Wagner, R., M'Kadmi, C., Nosjean, O., Fould, B., Marie, J., Fehrentz, J.-A., Martinez, J., Ferry, G., Boutin, J.A. and Banères, J.-L. (2016) 'Detergent-free Isolation of Functional G Protein-Coupled Receptors into Nanometric Lipid Particles', *Biochemistry*, 55(1), pp. 38–48. Available at: https://doi.org/10.1021/acs.biochem.5b01040.

Lohse, M.J., Andexinger, S., Pitcher, J., Trukawinski, S., Codina, J., Faure, J.P., Caron, M.G. and Lefkowitz, R.J. (1992) 'Receptor-specific desensitization with purified proteins. Kinase dependence and receptor specificity of beta-arrestin and arrestin in the beta 2-adrenergic receptor and rhodopsin systems', *The Journal of Biological Chemistry*, 267(12), pp. 8558–8564.

Luttrell, L.M., Ferguson, S.S., Daaka, Y., Miller, W.E., Maudsley, S., Della Rocca, G.J., Lin, F., Kawakatsu, H., Owada, K., Luttrell, D.K., Caron, M.G. and Lefkowitz, R.J. (1999) 'Beta-arrestin-dependent formation of beta2 adrenergic receptor-Src protein kinase complexes', *Science (New York, N.Y.)*, 283(5402), pp. 655–661. Available at: https://doi.org/10.1126/science.283.5402.655.

Lyons, J.A., Shahsavar, A., Paulsen, P.A., Pedersen, B.P. and Nissen, P. (2016) 'Expression strategies for structural studies of eukaryotic membrane proteins', *Current Opinion in Structural Biology*, 38, pp. 137–144. Available at: https://doi.org/10.1016/j.sbi.2016.06.011.

Lyumkis, D. (2019) 'Challenges and opportunities in cryo-EM single-particle analysis', *Journal of Biological Chemistry*, 294(13), pp. 5181–5197. Available at: https://doi.org/10.1074/jbc.REV118.005602.

Ma, S., Shen, Q., Zhao, L.-H., Mao, C., Zhou, X.E., Shen, D.-D., de Waal, P.W., Bi, P., Li, C., Jiang, Y., Wang, M.-W., Sexton, P.M., Wootten, D., Melcher, K., Zhang, Y. and Xu, H.E. (2020) 'Molecular Basis for Hormone Recognition and Activation of Corticotropin-Releasing Factor Receptors', *Molecular Cell*, 77(3), pp. 669-680.e4. Available at: https://doi.org/10.1016/j.molcel.2020.01.013.

Mackie, D.I., Nielsen, N.R., Harris, M., Singh, S., Davis, R.B., Dy, D., Ladds, G. and Caron, K.M. (2019) 'RAMP3 determines rapid recycling of atypical chemokine receptor-3 for guided angiogenesis', *Proceedings of the National Academy of Sciences of the United States of America*, 116(48), pp. 24093–24099. Available at: https://doi.org/10.1073/pnas.1905561116.

Maeda, K., Das, D., Nakata, H. and Mitsuya, H. (2012) 'CCR5 inhibitors: emergence, success, and challenges', *Expert Opinion on Emerging Drugs*, 17(2), pp. 135–145. Available at: https://doi.org/10.1517/14728214.2012.673584.

Maehle, A.-H. (2004) '"Receptive Substances": John Newport Langley (1852–1925) and his Path to a Receptor Theory of Drug Action', *Medical History*, 48(2), pp. 153–174.

Maehle, A.-H. (2009) 'A binding question: the evolution of the receptor concept', *Endeavour*, 33(4), pp. 135–140. Available at: https://doi.org/10.1016/j.endeavour.2009.09.001.

Mafi, A., Kim, S.-K. and Goddard, W.A. (2022) 'The mechanism for ligand activation of the GPCR–G protein complex', *Proceedings of the National Academy of Sciences of the United States of America*, 119(18), p. e2110085119. Available at: https://doi.org/10.1073/pnas.2110085119.

Mahon, M.J., Bonacci, T.M., Divieti, P. and Smrcka, A.V. (2006) 'A Docking Site for G Protein βγ Subunits on the Parathyroid Hormone 1 Receptor Supports Signaling through Multiple Pathways', *Molecular Endocrinology*, 20(1), pp. 136–146. Available at: https://doi.org/10.1210/me.2005-0169.

Mahoney, J.P. and Sunahara, R.K. (2016) 'Mechanistic insights into GPCR–G protein interactions', *Current Opinion in Structural Biology*, 41, pp. 247–254. Available at: https://doi.org/10.1016/j.sbi.2016.11.005.

Maletínská, L., Matyšková, R., Maixnerová, J., Sýkora, D., Pýchová, M., Špolcová, A., Blechová, M., Drápalová, J., Lacinová, Z., Haluzík, M. and Železná, B. (2011) 'The Peptidic GHS-R antagonist [D-Lys3]GHRP-6 markedly improves adiposity and related metabolic abnormalities in a mouse model of postmenopausal obesity', *Molecular and Cellular Endocrinology*, 343(1), pp. 55–62. Available at: https://doi.org/10.1016/j.mce.2011.06.006.

Mallipeddi, S., Zvonok, N. and Makriyannis, A. (2018) 'Expression, Purification and Characterization of the Human Cannabinoid 1 Receptor', *Scientific Reports*, 8(1), p. 2935. Available at: https://doi.org/10.1038/s41598-018-19749-5.

Manahan, C.L., Iglesias, P.A., Long, Y. and Devreotes, P.N. (2004) 'Chemoattractant Signaling in Dictyostelium Discoideum', *Annual Review of Cell and Developmental Biology*, 20(1), pp. 223–253. Available at: https://doi.org/10.1146/annurev.cellbio.20.011303.132633.

Manglik, A., Kruse, A.C., Kobilka, T.S., Thian, F.S., Mathiesen, J.M., Sunahara, R.K., Pardo, L., Weis, W.I., Kobilka, B.K. and Granier, S. (2012) 'Crystal structure of the μ-opioid receptor bound to a morphinan antagonist', *Nature*, 485(7398), pp. 321–326. Available at: https://doi.org/10.1038/nature10954.

Marassi, F.M., Ding, Y., Schwieters, C.D., Tian, Y. and Yao, Y. (2015) 'Backbone structure of Yersinia pestis Ail determined in micelles by NMR-restrained simulated annealing with implicit membrane solvation', *Journal of biomolecular NMR*, 63(1), pp. 59–65. Available at: https://doi.org/10.1007/s10858-015-9963-2.

Marblestone, J.G., Edavettal, S.C., Lim, Y., Lim, P., Zuo, X. and Butt, T.R. (2006) 'Comparison of SUMO fusion technology with traditional gene fusion systems: Enhanced expression and solubility with SUMO', *Protein Science*, 15(1), pp. 182–189. Available at: https://doi.org/10.1110/ps.051812706.

Mariani, V., Biasini, M., Barbato, A. and Schwede, T. (2013) 'lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests', *Bioinformatics*, 29(21), pp. 2722–2728. Available at: https://doi.org/10.1093/bioinformatics/btt473.

Marinko, J.T., Huang, H., Penn, W.D., Capra, J.A., Schlebach, J.P. and Sanders, C.R. (2019) 'Folding and Misfolding of Human Membrane Proteins in Health and Disease: From Single Molecules to Cellular Proteostasis', *Chemical Reviews*, 119(9), pp. 5537–5606. Available at: https://doi.org/10.1021/acs.chemrev.8b00532.

Markby, D.W., Onrust, R. and Bourne, H.R. (1993) 'Separate GTP Binding and GTPase Activating Domains of a Gα Subunit', *Science*, 262(5141), pp. 1895–1901. Available at: https://doi.org/10.1126/science.8266082.

Marley, J., Lu, M. and Bracken, C. (2001) 'A method for efficient isotopic labeling of recombinant proteins', *Journal of Biomolecular NMR*, 20(1), pp. 71–75. Available at: https://doi.org/10.1023/A:1011254402785.

Martin, D., Galisteo, R. and Gutkind, J.S. (2009) 'CXCL8/IL8 Stimulates Vascular Endothelial Growth Factor (VEGF) Expression and the Autocrine Activation of VEGFR2 in Endothelial Cells by Activating NFκB through the CBM (Carma3/Bcl10/Malt1) Complex*', *Journal of Biological Chemistry*, 284(10), pp. 6038–6042. Available at: https://doi.org/10.1074/jbc.C800207200.

Martin, J. and Sawyer, A. (2019) 'Elucidating the structure of membrane proteins', *BioTechniques*, 66(4), pp. 167–170. Available at: https://doi.org/10.2144/btn-2019-0030.

Martinelli, A. and Tuccinardi, T. (2008) 'Molecular modeling of adenosine receptors: new results and trends', *Medicinal Research Reviews*, 28(2), pp. 247–277. Available at: https://doi.org/10.1002/med.20106.

Martin-Fernandez, M., Tynan, C. and Webb, S. (2013) 'A "pocket guide" to total internal reflection fluorescence', *Journal of Microscopy*, 252(1), pp. 16–22. Available at: https://doi.org/10.1111/jmi.12070.

Mason, J.S., Bortolato, A., Congreve, M. and Marshall, F.H. (2012) 'New insights from structural biology into the druggability of G protein-coupled receptors', *Trends in Pharmacological Sciences*, 33(5), pp. 249–260. Available at: https://doi.org/10.1016/j.tips.2012.02.005.

Masuho, I., Kise, R., Gainza, P., Von Moo, E., Li, X., Tany, R., Wakasugi-Masuho, H., Correia, B.E. and Martemyanov, K.A. (2023) 'Rules and mechanisms governing G protein coupling selectivity of GPCRs', *Cell Reports*, 42(10), p. 113173. Available at: https://doi.org/10.1016/j.celrep.2023.113173.

Masuho, I., Ostrovskaya, O., Kramer, G.M., Jones, C.D., Xie, K. and Martemyanov, K.A. (2015) 'Distinct profiles of functional discrimination among G proteins determine the actions of G protein–coupled receptors', *Science Signaling*, 8(405), pp. ra123–ra123. Available at: https://doi.org/10.1126/scisignal.aab4068.

Mathieu, K., Javed, W., Vallet, S., Lesterlin, C., Candusso, M.-P., Ding, F., Xu, X.N., Ebel, C., Jault, J.-M. and Orelle, C. (2019) 'Functionality of membrane proteins overexpressed and purified from E. coli is highly dependent upon the strain', *Scientific Reports*, 9, p. 2654. Available at: https://doi.org/10.1038/s41598-019-39382-0.

Mazzoni, M.R., Malinski, J.A. and Hamm, H.E. (1991) 'Structural analysis of rod GTP-binding protein, Gt. Limited proteolytic digestion pattern of Gt with four proteases defines monoclonal antibody epitope', *Journal of Biological Chemistry*, 266(21), pp. 14072–14081. Available at: https://doi.org/10.1016/S0021-9258(18)92811-5.

McClendon, C.L., Kornev, A.P., Gilson, M.K. and Taylor, S.S. (2014) 'Dynamic architecture of a protein kinase', *Proceedings of the National Academy of Sciences of the United States of America*, 111(43), pp. E4623-4631. Available at: https://doi.org/10.1073/pnas.1418402111.

McCudden, C.R., Hains, M.D., Kimple, R.J., Siderovski, D.P. and Willard, F.S. (2005) 'G-protein signaling: back to the future', *Cellular and Molecular Life Sciences*, 62(5), pp. 551–577. Available at: https://doi.org/10.1007/s00018-004-4462-3.

McGlone, E.R., Manchanda, Y., Jones, B., Pickford, P., Inoue, A., Carling, D., Bloom, S.R., Tan, T. and Tomas, A. (2021) 'Receptor Activity-Modifying Protein 2 (RAMP2) alters glucagon receptor trafficking in hepatocytes with functional effects on receptor signalling', *Molecular Metabolism*, 53, p. 101296. Available at: https://doi.org/10.1016/j.molmet.2021.101296.

McGuffin, L.J. and Roche, D.B. (2010) 'Rapid model quality assessment for protein structure predictions using the comparison of multiple models without structural alignments', *Bioinformatics*, 26(2), pp. 182–188. Available at: https://doi.org/10.1093/bioinformatics/btp629.

McKee, E.E., Bentley, A.T., Smith, R.M. and Ciaccio, C.E. (1999) 'Origin of Guanine Nucleotides in Isolated Heart Mitochondria', *Biochemical and Biophysical Research Communications*, 257(2), pp. 466–472. Available at: https://doi.org/10.1006/bbrc.1999.0489.

McKenzie, E.A. and Abbott, W.M. (2018) 'Expression of recombinant proteins in insect and mammalian cells', *Methods*, 147, pp. 40–49. Available at: https://doi.org/10.1016/j.ymeth.2018.05.013.

McLatchie, L.M., Fraser, N.J., Main, M.J., Wise, A., Brown, J., Thompson, N., Solari, R., Lee, M.G. and Foord, S.M. (1998) 'RAMPs regulate the transport and ligand specificity of the calcitonin-receptor-like receptor', *Nature*, 393(6683), pp. 333–339. Available at: https://doi.org/10.1038/30666.

de Mendoza, A., Jones, J.W. and Friedrich, M. (2016) 'Methuselah/Methuselah-like G protein-coupled receptors constitute an ancient metazoan gene family', *Scientific Reports*, 6, p. 21801. Available at: https://doi.org/10.1038/srep21801.

de Mendoza, A., Sebé-Pedrós, A. and Ruiz-Trillo, I. (2014) 'The Evolution of the GPCR Signaling System in Eukaryotes: Modularity, Conservation, and the Transition to Metazoan Multicellularity', *Genome Biology and Evolution*, 6(3), pp. 606–619. Available at: https://doi.org/10.1093/gbe/evu038.

Mensah, G.A., Roth, G.A. and Fuster, V. (2019) 'The Global Burden of Cardiovascular Diseases and Risk Factors', *Journal of the American College of Cardiology*, 74(20), pp. 2529–2532. Available at: https://doi.org/10.1016/j.jacc.2019.10.009.

Meyer, A., Dierks, K., Hussein, R., Brillet, K., Brognaro, H. and Betzel, C. (2015) 'Systematic analysis of protein–detergent complexes applying dynamic light scattering to optimize solutions for crystallization trials', *Acta Crystallographica Section F: Structural Biology Communications*, 71(1), pp. 75–81. Available at: https://doi.org/10.1107/S2053230X14027149.

Michalke, K., Huyghe, C., Lichière, J., Gravière, M.-E., Siponen, M., Sciara, G., Lepaul, I., Wagner, R., Magg, C., Rudolph, R., Cambillau, C. and Desmyter, A. (2010) 'Mammalian G protein-coupled receptor expression in Escherichia coli: II. Refolding and biophysical characterization of mouse cannabinoid receptor 1 and human parathyroid hormone receptor 1', *Analytical Biochemistry*, 401(1), pp. 74–80. Available at: https://doi.org/10.1016/j.ab.2010.02.017.

Micrographia by Robert Hooke, 1665 | *The British Library* (no date). Available at: https://www.bl.uk/collection-items/micrographia-by-robert-hooke-1665 (Accessed: 3 October 2022).

Milić, D. and Veprintsev, D.B. (2015) 'Large-scale production and protein engineering of G protein-coupled receptors for structural studies', *Frontiers in Pharmacology*, 6. Available at: https://www.frontiersin.org/articles/10.3389/fphar.2015.00066 (Accessed: 3 November 2022).

Miller, D.E., Cook, K.R., Hemenway, E.A., Fang, V., Miller, A.L., Hales, K.G. and Hawley, R.S. (2018) 'The Molecular and Genetic Characterization of Second Chromosome Balancers in Drosophila melanogaster', *G3: Genes|Genomes|Genetics*, 8(4), pp. 1161–1171. Available at: https://doi.org/10.1534/g3.118.200021.

Mills, G.B. and Moolenaar, W.H. (2003) 'The emerging role of lysophosphatidic acid in cancer', *Nature Reviews Cancer*, 3(8), pp. 582–591. Available at: https://doi.org/10.1038/nrc1143.

Miroux, B. and Walker, J.E. (1996) 'Over-production of Proteins inEscherichia coli: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels', *Journal of Molecular Biology*, 260(3), pp. 289–298. Available at: https://doi.org/10.1006/jmbi.1996.0399.

Missale, C., Nash, S.R., Robinson, S.W., Jaber, M. and Caron, M.G. (1998) 'Dopamine receptors: from structure to function', *Physiological Reviews*, 78(1), pp. 189–225. Available at: https://doi.org/10.1152/physrev.1998.78.1.189.

Mitteroecker, P. and Gunz, P. (2009) 'Advances in Geometric Morphometrics', *Evolutionary Biology*, 36(2), pp. 235–247. Available at: https://doi.org/10.1007/s11692-009-9055-x.

Mizuta, H., Takakusaki, A., Suzuki, T., Otake, K., Dohmae, N. and Simizu, S. (2023) 'C-mannosylation regulates stabilization of RAMP1 protein and RAMP1-mediated cell migration', *The FEBS journal*, 290(1), pp. 196–208. Available at: https://doi.org/10.1111/febs.16592.

Mondal, M.S., Yamaguchi, H., Date, Y., Shimbara, T., Toshinai, K., Shimomura, Y., Mori, M. and Nakazato, M. (2003) 'A Role for Neuropeptide W in the Regulation of Feeding Behavior', *Endocrinology*, 144(11), pp. 4729–4733. Available at: https://doi.org/10.1210/en.2003-0536.

Montanya, E. (2012) 'A comparison of currently available GLP-1 receptor agonists for the treatment of type 2 diabetes', *Expert Opinion on Pharmacotherapy*, 13(10), pp. 1451–1467. Available at: https://doi.org/10.1517/14656566.2012.692777.

Moraes, I., Evans, G., Sanchez-Weatherby, J., Newstead, S. and Stewart, P.D.S. (2014) 'Membrane protein structure determination — The next generation', *Biochimica et Biophysica Acta*, 1838(1), pp. 78–87. Available at: https://doi.org/10.1016/j.bbamem.2013.07.010.

Moran, L.A., Horton, R.A., Scrimgeour, K.G. and Perry, M.D. (2014) *Principles of Biochemistry*. 5th edn. Harlow, Essex: Pearson. Available at: https://www.macmillanlearning.com/ed/uk/product/Lehninger-Principles-of-Biochemistry--8th-edition/p/1319381499 (Accessed: 13 October 2024).

Morfis, M., Tilakaratne, N., Furness, S.G.B., Christopoulos, G., Werry, T.D., Christopoulos, A. and Sexton, P.M. (2008) 'Receptor activity-modifying proteins differentially modulate the G protein-coupling efficiency of amylin receptors', *Endocrinology*, 149(11), pp. 5423–5431. Available at: https://doi.org/10.1210/en.2007-1735.

Mori, A., Chen, J.-F., Uchida, S., Durlach, C., King, S.M. and Jenner, P. (2022) 'The Pharmacological Potential of Adenosine A2A Receptor Antagonists for Treating Parkinson's Disease', *Molecules (Basel, Switzerland)*, 27(7), p. 2366. Available at: https://doi.org/10.3390/molecules27072366.

Morrison, E.A. and Henzler-Wildman, K.A. (2012) 'Reconstitution of integral membrane proteins into isotropic bicelles with improved sample stability and expanded lipid composition profile', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1818(3), pp. 814–820. Available at: https://doi.org/10.1016/j.bbamem.2011.12.020.

Mountjoy, K.G. (2010) 'Distribution and Function of Melanocortin Receptors within the Brain', in A. Catania (ed.) *Melanocortins: Multiple Actions and Therapeutic Potential*. New York, NY: Springer (Advances in Experimental Medicine and Biology), pp. 29–48. Available at: https://doi.org/10.1007/978-1-4419-6354-3_3.

Mu, J., Jiang, G., Brady, E., Dallas-Yang, Q., Liu, F., Woods, J., Zycband, E., Wright, M., Li, Z., Lu, K., Zhu, L., Shen, X., SinhaRoy, R., Candelore, M.L., Qureshi, S.A., Shen, D.-M., Zhang, F., Parmee, E.R. and Zhang, B.B. (2011) 'Chronic treatment with a glucagon receptor antagonist lowers glucose and moderately raises circulating glucagon and glucagon-like peptide 1 without severe alpha cell hypertrophy in diet-induced obese mice', *Diabetologia*, 54(9), pp. 2381–2391. Available at: https://doi.org/10.1007/s00125-011-2217-2.

Murakami, M. and Kouyama, T. (2008) 'Crystal structure of squid rhodopsin', *Nature*, 453(7193), pp. 363–367. Available at: https://doi.org/10.1038/nature06925.

Muranaka, N., Miura, M., Taira, H. and Hohsaka, T. (2007) 'Incorporation of Unnatural Non-α-Amino Acids into the N Terminus of Proteins in a Cell-Free Translation System', *ChemBioChem*, 8(14), pp. 1650–1653. Available at: https://doi.org/10.1002/cbic.200700249.

Murph, M.M., Hurst-Kennedy, J., Newton, V., Brindley, D.N. and Radhakrishna, H. (2007) 'Lysophosphatidic Acid Decreases the Nuclear Localization and Cellular Abundance of the p53 Tumor Suppressor in A549 Lung Carcinoma Cells', *Molecular Cancer Research*, 5(11), pp. 1201–1211. Available at: https://doi.org/10.1158/1541-7786.MCR-06-0338.

Muto, T., Tsuchiya, D., Morikawa, K. and Jingami, H. (2007) 'Structures of the extracellular regions of the group II/III metabotropic glutamate receptors', *Proceedings of the National Academy of Sciences of the United States of America*, 104(10), pp. 3759–3764. Available at: https://doi.org/10.1073/pnas.0611577104.

Nagar, H., Kim, S., Lee, I., Choi, S.-J., Piao, S., Jeon, B.H., Shong, M. and Kim, C.-S. (2021) 'CRIF1 deficiency suppresses endothelial cell migration via upregulation of RhoGDI2', *PLoS ONE*, 16(8), p. e0256646. Available at: https://doi.org/10.1371/journal.pone.0256646.

Naydenova, K. and Russo, C.J. (2017) 'Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy', *Nature Communications*, 8(1), p. 629. Available at: https://doi.org/10.1038/s41467-017-00782-3.

Newstead, S., Kim, H., von Heijne, G., Iwata, S. and Drew, D. (2007) 'High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in Saccharomyces cerevisiae', *Proceedings of the National Academy of Sciences*, 104(35), pp. 13936–13941. Available at: https://doi.org/10.1073/pnas.0704546104.

Nguyen, A.H., Thomsen, A.R.B., Cahill, T.J., Huang, R., Huang, L.-Y., Marcink, T., Clarke, O.B., Heissel, S., Masoudi, A., Ben-Hail, D., Samaan, F., Dandey, V.P., Tan, Y.Z., Hong, C., Mahoney, J.P., Triest, S., Little, J., Chen, X., Sunahara, R., Steyaert, J., Molina, H., Yu, Z., des Georges, A. and Lefkowitz, R.J. (2019) 'Structure of an endosomal signaling GPCR-G protein-β-arrestin megacomplex', *Nature Structural & Molecular Biology*, 26(12), pp. 1123–1131. Available at: https://doi.org/10.1038/s41594-019-0330-y.

Nichols, D.E. and Nichols, C.D. (2008) 'Serotonin Receptors', *Chemical Reviews*, 108(5), pp. 1614–1641. Available at: https://doi.org/10.1021/cr078224o.

Nicolson, G.L. and Ferreira de Mattos, G. (2021) 'A Brief Introduction to Some Aspects of the Fluid–Mosaic Model of Cell Membrane Structure and Its Importance in Membrane Lipid Replacement', *Membranes*, 11(12), p. 947. Available at: https://doi.org/10.3390/membranes11120947.

Nicolson, G.L. and Ferreira de Mattos, G. (2022) 'Fifty Years of the Fluid–Mosaic Model of Biomembrane Structure and Organization and Its Importance in Biomedicine with Particular Emphasis on Membrane Lipid Replacement', *Biomedicines*, 10(7), p. 1711. Available at: https://doi.org/10.3390/biomedicines10071711.

Nietlispach, D. and Gautier, A. (2011) 'Solution NMR studies of polytopic α-helical membrane proteins', *Current Opinion in Structural Biology*, 21(4), pp. 497–508. Available at: https://doi.org/10.1016/j.sbi.2011.06.009.

de Nigris, F., Schiano, C., Infante, T. and Napoli, C. (2012) 'CXCR4 Inhibitors: Tumor Vasculature and Therapeutic Challenges', *Recent Patents on Anti-Cancer Drug Discovery*, 7(3), pp. 251–264. Available at: https://doi.org/10.2174/157489212801820039.

Nivedha, A.K., Lee, S. and Vaidehi, N. (2023) 'Biased agonists differentially modulate the receptor conformation ensembles in Angiotensin II type 1 receptor', *Journal of Molecular Graphics and Modelling*, 118, p. 108365. Available at: https://doi.org/10.1016/j.jmgm.2022.108365.

Nussinov, R., Zhang, M., Liu, Y. and Jang, H. (2022) 'AlphaFold, Artificial Intelligence (AI), and Allostery', *The Journal of Physical Chemistry B*, 126(34), pp. 6372–6383. Available at: https://doi.org/10.1021/acs.jpcb.2c04346.

O'higgins, P. (2000) 'The study of morphological variation in the hominid fossil record: biology, landmarks and geometry', *Journal of Anatomy*, 197(1), pp. 103–120. Available at: https://doi.org/10.1046/j.1469-7580.2000.19710103.x.

Olson, T.S. and Ley, K. (2002) 'Chemokines and chemokine receptors in leukocyte trafficking', *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 283(1), pp. R7–R28. Available at: https://doi.org/10.1152/ajpregu.00738.2001.

Oluwole, A.O., Danielczak, B., Meister, A., Babalola, J.O., Vargas, C. and Keller, S. (2017) 'Solubilization of Membrane Proteins into Functional Lipid-Bilayer Nanodiscs Using a Diisobutylene/Maleic Acid Copolymer', *Angewandte Chemie International Edition*, 56(7), pp. 1919–1924. Available at: https://doi.org/10.1002/anie.201610778.

O'Mahony, L., Akdis, M. and Akdis, C.A. (2011) 'Regulation of the immune response and inflammation by histamine and histamine receptors', *Journal of Allergy and Clinical Immunology*, 128(6), pp. 1153–1162. Available at: https://doi.org/10.1016/j.jaci.2011.06.051.

Opella, S.J. and Marassi, F.M. (2017) 'Applications of NMR to membrane proteins', *Archives of biochemistry and biophysics*, 628, pp. 92–101. Available at: https://doi.org/10.1016/j.abb.2017.05.011.

Oshikawa, S., Tanoue, A., Koshimizu, T., Kitagawa, Y. and Tsujimoto, G. (2004) 'Vasopressin Stimulates Insulin Release from Islet Cells through V1b Receptors: a Combined Pharmacological/Knockout Approach', *Molecular Pharmacology*, 65(3), pp. 623–629. Available at: https://doi.org/10.1124/mol.65.3.623.

Overington, J.P., Al-Lazikani, B. and Hopkins, A.L. (2006) 'How many drug targets are there?', *Nature Reviews Drug Discovery*, 5(12), pp. 993–996. Available at: https://doi.org/10.1038/nrd2199.

Padmanabha Das, K.M., Shih, W.M., Wagner, G. and Nasr, M.L. (2020) 'Large Nanodiscs: A Potential Game Changer in Structural Biology of Membrane Protein Complexes and Virus Entry', *Frontiers in Bioengineering and Biotechnology*, 8. Available at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00539 (Accessed: 27 February 2024).

Pak, M.A., Markhieva, K.A., Novikova, M.S., Petrov, D.S., Vorobyev, I.S., Maksimova, E.S., Kondrashov, F.A. and Ivankov, D.N. (2023) 'Using AlphaFold to predict the impact of single mutations on protein stability and function', *PLOS ONE*, 18(3), p. e0282689. Available at: https://doi.org/10.1371/journal.pone.0282689.

Palczewski, K., Buczyłko, J., Lebioda, L., Crabb, J.W. and Polans, A.S. (1993) 'Identification of the N-terminal region in rhodopsin kinase involved in its interaction with rhodopsin', *The Journal of Biological Chemistry*, 268(8), pp. 6004–6013.

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., Yamamoto, M. and Miyano, M. (2000) 'Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor', *Science*, 289(5480), pp. 739–745. Available at: https://doi.org/10.1126/science.289.5480.739.

Pandey, A., Sarker, M., Liu, X.-Q. and Rainey, J.K. (2014) 'Small expression tags enhance bacterial expression of the first three transmembrane segments of the apelin receptor', *Biochemistry and Cell Biology*, 92(4), pp. 269–278. Available at: https://doi.org/10.1139/bcb-2014-0009.

Pandey, A., Shin, K., Patterson, R.E., Liu, X.-Q. and Rainey, J.K. (2016) 'Current strategies for protein production and purification enabling membrane protein structural biology', *Biochemistry and cell biology = Biochimie et biologie cellulaire*, 94(6), pp. 507–527. Available at: https://doi.org/10.1139/bcb-2015-0143.

Panneels, V., Kock, I., Krijnse-Locker, J., Rezgaoui, M. and Sinning, I. (2011) 'Drosophila Photoreceptor Cells Exploited for the Production of Eukaryotic Membrane Proteins: Receptors, Transporters and Channels', *PLoS ONE*, 6(4), p. e18478. Available at: https://doi.org/10.1371/journal.pone.0018478.

Paradis, P., Dali-Youcef, N., Paradis, F.W., Thibault, G. and Nemer, M. (2000) 'Overexpression of angiotensin II type I receptor in cardiomyocytes induces cardiac hypertrophy and remodeling', *Proceedings of the National Academy of Sciences*, 97(2), pp. 931–936. Available at: https://doi.org/10.1073/pnas.97.2.931.

Parameswaran, N. and Spielman, W.S. (2006) 'RAMPs: The past, present and future', *Trends in Biochemical Sciences*, 31(11), pp. 631–638. Available at: https://doi.org/10.1016/j.tibs.2006.09.006.

Park, S.H., Casagrande, F., Das, B.B., Albrecht, L., Chu, M. and Opella, S.J. (2011) 'Local and Global Dynamics of the G Protein-Coupled Receptor CXCR1', *Biochemistry*, 50(12), pp. 2371–2380. Available at: https://doi.org/10.1021/bi101568j.

Park, S.H., Das, B.B., De Angelis, A.A., Scrima, M. and Opella, S.J. (2010) 'Mechanically, Magnetically, and "Rotationally Aligned" Membrane Proteins in Phospholipid Bilayers Give Equivalent Angular Constraints for NMR Structure Determination', *The journal of physical chemistry. B*, 114(44), pp. 13995–14003. Available at: https://doi.org/10.1021/jp106043w.

Parker, J.L. and Newstead, S. (2013) 'Phasing statistics for alpha helical membrane protein structures', *Protein Science: A Publication of the Protein Society*, 22(11), pp. 1664–1668. Available at: https://doi.org/10.1002/pro.2341.

Parsons, M.E. and Ganellin, C.R. (2006) 'Histamine and its receptors', *British Journal of Pharmacology*, 147 Suppl 1(Suppl 1), pp. S127-135. Available at: https://doi.org/10.1038/sj.bjp.0706440.

Pasternak, G.W. (2014) 'Opioids and their receptors: Are we there yet?', *Neuropharmacology*, 76 Pt B(0 0), pp. 198–203. Available at: https://doi.org/10.1016/j.neuropharm.2013.03.039.

Patterson, G.H., Piston, D.W. and Barisas, B.G. (2000) 'Förster Distances between Green Fluorescent Protein Pairs', *Analytical Biochemistry*, 284(2), pp. 438–440. Available at: https://doi.org/10.1006/abio.2000.4708.

Pearce, L.R., Komander, D. and Alessi, D.R. (2010) 'The nuts and bolts of AGC protein kinases', *Nature Reviews Molecular Cell Biology*, 11(1), pp. 9–22. Available at: https://doi.org/10.1038/nrm2822.

Penczek, P.A. (2010a) 'Chapter Three - Resolution Measures in Molecular Electron Microscopy', in G.J. Jensen (ed.) *Methods in Enzymology*. Academic Press (Cryo-EM, Part B: 3-D Reconstruction), pp. 73–100. Available at: https://doi.org/10.1016/S0076-6879(10)82003-8.

Penczek, P.A. (2010b) 'Chapter Two - Image Restoration in Cryo-Electron Microscopy', in G.J. Jensen (ed.) *Methods in Enzymology*. Academic Press (Cryo-EM, Part B: 3-D Reconstruction), pp. 35–72. Available at: https://doi.org/10.1016/S0076-6879(10)82002-6.

Peng, X., Yang, L., Liu, Z., Lou, S., Mei, S., Li, M., Chen, Z. and Zhang, H. (2022) 'Structural basis for recognition of antihistamine drug by human histamine receptor', *Nature Communications*, 13(1), p. 6105. Available at: https://doi.org/10.1038/s41467-022-33880-y.

Penrice, S.C. and Deeming, Dc. (2020) 'Morphometrics of feeding anatomy in stereospondyl amphibians', *Australasian Palaeontological Memoirs*, (51), pp. 131–140. Available at: https://doi.org/10.3316/informit.282127522636993.

Periasamy, A., Wallrabe, H., Chen, Y. and Barroso, M. (2008) 'Chapter 22 Quantitation of Protein–Protein Interactions: Confocal FRET Microscopy', in *Methods in Cell Biology*. Academic Press (Biophysical Tools for Biologists, Volume Two: In Vivo Techniques), pp. 569–598. Available at: https://doi.org/10.1016/S0091-679X(08)00622-5.

Petrovskaya, L.E., Shulga, A.A., Bocharova, O.V., Ermolyuk, Ya.S., Kryukova, E.A., Chupin, V.V., Blommers, M.J.J., Arseniev, A.S. and Kirpichnikov, M.P. (2010) 'Expression of G-protein coupled receptors in Escherichia coli for structural studies', *Biochemistry (Moscow)*, 75(7), pp. 881–891. Available at: https://doi.org/10.1134/S0006297910070102.

Pfleger, J., Gresham, K. and Koch, W.J. (2019) 'G protein-coupled receptor kinases as therapeutic targets in the heart', *Nature Reviews Cardiology*, 16(10), pp. 612–622. Available at: https://doi.org/10.1038/s41569-019-0220-3.

Pham, V., Zhu, Y., Dal Maso, E., Reynolds, C.A., Deganutti, G., Atanasio, S., Hick, C.A., Yang, D., Christopoulos, A., Hay, D.L., Furness, S.G.B., Wang, M.-W., Wootten, D. and Sexton, P.M. (2019) 'Deconvoluting the Molecular Control of Binding and Signaling at the Amylin 3 Receptor: RAMP3 Alters Signal Propagation through Extracellular Loops of the Calcitonin Receptor', *ACS Pharmacology & Translational Science*, 2(3), pp. 183–197. Available at: https://doi.org/10.1021/acsptsci.9b00010.

Phelps, C.B. and Brand, A.H. (1998) 'Ectopic Gene Expression inDrosophilaUsing GAL4 System', *Methods*, 14(4), pp. 367–379. Available at: https://doi.org/10.1006/meth.1998.0592.

Pin, J.-P. and Bettler, B. (2016) 'Organization and functions of mGlu and GABAB receptor complexes', *Nature*, 540(7631), pp. 60–68. Available at: https://doi.org/10.1038/nature20566.

Pioszak, A.A. and Hay, D.L. (2020) 'RAMPs as allosteric modulators of the calcitonin and calcitonin-like class B G protein-coupled receptors', *Advances in pharmacology (San Diego, Calif.)*, 88, pp. 115–141. Available at: https://doi.org/10.1016/bs.apha.2020.01.001.

Pitcher, J.A., Freedman, N.J. and Lefkowitz, R.J. (1998) 'G Protein–Coupled Receptor Kinases', *Annual Review of Biochemistry*, 67(1), pp. 653–692. Available at: https://doi.org/10.1146/annurev.biochem.67.1.653.

Pitcher, J.A., Inglese, J., Higgins, J.B., Arriza, J.L., Casey, P.J., Kim, C., Benovic, J.L., Kwatra, M.M., Caron, M.G. and Lefkowitz, R.J. (1992) 'Role of βγ Subunits of G Proteins in Targeting the β-Adrenergic Receptor Kinase to Membrane-Bound Receptors', *Science*, 257(5074), pp. 1264–1267. Available at: https://doi.org/10.1126/science.1325672.

Pitson, S.M., Xia, P., Leclercq, T.M., Moretti, P.A.B., Zebol, J.R., Lynn, H.E., Wattenberg, B.W. and Vadas, M.A. (2004) 'Phosphorylation-dependent translocation of sphingosine kinase to the plasma membrane drives its oncogenic signalling', *Journal of Experimental Medicine*, 201(1), pp. 49–54. Available at: https://doi.org/10.1084/jem.20040559.

Polakis, P. (2000) 'Wnt signaling and cancer', *Genes & Development*, 14(15), pp. 1837–1851. Available at: https://doi.org/10.1101/gad.14.15.1837.

Pollakis, G. and A Paxton, W. (2010) 'HIV-1 (co)Receptors: Implications for Vaccine and Therapy Design', *Current Pharmaceutical Design*, 16(33), pp. 3701–3715. Available at: https://doi.org/10.2174/138161210794079146.

Pollock, N.L., Lee, S.C., Patel, J.H., Gulamhussein, A.A. and Rothnie, A.J. (2018) 'Structure and function of membrane proteins encapsulated in a polymer-bound lipid bilayer', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1860(4), pp. 809–817. Available at: https://doi.org/10.1016/j.bbamem.2017.08.012.

Polly, P.D., Stayton, C.T., Dumont, E.R., Pierce, S.E., Rayfield, E.J. and Angielczyk, K.D. (2016) 'Combining geometric morphometrics and finite element analysis with evolutionary modeling: towards a synthesis', *Journal of Vertebrate Paleontology*, 36(4), p. e1111225. Available at: https://doi.org/10.1080/02724634.2016.1111225.

Popot, J.-L. and Engelman, D.M. (2000) 'Helical Membrane Protein Folding, Stability, and Evolution', *Annual Review of Biochemistry*, 69(1), pp. 881–922. Available at: https://doi.org/10.1146/annurev.biochem.69.1.881.

Poyner, D. and Wheatley, M. (2009) *G Protein-Coupled Receptors: Essential Methods*. John Wiley & Sons.

Poyner, D.R. and Hay, D.L. (2012) 'Secretin family (Class B) G protein-coupled receptors – from molecular to clinical perspectives', *British Journal of Pharmacology*, 166(1), pp. 1–3. Available at: https://doi.org/10.1111/j.1476-5381.2011.01810.x.

Prakash, J., Herlin, M., Kumar, J., Garg, G., Akesson, K.E., Grabowski, P.S., Skerry, T.M., Richards, G.O. and McGuigan, F.E.A. (2019) 'Analysis of RAMP3 gene polymorphism with body composition and bone density in young and elderly women', *Gene: X*, 2, p. 100009. Available at: https://doi.org/10.1016/j.gene.2019.100009.

Premont, R.T., Macrae, A.D., Aparicio, S.A.J.R., Kendall, H.E., Welch, J.E. and Lefkowitz, R.J. (1999) 'The GRK4 Subfamily of G Protein-coupled Receptor Kinases: ALTERNATIVE SPLICING, GENE ORGANIZATION, AND SEQUENCE CONSERVATION *', *Journal of Biological Chemistry*, 274(41), pp. 29381–29389. Available at: https://doi.org/10.1074/jbc.274.41.29381.

Prosser, R.S. and Alonzi, N.A. (2023) 'Discerning conformational dynamics and binding kinetics of GPCRs by 19F NMR', *Current Opinion in Pharmacology*, 72, p. 102377. Available at: https://doi.org/10.1016/j.coph.2023.102377.

Punjani, A. and Fleet, D.J. (2021) '3D Variability Analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM'. bioRxiv, p. 2020.04.08.032466. Available at: https://doi.org/10.1101/2020.04.08.032466.

Pyne, S. and Pyne, N.J. (2000) 'Sphingosine 1-phosphate signalling in mammalian cells', *Biochemical Journal*, 349(2), pp. 385–402. Available at: https://doi.org/10.1042/bj3490385.

Qi, X., Liu, H., Thompson, B., McDonald, J., Zhang, C. and Li, X. (2019) 'Cryo-EM structure of oxysterol-bound human Smoothened coupled to a heterotrimeric Gi', *Nature*, 571(7764), pp. 279–283. Available at: https://doi.org/10.1038/s41586-019-1286-0.

Qureshi, S.A., Rios Candelore, M., Xie, D., Yang, X., Tota, L.M., Ding, V.D.-H., Li, Z., Bansal, A., Miller, C., Cohen, S.M., Jiang, G., Brady, E., Saperstein, R., Duffy, J.L., Tata, J.R., Chapman, K.T., Moller, D.E. and Zhang, B.B. (2004) 'A Novel Glucagon Receptor Antagonist Inhibits Glucagon-Mediated Biological Effects', *Diabetes*, 53(12), pp. 3267–3273. Available at: https://doi.org/10.2337/diabetes.53.12.3267.

Rajagopalan, L. and Rajarathnam, K. (2006) 'Structural basis of chemokine receptor function--a model for binding affinity and ligand selectivity', *Bioscience Reports*, 26(5), pp. 325–339. Available at: https://doi.org/10.1007/s10540-006-9025-9.

Ramadurai, S., Holt, A., Krasnikov, V., van den Bogaart, G., Killian, J.A. and Poolman, B. (2009) 'Lateral Diffusion of Membrane Proteins', *Journal of the American Chemical Society*, 131(35), pp. 12650–12656. Available at: https://doi.org/10.1021/ja902853g.

Ramón, A. and Marín, M. (2011) 'Advances in the production of membrane proteins in Pichia pastoris', *Biotechnology Journal*, 6(6), pp. 700–706. Available at: https://doi.org/10.1002/biot.201100146.

Rao, V.S., Srinivas, K., Sujini, G.N. and Kumar, G.N.S. (2014) 'Protein-Protein Interaction Detection: Methods and Analysis', *International Journal of Proteomics*, 2014, p. 147648. Available at: https://doi.org/10.1155/2014/147648.

Rasmussen, Søren 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., Schnapp, A., Konetzki, I., Sunahara, R.K., Gellman, S.H., Pautsch, A., Steyaert, J., Weis, W.I. and Kobilka, B.K. (2011) 'Structure of a nanobody-stabilized active state of the β2 adrenoceptor', *Nature*, 469(7329), pp. 175–180. Available at: https://doi.org/10.1038/nature09648.

Rasmussen, S.G.F., Choi, H.-J., Rosenbaum, D.M., Kobilka, T.S., Thian, F.S., Edwards, P.C., Burghammer, M., Ratnala, V.R.P., Sanishvili, R., Fischetti, R.F., Schertler, G.F.X., Weis, W.I. and Kobilka, B.K. (2007) 'Crystal structure of the human β2 adrenergic G-protein-coupled receptor', *Nature*, 450(7168), pp. 383–387. Available at: https://doi.org/10.1038/nature06325.

Rasmussen, Søren G.F., DeVree, B.T., Zou, Y., Kruse, A.C., Chung, K.Y., Kobilka, T.S., Thian, F.S., Chae, P.S., Pardon, E., Calinski, D., Mathiesen, J.M., Shah, S.T.A., Lyons, J.A., Caffrey, M., Gellman, S.H., Steyaert, J., Skiniotis, G., Weis, W.I., Sunahara, R.K. and Kobilka, B.K. (2011) 'Crystal Structure of the β2Adrenergic Receptor-Gs protein complex', *Nature*, 477(7366), pp. 549–555. Available at: https://doi.org/10.1038/nature10361.

Ratner, R.E., Dickey, R., Fineman, M., Maggs, D.G., Shen, L., Strobel, S.A., Weyer, C. and Kolterman, O.G. (2004) 'Amylin replacement with pramlintide as an adjunct to insulin therapy improves long-term glycaemic and weight control in Type 1 diabetes mellitus: a 1-year, randomized controlled trial', *Diabetic Medicine: A Journal of the British Diabetic Association*, 21(11), pp. 1204–1212. Available at: https://doi.org/10.1111/j.1464-5491.2004.01319.x.

Raunser, S. and Walz, T. (2009) 'Electron crystallography as a technique to study the structure on membrane proteins in a lipidic environment', *Annual Review of Biophysics*, 38, pp. 89–105. Available at: https://doi.org/10.1146/annurev.biophys.050708.133649.

Rayasam, G.V., Tulasi, V.K., Davis, J.A. and Bansal, V.S. (2007) 'Fatty acid receptors as new therapeutic targets for diabetes', *Expert Opinion on Therapeutic Targets*, 11(5), pp. 661–671. Available at: https://doi.org/10.1517/14728222.11.5.661.

Reeves, P.J., Callewaert, N., Contreras, R. and Khorana, H.G. (2002) 'Structure and function in rhodopsin: High-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line', *Proceedings of the National Academy of Sciences*, 99(21), pp. 13419–13424. Available at: https://doi.org/10.1073/pnas.212519299.

Reeves, P.J., Kim, J.-M. and Khorana, H.G. (2002) 'Structure and function in rhodopsin: A tetracycline-inducible system in stable mammalian cell lines for high-level expression of opsin mutants', *Proceedings of the National Academy of Sciences*, 99(21), pp. 13413–13418. Available at: https://doi.org/10.1073/pnas.212519199.

Regen, S.L. (2020) 'The Origin of Lipid Rafts', *Biochemistry*, 59(49), pp. 4617–4621. Available at: https://doi.org/10.1021/acs.biochem.0c00851.

Reich, R. and Martin, G.R. (1996) 'Identification of arachidonic acid pathways required for the invasive and metastatic activity of malignant tumor cells', *Prostaglandins*, 51(1), pp. 1–17. Available at: https://doi.org/10.1016/0090-6980(95)00154-9.

Renaud, J.-P., Chari, A., Ciferri, C., Liu, W., Rémigy, H.-W., Stark, H. and Wiesmann, C. (2018) 'Cryo-EM in drug discovery: achievements, limitations and prospects', *Nature Reviews Drug Discovery*, 17(7), pp. 471–492. Available at: https://doi.org/10.1038/nrd.2018.77.

Reuter, U., Goadsby, P.J., Lanteri-Minet, M., Wen, S., Hours-Zesiger, P., Ferrari, M.D. and Klatt, J. (2018) 'Efficacy and tolerability of erenumab in patients with episodic migraine in whom two-to-four previous preventive treatments were unsuccessful: a randomised, double-blind, placebo-controlled, phase 3b study', *Lancet (London, England)*, 392(10161), pp. 2280–2287. Available at: https://doi.org/10.1016/S0140-6736(18)32534-0.

Rivera, J., Proia, R.L. and Olivera, A. (2008) 'The alliance of sphingosine-1-phosphate and its receptors in immunity', *Nature Reviews Immunology*, 8(10), pp. 753–763. Available at: https://doi.org/10.1038/nri2400.

Robert, X., Kassis-Sahyoun, J., Ceres, N., Martin, J., Sawaya, M.R., Read, R.J., Gouet, P., Falson, P. and Chaptal, V. (2017) 'X-ray diffraction reveals the intrinsic difference in the physical properties of membrane and soluble proteins', *Scientific Reports*, 7(1), p. 17013. Available at: https://doi.org/10.1038/s41598-017-17216-1.

Robertson, N., Jazayeri, A., Errey, J., Baig, A., Hurrell, E., Zhukov, A., Langmead, C.J., Weir, M. and Marshall, F.H. (2011) 'The properties of thermostabilised G protein-coupled receptors (StaRs) and their use in drug discovery', *Neuropharmacology*, 60(1), pp. 36–44. Available at: https://doi.org/10.1016/j.neuropharm.2010.07.001.

Rosa, M., Noel, T., Harris, M. and Ladds, G. (2021) 'Emerging roles of adhesion G protein-coupled receptors', *Biochemical Society Transactions*, 49(4), pp. 1695–1709. Available at: https://doi.org/10.1042/BST20201144.

Rosenbaum, D.M., Cherezov, V., Hanson, M.A., Rasmussen, S.G.F., Thian, F.S., Kobilka, T.S., Choi, H.-J., Yao, X.-J., Weis, W.I., Stevens, R.C. and Kobilka, B.K. (2007) 'GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function', *Science (New York, N.Y.)*, 318(5854), pp. 1266–1273. Available at: https://doi.org/10.1126/science.1150609.

Rosendo-Pineda, M.J., Moreno, C.M. and Vaca, L. (2020) 'Role of ion channels during cell division', *Cell Calcium*, 91, p. 102258. Available at: https://doi.org/10.1016/j.ceca.2020.102258.

Rosenkilde, M.M., Benned-Jensen, T., Frimurer, T.M. and Schwartz, T.W. (2010) 'The minor binding pocket: a major player in 7TM receptor activation', *Trends in Pharmacological Sciences*, 31(12), pp. 567–574. Available at: https://doi.org/10.1016/j.tips.2010.08.006.

Routledge, S.J., Mikaliunaite, L., Patel, A., Clare, M., Cartwright, S.P., Bawa, Z., Wilks, M.D.B., Low, F., Hardy, D., Rothnie, A.J. and Bill, R.M. (2016) 'The synthesis of recombinant membrane proteins in yeast for structural studies', *Methods (San Diego, Calif.)*, 95, pp. 26–37. Available at: https://doi.org/10.1016/j.ymeth.2015.09.027.

Routledge, S.J., Simms, J., Clark, A., Yeung, H.Y., Wigglesworth, M.J., Dickerson, I.M., Kitchen, P., Ladds, G. and Poyner, D.R. (2020) 'Receptor component protein, an endogenous allosteric modulator of family B G protein coupled receptors', *Biochimica Et Biophysica Acta. Biomembranes*, 1862(3), p. 183174. Available at: https://doi.org/10.1016/j.bbamem.2019.183174.

Rozengurt, E., Sinnett-Smith, J. and Kisfalvi, K. (2010) 'Crosstalk between Insulin/Insulin-like Growth Factor-1 Receptors and G Protein-Coupled Receptor Signaling Systems: A Novel Target for the Antidiabetic Drug Metformin in Pancreatic Cancer', *Clinical Cancer Research*, 16(9), pp. 2505–2511. Available at: https://doi.org/10.1158/1078-0432.CCR-09-2229.

Russell, F.A., King, R., Smillie, S.-J., Kodji, X. and Brain, S.D. (2014) 'Calcitonin Gene-Related Peptide: Physiology and Pathophysiology', *Physiological Reviews*, 94(4), pp. 1099–1142. Available at: https://doi.org/10.1152/physrev.00034.2013.

Ryan, G.J., Jobe, L.J. and Martin, R. (2005) 'Pramlintide in the treatment of type 1 and type 2 diabetes mellitus', *Clinical Therapeutics*, 27(10), pp. 1500–1512. Available at: https://doi.org/10.1016/j.clinthera.2005.10.009.

Saarenpää, T., Jaakola, V.-P. and Goldman, A. (2015) 'Chapter Nine - Baculovirus-Mediated Expression of GPCRs in Insect Cells', in A.K. Shukla (ed.) *Methods in Enzymology*. Academic Press (Membrane Proteins—Production and Functional Characterization), pp. 185–218. Available at: https://doi.org/10.1016/bs.mie.2014.12.033.

Sahin, C., Reid, D.J., Marty, M.T. and Landreh, M. (2020) 'Scratching the surface: native mass spectrometry of peripheral membrane protein complexes', *Biochemical Society Transactions*, 48(2), pp. 547–558. Available at: https://doi.org/10.1042/BST20190787.

Saito, Y. and Nagasaki, H. (2008) 'The Melanin-Concentrating Hormone System and Its Physiological Functions', in O. Civelli and Q.-Y. Zhou (eds) *Orphan G Protein-Coupled Receptors and Novel Neuropeptides*. Berlin, Heidelberg: Springer (Results and Problems in Cell Differentiation), pp. 159–179. Available at: https://doi.org/10.1007/400_2007_052.

Sala, D., Engelberger, F., Mchaourab, H.S. and Meiler, J. (2023) 'Modeling conformational states of proteins with AlphaFold', *Current Opinion in Structural Biology*, 81, p. 102645. Available at: https://doi.org/10.1016/j.sbi.2023.102645.

Sallese, M., Mariggiò, S., D'Urbano, E., Iacovelli, L. and Blasi, A.D. (2000) 'Selective Regulation of Gq Signaling by G Protein-Coupled Receptor Kinase 2: Direct Interaction of Kinase N Terminus with Activated Gαq', *Molecular Pharmacology*, 57(4), pp. 826–831. Available at: https://doi.org/10.1124/mol.57.4.826.

Salman, M.M., Kitchen, P., Halsey, A., Wang, M.X., Törnroth-Horsefield, S., Conner, A.C., Badaut, J., Iliff, J.J. and Bill, R.M. (2022) 'Emerging roles for dynamic aquaporin-4 subcellular relocalization in CNS water homeostasis', *Brain: A Journal of Neurology*, 145(1), pp. 64–75. Available at: https://doi.org/10.1093/brain/awab311.

Sanders, C.R. and Prosser, R.S. (1998) 'Bicelles: a model membrane system for all seasons?', *Structure*, 6(10), pp. 1227–1234. Available at: https://doi.org/10.1016/S0969-2126(98)00123-3.

Santos, R., Ursu, O., Gaulton, A., Bento, A.P., Donadi, R.S., Bologa, C.G., Karlsson, A., Al-Lazikani, B., Hersey, A., Oprea, T.I. and Overington, J.P. (2017) 'A comprehensive map of molecular drug targets', *Nature Reviews Drug Discovery*, 16(1), pp. 19–34. Available at: https://doi.org/10.1038/nrd.2016.230.

Sarkar, K., Joedicke, L., Westwood, M., Burnley, R., Wright, M., McMillan, D. and Byrne, B. (2019) 'Modulation of PTH1R signaling by an ECD binding antibody results in inhibition of β-arrestin 2 coupling', *Scientific Reports*, 9, p. 14432. Available at: https://doi.org/10.1038/s41598-019-51016-z.

Schauperl, M. and Denny, R.A. (2022) 'AI-Based Protein Structure Prediction in Drug Discovery: Impacts and Challenges', *Journal of Chemical Information and Modeling*, 62(13), pp. 3142–3156. Available at: https://doi.org/10.1021/acs.jcim.2c00026.

Scheerer, P., Park, J.H., Hildebrand, P.W., Kim, Y.J., Krauß, N., Choe, H.-W., Hofmann, K.P. and Ernst, O.P. (2008) 'Crystal structure of opsin in its G-protein-interacting conformation', *Nature*, 455(7212), pp. 497–502. Available at: https://doi.org/10.1038/nature07330.

Schiller, B., Hykollari, A., Yan, S., Paschinger, K. and Wilson, I.B.H. (2012) 'Complicated N-linked glycans in simple organisms', *Biological Chemistry*, 393(8), pp. 661–673. Available at: https://doi.org/10.1515/hsz-2012-0150.

Schinkel, A.H. and Jonker, J.W. (2012) 'Mammalian drug efflux transporters of the ATP binding cassette (ABC) family: an overview', *Advanced Drug Delivery Reviews*, 64, pp. 138–153. Available at: https://doi.org/10.1016/j.addr.2012.09.027.

Schmidt, T.G. and Skerra, A. (2007) 'The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins', *Nature Protocols*, 2(6), pp. 1528–1535. Available at: https://doi.org/10.1038/nprot.2007.209.

Schwartz, T.W. and Frimurer, T.M. (2017) 'Full monty of family B GPCRs', *Nature Chemical Biology*, 13(8), pp. 819–821. Available at: https://doi.org/10.1038/nchembio.2438.

Schwartz, T.W., Frimurer, T.M., Holst, B., Rosenkilde, M.M. and Elling, C.E. (2006) 'Molecular mechanism of 7TM receptor activation--a global toggle switch model', *Annual Review of Pharmacology and Toxicology*, 46, pp. 481–519. Available at: https://doi.org/10.1146/annurev.pharmtox.46.120604.141218.

Scott, H.L., Kennison, K.B., Enoki, T.A., Doktorova, M., Kinnun, J.J., Heberle, F.A. and Katsaras, J. (2021) 'Model Membrane Systems Used to Study Plasma Membrane Lipid Asymmetry', *Symmetry*, 13(8), p. 1356. Available at: https://doi.org/10.3390/sym13081356.

Scuteri, D., Tonin, P., Nicotera, P., Bagetta, G. and Corasaniti, M.T. (2022) 'Real world considerations for newly approved CGRP receptor antagonists in migraine care', *Expert Review of Neurotherapeutics*, 22(3), pp. 221–230. Available at: https://doi.org/10.1080/14737175.2022.2049758.

Seddon, A.M., Curnow, P. and Booth, P.J. (2004) 'Membrane proteins, lipids and detergents: not just a soap opera', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1666(1), pp. 105–117. Available at: https://doi.org/10.1016/j.bbamem.2004.04.011.

Sgro, G.G. and Costa, T.R.D. (2018) 'Cryo-EM Grid Preparation of Membrane Protein Samples for Single Particle Analysis', *Frontiers in Molecular Biosciences*, 5, p. 74. Available at: https://doi.org/10.3389/fmolb.2018.00074.

Shao, L., Chen, Y., Zhang, S., Zhang, Z., Cao, Y., Yang, D. and Wang, M.-W. (2022) 'Modulating effects of RAMPs on signaling profiles of the glucagon receptor family', *Acta Pharmaceutica Sinica. B*, 12(2), pp. 637–650. Available at: https://doi.org/10.1016/j.apsb.2021.07.028.

Shashkova, S. and Leake, M.C. (2017) 'Single-molecule fluorescence microscopy review: shedding new light on old problems', *Bioscience Reports*, 37(4), p. BSR20170031. Available at: https://doi.org/10.1042/BSR20170031.

Shi, C.-S., Lee, S.B., Sinnarajah, S., Dessauer, C.W., Rhee, S.G. and Kehrl, J.H. (2001) 'Regulator of G-protein Signaling 3 (RGS3) Inhibits Gβ1γ2-induced Inositol Phosphate Production, Mitogen-activated Protein Kinase Activation, and Akt Activation *', *Journal of Biological Chemistry*, 276(26), pp. 24293–24300. Available at: https://doi.org/10.1074/jbc.M100089200.

Shibata, Y., White, J.F., Serrano-Vega, M.J., Magnani, F., Aloia, A.L., Grisshammer, R. and Tate, C.G. (2009) 'Thermostabilization of the Neurotensin Receptor NTS1', *Journal of Molecular Biology*, 390(2), pp. 262–277. Available at: https://doi.org/10.1016/j.jmb.2009.04.068.

Shihoya, W., Nishizawa, T., Okuta, A., Tani, K., Dohmae, N., Fujiyoshi, Y., Nureki, O. and Doi, T. (2016) 'Activation mechanism of endothelin ETB receptor by endothelin-1', *Nature*, 537(7620), pp. 363–368. Available at: https://doi.org/10.1038/nature19319.

Shimizu, K., Cao, W., Saad, G., Shoji, M. and Terada, T. (2018) 'Comparative analysis of membrane protein structure databases', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1860(5), pp. 1077–1091. Available at: https://doi.org/10.1016/j.bbamem.2018.01.005.

Shimizu, N., Tanaka, A., Oue, A., Mori, T., Apichartpiyakul, C. and Hoshino, H. (2008) 'A short amino acid sequence containing tyrosine in the N-terminal region of G protein-coupled receptors is critical for their potential use as co-receptors for human and simian immunodeficiency viruses', *Journal of General Virology*, 89(12), pp. 3126–3136. Available at: https://doi.org/10.1099/vir.0.2008/002188-0.

Shukla, A.K., Manglik, A., Kruse, A.C., Xiao, K., Reis, R.I., Tseng, W.-C., Staus, D.P., Hilger, D., Uysal, S., Huang, L.-Y., Paduch, M., Tripathi-Shukla, P., Koide, A., Koide, S., Weis, W.I., Kossiakoff, A.A., Kobilka, B.K. and Lefkowitz, R.J. (2013) 'Structure of active β-arrestin-1

bound to a G-protein-coupled receptor phosphopeptide', *Nature*, 497(7447), pp. 137–141. Available at: https://doi.org/10.1038/nature12120.

Siderovski, D.P., Hessel, A., Chung, S., Mak, T.W. and Tyers, M. (1996) 'A new family of regulators of G-protein-coupled receptors?', *Current Biology*, 6(2), pp. 211–212. Available at: https://doi.org/10.1016/S0960-9822(02)00454-2.

Simon, M.I., Strathmann, M.P. and Gautam, N. (1991) 'Diversity of G Proteins in Signal Transduction', *Science*, 252(5007), pp. 802–808. Available at: https://doi.org/10.1126/science.1902986.

Singer, S.J. and Nicolson, G.L. (1972) 'The fluid mosaic model of the structure of cell membranes', *Science (New York, N.Y.)*, 175(4023), pp. 720–731. Available at: https://doi.org/10.1126/science.175.4023.720.

Singh, D., Mathur, A., Arora, S., Roy, S. and Mahindroo, N. (2022) 'Journey of organ on a chip technology and its role in future healthcare scenario', *Applied Surface Science Advances*, 9, p. 100246. Available at: https://doi.org/10.1016/j.apsadv.2022.100246.

Singh, S., Gras, A., Fiez-Vandal, C., Ruprecht, J., Rana, R., Martinez, M., Strange, P.G., Wagner, R. and Byrne, B. (2008) 'Large-scale functional expression of WT and truncated human adenosine A2A receptor in Pichia pastoris bioreactor cultures', *Microbial Cell Factories*, 7(1), p. 28. Available at: https://doi.org/10.1186/1475-2859-7-28.

Slice, D.E. (2005) 'Modern Morphometrics', in D.E. Slice (ed.) *Modern Morphometrics in Physical Anthropology*. Boston, MA: Springer US (Developments in Primatology: Progress and Prospects), pp. 1–45. Available at: https://doi.org/10.1007/0-387-27614-9_1.

Sligar, S.G. and Denisov, I.G. (2021) 'Nanodiscs: A toolkit for membrane protein science', *Protein Science : A Publication of the Protein Society*, 30(2), pp. 297–315. Available at: https://doi.org/10.1002/pro.3994.

Slinger, E., Langemeijer, E., Siderius, M., Vischer, H.F. and Smit, M.J. (2011) 'Herpesvirus-encoded GPCRs rewire cellular signaling', *Molecular and Cellular Endocrinology*, 331(2), pp. 179–184. Available at: https://doi.org/10.1016/j.mce.2010.04.007.

Slot, A.J., Molinski, S.V. and Cole, S.P.C. (2011) 'Mammalian multidrug-resistance proteins (MRPs)', *Essays in Biochemistry*. Edited by F.J. Sharom, 50, pp. 179–207. Available at: https://doi.org/10.1042/bse0500179.

Smith, S.O. (2010) 'Structure and activation of the visual pigment rhodopsin', *Annual Review of Biophysics*, 39, pp. 309–328. Available at: https://doi.org/10.1146/annurev-biophys-101209-104901.

Sobhanifar, S., Reckel, S., Junge, F., Schwarz, D., Kai, L., Karbyshev, M., Löhr, F., Bernhard, F. and Dötsch, V. (2009) 'Cell-free expression and stable isotope labelling strategies for membrane proteins', *Journal of Biomolecular NMR*, 46(1), p. 33. Available at: https://doi.org/10.1007/s10858-009-9364-5.

Sonnleitner, A., Schütz, G.J. and Schmidt, Th. (1999) 'Free Brownian Motion of Individual Lipid Molecules in Biomembranes', *Biophysical Journal*, 77(5), pp. 2638–2642. Available at: https://doi.org/10.1016/S0006-3495(99)77097-9.

Soudy, R., Kimura, R., Patel, A., Fu, W., Kaur, K., Westaway, D., Yang, J. and Jhamandas, J. (2019) 'Short amylin receptor antagonist peptides improve memory deficits in Alzheimer's

disease mouse model', *Scientific Reports*, 9, p. 10942. Available at: https://doi.org/10.1038/s41598-019-47255-9.

Srivastava, A., Yano, J., Hirozane, Y., Kefala, G., Gruswitz, F., Snell, G., Lane, W., Ivetac, A., Aertgeerts, K., Nguyen, J., Jennings, A. and Okada, K. (2014) 'High-resolution structure of the human GPR40 receptor bound to allosteric agonist TAK-875', *Nature*, 513(7516), pp. 124–127. Available at: https://doi.org/10.1038/nature13494.

St. Croix, C.M., Shand, S.H. and Watkins, S.C. (2005) 'Confocal microscopy: comparisons, applications, and problems', *BioTechniques*, 39(6S), pp. S2–S5. Available at: https://doi.org/10.2144/000112089.

Standfuss, J., Edwards, P.C., D'Antona, A., Fransen, M., Xie, G., Oprian, D.D. and Schertler, G.F.X. (2011) 'The structural basis of agonist-induced activation in constitutively active rhodopsin', *Nature*, 471(7340), pp. 656–660. Available at: https://doi.org/10.1038/nature09795.

Staus, D.P., Strachan, R.T., Manglik, A., Pani, B., Kahsai, A.W., Kim, T.H., Wingler, L.M., Ahn, S., Chatterjee, A., Masoudi, A., Kruse, A.C., Pardon, E., Steyaert, J., Weis, W.I., Prosser, R.S., Kobilka, B.K., Costa, T. and Lefkowitz, R.J. (2016) 'Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation', *Nature*, 535(7612), pp. 448–452. Available at: https://doi.org/10.1038/nature18636.

Steiner, S., Muff, R., Gujer, R., Fischer, J.A. and Born, W. (2002) 'The transmembrane domain of receptor-activity-modifying protein 1 is essential for the functional expression of a calcitonin gene-related peptide receptor', *Biochemistry*, 41(38), pp. 11398–11404. Available at: https://doi.org/10.1021/bi020279r.

Stetsenko, A. and Guskov, A. (2017) 'An Overview of the Top Ten Detergents Used for Membrane Protein Crystallization', *Crystals*, 7(7), p. 197. Available at: https://doi.org/10.3390/cryst7070197.

Stewart, A. and Fisher, R.A. (2015) 'Introduction: G Protein-coupled Receptors and RGS Proteins', *Progress in Molecular Biology and Translational Science*, 133, pp. 1–11. Available at: https://doi.org/10.1016/bs.pmbts.2015.03.002.

Strader, C.D., Fong, T.M., Tota, M.R., Underwood, D. and Dixon, R.A.F. (1994) 'Structure and Function of G Protein-Coupled Receptors', *Annual Review of Biochemistry*, 63(1), pp. 101–132. Available at: https://doi.org/10.1146/annurev.bi.63.070194.000533.

Strieter, R.M., Burdick, M.D., Mestas, J., Gomperts, B., Keane, M.P. and Belperio, J.A. (2006) 'Cancer CXC chemokine networks and tumour angiogenesis', *European Journal of Cancer*, 42(6), pp. 768–778. Available at: https://doi.org/10.1016/j.ejca.2006.01.006.

Strosberg, A.D. (1993) 'Structure, function, and regulation of adrenergic receptors', *Protein Science: A Publication of the Protein Society*, 2(8), pp. 1198–1209. Available at: https://doi.org/10.1002/pro.5560020802.

Studier, F.W. (1991) 'Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system', *Journal of Molecular Biology*, 219(1), pp. 37–44. Available at: https://doi.org/10.1016/0022-2836(91)90855-Z.

Sun, L. and Ye, R.D. (2012) 'Role of G protein-coupled receptors in inflammation', *Acta Pharmacologica Sinica*, 33(3), pp. 342–350. Available at: https://doi.org/10.1038/aps.2011.200.

Sun, Y., Wallrabe, H., Seo, S.-A. and Periasamy, A. (2011) 'FRET microscopy in 2010: The legacy of Theodor Förster on the 100th anniversary of his birth', *Chemphyschem : a European journal of chemical physics and physical chemistry*, 12(3), pp. 462–474. Available at: https://doi.org/10.1002/cphc.201000664.

Suresh, P. and Wanchu, A. (2006) 'Chemokines and chemokine receptors in HIV infection: Role in pathogenesis and therapeutics', *Journal of Postgraduate Medicine*, 52(3), p. 210.

Sušac, L., Horst, R. and Wüthrich, K. (2014) 'Outer Membrane Protein A from E. coli Characterized in Lipid Bilayer Nanodiscs and Detergent Micelles by NMR in Solution', *Chembiochem : a European journal of chemical biology*, 15(7), pp. 995–1000. Available at: https://doi.org/10.1002/cbic.201300729.

Sutkeviciute, I. and Vilardaga, J.-P. (2020) 'Structural insights into emergent signaling modes of G protein–coupled receptors', *The Journal of Biological Chemistry*, 295(33), pp. 11626–11642. Available at: https://doi.org/10.1074/jbc.REV120.009348.

Syrovatkina, V., Alegre, K.O., Dey, R. and Huang, X.-Y. (2016) 'Regulation, Signaling, and Physiological Functions of G-Proteins', *Journal of Molecular Biology*, 428(19), pp. 3850–3868. Available at: https://doi.org/10.1016/j.jmb.2016.08.002.

Takegawa, K., Tohda, H., Sasaki, M., Idiris, A., Ohashi, T., Mukaiyama, H., Giga-Hama, Y. and Kumagai, H. (2009) 'Production of heterologous proteins using the fission-yeast (Schizosaccharomyces pombe) expression system', *Biotechnology and Applied Biochemistry*, 53(Pt 4), pp. 227–235. Available at: https://doi.org/10.1042/BA20090048.

Tan, S., Tan, H.T. and Chung, M.C.M. (2008) 'Membrane proteins and membrane proteomics', *Proteomics*, 8(19), pp. 3924–3932. Available at: https://doi.org/10.1002/pmic.200800597.

Tang, C.-M. and Insel, P.A. (2004) 'GPCR Expression in the Heart: "New" Receptors in Myocytes and Fibroblasts', *Trends in Cardiovascular Medicine*, 14(3), pp. 94–99. Available at: https://doi.org/10.1016/j.tcm.2003.12.007.

Tang, W., Knox, R.W. and Nevzorov, A.A. (2012) 'A spectroscopic assignment technique for membrane proteins reconstituted in magnetically aligned bicelles', *Journal of biomolecular NMR*, 54(3), pp. 307–316. Available at: https://doi.org/10.1007/s10858-012-9673-y.

Tapaneeyakorn, S., Goddard, A.D., Oates, J., Willis, C.L. and Watts, A. (2011) 'Solution- and solid-state NMR studies of GPCRs and their ligands', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1808(6), pp. 1462–1475. Available at: https://doi.org/10.1016/j.bbamem.2010.10.003.

Tarling, E.J., Vallim, T.Q. de A. and Edwards, P.A. (2013) 'Role of ABC transporters in lipid transport and human disease', *Trends in Endocrinology & Metabolism*, 24(7), pp. 342–350. Available at: https://doi.org/10.1016/j.tem.2013.01.006.

Tate, C. g. (2001) 'Overexpression of mammalian integral membrane proteins for structural studies', *FEBS Letters*, 504(3), pp. 94–98. Available at: https://doi.org/10.1016/S0014-5793(01)02711-9.

Taylor, C.W. and Laude, A.J. (2002) 'IP3 receptors and their regulation by calmodulin and cytosolic Ca2+', *Cell Calcium*, 32(5), pp. 321–334. Available at: https://doi.org/10.1016/S0143416002001859.

Taylor, P.D., Toseland, C.P., Attwood, T.K. and Flower, D.R. (2006) 'A predictor of membrane class: Discriminating α-helical and β-barrel membrane proteins from non-membranous proteins', *Bioinformation*, 1(6), pp. 208–213.

Tehan, B.G., Bortolato, A., Blaney, F.E., Weir, M.P. and Mason, J.S. (2014) 'Unifying Family A GPCR Theories of Activation', *Pharmacology & Therapeutics*, 143(1), pp. 51–60. Available at: https://doi.org/10.1016/j.pharmthera.2014.02.004.

Tepper, S.J. (2018) 'History and Review of anti-Calcitonin Gene-Related Peptide (CGRP) Therapies: From Translational Research to Treatment', *Headache*, 58 Suppl 3, pp. 238–275. Available at: https://doi.org/10.1111/head.13379.

Teske, J.A., Billington, C.J. and Kotz, C.M. (2010) 'Hypocretin/orexin and energy expenditure', *Acta Physiologica*, 198(3), pp. 303–312. Available at: https://doi.org/10.1111/j.1748-1716.2010.02075.x.

Thoma, J. and Burmann, B.M. (2021) 'Fake It 'Till You Make It—The Pursuit of Suitable Membrane Mimetics for Membrane Protein Biophysics', *International Journal of Molecular Sciences*, 22(1), p. 50. Available at: https://doi.org/10.3390/ijms22010050.

Thomas, J.A. and Tate, C.G. (2014) 'Quality Control in Eukaryotic Membrane Protein Overproduction', *Journal of Molecular Biology*, 426(24), pp. 4139–4154. Available at: https://doi.org/10.1016/j.jmb.2014.10.012.

Thomas, S.M., Bhola, N.E., Zhang, Q., Contrucci, S.C., Wentzel, A.L., Freilino, M.L., Gooding, W.E., Siegfried, J.M., Chan, D.C. and Grandis, J.R. (2006) 'Cross-talk between G Protein–Coupled Receptor and Epidermal Growth Factor Receptor Signaling Pathways Contributes to Growth and Invasion of Head and Neck Squamous Cell Carcinoma', *Cancer Research*, 66(24), pp. 11831–11839. Available at: https://doi.org/10.1158/0008-5472.CAN-06-2876.

Thompson, M.D., Cole, D.E.C. and Jose, P.A. (2008) 'Pharmacogenomics of G Protein-Coupled Receptor', in Q. Yan (ed.) *Pharmacogenomics in Drug Discovery and Development: From Bench to Bedside*. Totowa, NJ: Humana Press (Methods in Molecular Biology™), pp. 77–107. Available at: https://doi.org/10.1007/978-1-59745-205-2_6.

Thomsen, A.R.B., Hvidtfeldt, M. and Bräuner-Osborne, H. (2012) 'Biased agonism of the calcium-sensing receptor', *Cell Calcium*, 51(2), pp. 107–116. Available at: https://doi.org/10.1016/j.ceca.2011.11.009.

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., Huang, L., Breton, B., Heydenreich, F.M., Sunahara, R.K., Skiniotis, G., Bouvier, M. and Lefkowitz, R.J. (2016) 'GPCR-G Protein-β-Arrestin Super-Complex Mediates Sustained G Protein Signaling', *Cell*, 166(4), pp. 907–919. Available at: https://doi.org/10.1016/j.cell.2016.07.004.

Tian, F., Song, Z. and Cross, T.A. (1998) 'Orientational constraints derived from hydrated powder samples by two-dimensional PISEMA', *Journal of Magnetic Resonance (San Diego, Calif.: 1997)*, 135(1), pp. 227–231. Available at: https://doi.org/10.1006/jmre.1998.1544.

Tiefenauer, L. and Demarche, S. (2012) 'Challenges in the Development of Functional Assays of Membrane Proteins', *Materials*, 5(11), pp. 2205–2242. Available at: https://doi.org/10.3390/ma5112205.

Tilley, D.G. and Rockman, H.A. (2006) 'Role of β-adrenergic receptor signaling and desensitization in heart failure: new concepts and prospects for treatment', *Expert Review of*

Cardiovascular Therapy, 4(3), pp. 417–432. Available at: https://doi.org/10.1586/14779072.4.3.417.

Tivol, W.F., Briegel, A. and Jensen, G.J. (2008) 'An Improved Cryogen for Plunge Freezing', *Microscopy and Microanalysis*, 14(5), pp. 375–379. Available at: https://doi.org/10.1017/S1431927608080781.

Tsou, P.-S., Lu, C., Gurrea-Rubio, M., Muraoka, S., Campbell, P.L., Wu, Q., Model, E.N., Lind, M.E., Vichaikul, S., Mattichak, M.N., Brodie, W.D., Hervoso, J.L., Ory, S., Amarista, C.I., Pervez, R., Junginger, L., Ali, M., Hodish, G., O'Mara, M.M., Ruth, J.H., Robida, A.M., Alt, A.J., Zhang, C., Urquhart, A.G., Lawton, J.N., Chung, K.C., Maerz, T., Saunders, T.L., Groppi, V.E., Fox, D.A. and Amin, M.A. (2022) 'Soluble CD13 induces inflammatory arthritis by activating the bradykinin receptor B1', *The Journal of Clinical Investigation*, 132(11), p. e151827. Available at: https://doi.org/10.1172/JCI151827.

Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., Velankar, S., Kleywegt, G.J., Bateman, A., Evans, R., Pritzel, A., Figurnov, M., Ronneberger, O., Bates, R., Kohl, S.A.A., Potapenko, A., Ballard, A.J., Romera-Paredes, B., Nikolov, S., Jain, R., Clancy, E., Reiman, D., Petersen, S., Senior, A.W., Kavukcuoglu, K., Birney, E., Kohli, P., Jumper, J. and Hassabis, D. (2021) 'Highly accurate protein structure prediction for the human proteome', *Nature*, 596(7873), pp. 590–596. Available at: https://doi.org/10.1038/s41586-021-03828-1.

Tuteja, N. (2009) 'Signaling through G protein coupled receptors', *Plant Signaling & Behavior*, 4(10), pp. 942–947.

Udawela, M., Christopoulos, G., Morfis, M., Christopoulos, A., Ye, S., Tilakaratne, N. and Sexton, P.M. (2006) 'A critical role for the short intracellular C terminus in receptor activity-modifying protein function', *Molecular Pharmacology*, 70(5), pp. 1750–1760. Available at: https://doi.org/10.1124/mol.106.024257.

Udawela, M., Christopoulos, G., Morfis, M., Tilakaratne, N., Christopoulos, A. and Sexton, P.M. (2008) 'The effects of C-terminal truncation of receptor activity modifying proteins on the induction of amylin receptor phenotype from human CTb receptors', *Regulatory Peptides*, 145(1–3), pp. 65–71. Available at: https://doi.org/10.1016/j.regpep.2007.08.003.

Udawela, M., Hay, D.L. and Sexton, P.M. (2004) 'The receptor activity modifying protein family of G protein coupled receptor accessory proteins', *Seminars in Cell & Developmental Biology*, 15(3), pp. 299–308. Available at: https://doi.org/10.1016/j.semcdb.2003.12.019.

Umezu-Goto, M., Tanyi, J., Lahad, J., Liu, S., Yu, S., Lapushin, R., Hasegawa, Y., Lu, Y., Trost, R., Bevers, T., Jonasch, E., Aldape, K., Liu, J., James, R.D., Ferguson, C.G., Xu, Y., Prestwich, G.D. and Mills, G.B. (2004) 'Lysophosphatidic acid production and action: Validated targets in cancer?', *Journal of Cellular Biochemistry*, 92(6), pp. 1115–1140. Available at: https://doi.org/10.1002/jcb.20113.

Unger, T. and Peleg, Y. (2012) 'Recombinant Protein Expression in the Baculovirus-Infected Insect Cell System', in E.D. Zanders (ed.) *Chemical Genomics and Proteomics: Reviews and Protocols*. Totowa, NJ: Humana Press (Methods in Molecular Biology), pp. 187–199. Available at: https://doi.org/10.1007/978-1-61779-349-3_13.

Vaidehi, N., Grisshammer, R. and Tate, C.G. (2016) 'How do mutations thermostabilize G protein-coupled receptors?', *Trends in pharmacological sciences*, 37(1), pp. 37–46. Available at: https://doi.org/10.1016/j.tips.2015.09.005.

Vaitsopoulou, A., Depping, P., Bill, R.M., Goddard, A.D. and Rothnie, A.J. (2022) 'Membrane Protein Production in Insect Cells', *Methods in Molecular Biology (Clifton, N.J.)*, 2507, pp. 223–240. Available at: https://doi.org/10.1007/978-1-0716-2368-8_12.

Van Horn, W.D., Kim, H.-J., Ellis, C.D., Hadziselimovic, A., Sulistijo, E.S., Karra, M.D., Tian, C., Sönnichsen, F.D. and Sanders, C.R. (2009) 'Solution NMR Structure of Membrane-Integral Diacylglycerol Kinase', *Science (New York, N.Y.)*, 324(5935), pp. 1726–1729. Available at: https://doi.org/10.1126/science.1171716.

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Žídek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., Figurnov, M., Cowie, A., Hobbs, N., Kohli, P., Kleywegt, G., Birney, E., Hassabis, D. and Velankar, S. (2022) 'AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models', *Nucleic Acids Research*, 50(D1), pp. D439–D444. Available at: https://doi.org/10.1093/nar/gkab1061.

Velazhahan, V., Ma, N., Pándy-Szekeres, G., Kooistra, A.J., Lee, Y., Gloriam, D.E., Vaidehi, N. and Tate, C.G. (2021) 'Structure of the class D GPCR Ste2 dimer coupled to two G proteins', *Nature*, 589(7840), pp. 148–153. Available at: https://doi.org/10.1038/s41586-020-2994-1.

Verkleij, A.J., Leunissen-Bijvelt, J., de Kruijff, B., Hope, M. and Cullis, P.R. (1984) 'Non-bilayer structures in membrane fusion', *Ciba Foundation Symposium*, 103, pp. 45–59. Available at: https://doi.org/10.1002/9780470720844.ch4.

Vestergaard, M., Kraft, J.F., Vosegaard, T., Thøgersen, L. and Schiøtt, B. (2015) 'Bicelles and Other Membrane Mimics: Comparison of Structure, Properties, and Dynamics from MD Simulations', *The Journal of Physical Chemistry B*, 119(52), pp. 15831–15843. Available at: https://doi.org/10.1021/acs.jpcb.5b08463.

Vieira Gomes, A.M., Souza Carmo, T., Silva Carvalho, L., Mendonça Bahia, F. and Parachin, N.S. (2018) 'Comparison of Yeasts as Hosts for Recombinant Protein Production', *Microorganisms*, 6(2), p. 38. Available at: https://doi.org/10.3390/microorganisms6020038.

Vischer, H.F., Vink, C. and Smit, M.J. (2006) 'A Viral Conspiracy: Hijacking the Chemokine System Through Virally Encoded Pirated Chemokine Receptors', in T.E. Lane (ed.) *Chemokines and Viral Infection*. Berlin, Heidelberg: Springer (Current Topics in Microbiology and Immunology), pp. 121–154. Available at: https://doi.org/10.1007/978-3-540-33397-5_6.

Vogel, R., Mahalingam, M., Lüdeke, S., Huber, T., Siebert, F. and Sakmar, T.P. (2008) 'Functional role of the "ionic lock"--an interhelical hydrogen-bond network in family A heptahelical receptors', *Journal of Molecular Biology*, 380(4), pp. 648–655. Available at: https://doi.org/10.1016/j.jmb.2008.05.022.

Wagner, S., Baars, L., Ytterberg, A.J., Klussmeier, A., Wagner, C.S., Nord, O., Nygren, P.-Å., van Wijk, K.J. and de Gier, J.-W. (2007) 'Consequences of Membrane Protein Overexpression in Escherichia coli*', *Molecular & Cellular Proteomics*, 6(9), pp. 1527–1550. Available at: https://doi.org/10.1074/mcp.M600431-MCP200.

Wagner, S., Klepsch, M.M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Högbom, M., van Wijk, K.J., Slotboom, D.J., Persson, J.O. and de Gier, J.-W. (2008) 'Tuning Escherichia coli for membrane protein overexpression', *Proceedings of the National Academy of Sciences*, 105(38), pp. 14371–14376. Available at: https://doi.org/10.1073/pnas.0804090105.

Wallukat, G. (2002) 'The beta-adrenergic receptors', *Herz*, 27(7), pp. 683–690. Available at: https://doi.org/10.1007/s00059-002-2434-z.

Wang, D. and DuBois, R.N. (2006) 'Prostaglandins and Cancer', *Gut*, 55(1), pp. 115–122. Available at: https://doi.org/10.1136/gut.2004.047100.

Wang, Y., Zhang, S., Li, F., Zhou, Y., Zhang, Y., Wang, Z., Zhang, R., Zhu, J., Ren, Y., Tan, Y., Qin, C., Li, Y., Li, X., Chen, Y. and Zhu, F. (2020) 'Therapeutic target database 2020: enriched resource for facilitating research and early development of targeted therapeutics', *Nucleic Acids Research*, 48(D1), pp. D1031–D1041. Available at: https://doi.org/10.1093/nar/gkz981.

Ward, B.K., Magno, A.L., Walsh, J.P. and Ratajczak, T. (2012) 'The role of the calcium-sensing receptor in human disease', *Clinical Biochemistry*, 45(12), pp. 943–953. Available at: https://doi.org/10.1016/j.clinbiochem.2012.03.034.

Warne, T., Serrano-Vega, M.J., Baker, J.G., Moukhametzianov, R., Edwards, P.C., Henderson, R., Leslie, A.G.W., Tate, C.G. and Schertler, G.F.X. (2008) 'Structure of a β1-adrenergic G protein-coupled receptor', *Nature*, 454(7203), pp. 486–491. Available at: https://doi.org/10.1038/nature07101.

Warne, T. and Tate, C.G. (2013) 'The importance of interactions with helix 5 in determining the efficacy of β-adrenoceptor ligands', *Biochemical Society Transactions*, 41(1), pp. 159–165. Available at: https://doi.org/10.1042/BST20120228.

Warner, D.R., Weng, G., Yu, S., Matalon, R. and Weinstein, L.S. (1998) 'A Novel Mutation in the Switch 3 Region of Gsα in a Patient with Albright Hereditary Osteodystrophy Impairs GDP Binding and Receptor Activation *', *Journal of Biological Chemistry*, 273(37), pp. 23976–23983. Available at: https://doi.org/10.1074/jbc.273.37.23976.

Watanabe, M., Houten, S.M., Mataki, C., Christoffolete, M.A., Kim, B.W., Sato, H., Messaddeq, N., Harney, J.W., Ezaki, O., Kodama, T., Schoonjans, K., Bianco, A.C. and Auwerx, J. (2006) 'Bile acids induce energy expenditure by promoting intracellular thyroid hormone activation', *Nature*, 439(7075), pp. 484–489. Available at: https://doi.org/10.1038/nature04330.

Wattiez, A.-S., Sowers, L.P. and Russo, A.F. (2020) 'Calcitonin gene-related peptide (CGRP): Role in migraine pathophysiology and therapeutic targeting', *Expert opinion on therapeutic targets*, 24(2), pp. 91–100. Available at: https://doi.org/10.1080/14728222.2020.1724285.

Webster, M. and Sheets, H.D. (2010) 'A Practical Introduction to Landmark-Based Geometric Morphometrics', *The Paleontological Society Papers*, 16, pp. 163–188. Available at: https://doi.org/10.1017/S1089332600001868.

Weichseldorfer, M., Tagaya, Y., Reitz, M., DeVico, A.L. and Latinovic, O.S. (2022) 'Identifying CCR5 coreceptor populations permissive for HIV-1 entry and productive infection: implications for in vivo studies', *Journal of Translational Medicine*, 20(1), p. 39. Available at: https://doi.org/10.1186/s12967-022-03243-8.

Weis, W.I. and Kobilka, B.K. (2018) 'The Molecular Basis of G Protein-Coupled Receptor Activation', *Annual Review of Biochemistry*, 87, pp. 897–919. Available at: https://doi.org/10.1146/annurev-biochem-060614-033910.

Weiss, H.M., Haase, W., Michel, H. and Reilander, H. (1998) 'Comparative biochemical and pharmacological characterization of the mouse 5HT5A 5-hydroxytryptamine receptor and the

human β2-adrenergic receptor produced in the methylotrophic yeast Pichia pastoris', *Biochemical Journal*, 330(3), pp. 1137–1147. Available at: https://doi.org/10.1042/bj3301137.

van der Westhuizen, E.T., Breton, B., Christopoulos, A. and Bouvier, M. (2014) 'Quantification of ligand bias for clinically relevant β2-adrenergic receptor ligands: implications for drug taxonomy', *Molecular Pharmacology*, 85(3), pp. 492–509. Available at: https://doi.org/10.1124/mol.113.088880.

Weston, C., Lu, J., Li, N., Barkan, K., Richards, G.O., Roberts, D.J., Skerry, T.M., Poyner, D., Pardamwar, M., Reynolds, C.A., Dowell, S.J., Willars, G.B. and Ladds, G. (2015) 'Modulation of Glucagon Receptor Pharmacology by Receptor Activity-modifying Protein-2 (RAMP2)', *The Journal of Biological Chemistry*, 290(38), pp. 23009–23022. Available at: https://doi.org/10.1074/jbc.M114.624601.

Weston, C., Winfield, I., Harris, M., Hodgson, R., Shah, A., Dowell, S.J., Mobarec, J.C., Woodlock, D.A., Reynolds, C.A., Poyner, D.R., Watkins, H.A. and Ladds, G. (2016) 'Receptor Activity-modifying Protein-directed G Protein Signaling Specificity for the Calcitonin Gene-related Peptide Family of Receptors', *The Journal of Biological Chemistry*, 291(42), pp. 21925–21944. Available at: https://doi.org/10.1074/jbc.M116.751362.

Wheatley, M., Charlton, J., Jamshad, M., Routledge, S.J., Bailey, S., La-Borde, P.J., Azam, M.T., Logan, R.T., Bill, R.M., Dafforn, T.R. and Poyner, D.R. (2016) 'GPCR–styrene maleic acid lipid particles (GPCR–SMALPs): their nature and potential', *Biochemical Society Transactions*, 44(2), pp. 619–623. Available at: https://doi.org/10.1042/BST20150284.

Wheatley, M. and Hawtin, S. (1999) 'Glycosylation of G-protein-coupled receptors for hormones central to normal reproductive functioning: its occurrence and role', *Human Reproduction Update*, 5(4), pp. 356–364. Available at: https://doi.org/10.1093/humupd/5.4.356.

Wheeler-Jones, C.P.D., Farrar, C. and Garonna, E. (2009) 'Protease-activated receptors, cyclo-oxygenases and pro-angiogenic signalling in endothelial cells', *Biochemical Society Transactions*, 37(6), pp. 1179–1183. Available at: https://doi.org/10.1042/BST0371179.

White, A.D., Peña, K.A., Clark, L.J., Maria, C.S., Liu, S., Jean-Alphonse, F.G., Lee, J.Y., Lei, S., Cheng, Z., Tu, C.-L., Fang, F., Szeto, N., Gardella, T.J., Xiao, K., Gellman, S.H., Bahar, I., Sutkeviciute, I., Chang, W. and Vilardaga, J.-P. (2021) 'Spatial bias in cAMP generation determines biological responses to PTH type 1 receptor activation', *Science Signaling*, 14(703), p. eabc5944. Available at: https://doi.org/10.1126/scisignal.abc5944.

White, J.F., Noinaj, N., Shibata, Y., Love, J., Kloss, B., Xu, F., Gvozdenovic-Jeremic, J., Shah, P., Shiloach, J., Tate, C.G. and Grisshammer, R. (2012) 'Structure of the agonist-bound neurotensin receptor', *Nature*, 490(7421), pp. 508–513. Available at: https://doi.org/10.1038/nature11558.

Wickman, K.D., Iñiguez-Lluhi, J.A., Davenport, P.A., Taussig, R., Krapivinsky, G.B., Linder, M.E., Gilman, A.G. and Clapham, D.E. (1994) 'Recombinant G-protein βγ-subunits activate the muscarinic-gated atrial potassium channel', *Nature*, 368(6468), pp. 255–257. Available at: https://doi.org/10.1038/368255a0.

Williamson, I.M., Alvis, S.J., East, J.M. and Lee, A.G. (2003) 'The potassium channel KcsA and its interaction with the lipid bilayer', *Cellular and Molecular Life Sciences CMLS*, 60(8), pp. 1581–1590. Available at: https://doi.org/10.1007/s00018-003-3172-y.

Wingler, L.M. and Lefkowitz, R.J. (2020) 'Conformational basis of G protein-coupled receptor signaling versatility', *Trends in cell biology*, 30(9), pp. 736–747. Available at: https://doi.org/10.1016/j.tcb.2020.06.002.

Wingler, L.M., Skiba, M.A., McMahon, C., Staus, D.P., Kleinhenz, A.L.W., Suomivuori, C.-M., Latorraca, N.R., Dror, R.O., Lefkowitz, R.J. and Kruse, A.C. (2020) 'Angiotensin and biased analogs induce structurally distinct active conformations within a GPCR', *Science (New York, N.Y.)*, 367(6480), pp. 888–892. Available at: https://doi.org/10.1126/science.aay9813.

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. and Goddard, A.D. (2020) 'Expression and purification of recombinant G protein-coupled receptors: A review', *Protein Expression and Purification*, 167, p. 105524. Available at: https://doi.org/10.1016/j.pep.2019.105524.

Wiseman, D.N., Samra, N., Román Lara, M.M., Penrice, S.C. and Goddard, A.D. (2021) 'The Novel Application of Geometric Morphometrics with Principal Component Analysis to Existing G Protein-Coupled Receptor (GPCR) Structures', *Pharmaceuticals*, 14(10), p. 953. Available at: https://doi.org/10.3390/ph14100953.

Wlodawer, A., Minor, W., Dauter, Z. and Jaskolski, M. (2013) 'Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination', *The FEBS journal*, 280(22), pp. 5705–5736. Available at: https://doi.org/10.1111/febs.12495.

Woolley, M.J., Reynolds, C.A., Simms, J., Walker, C.S., Mobarec, J.C., Garelja, M.L., Conner, A.C., Poyner, D.R. and Hay, D.L. (2017) 'Receptor activity-modifying protein dependent and independent activation mechanisms in the coupling of calcitonin gene-related peptide and adrenomedullin receptors to Gs', *Biochemical Pharmacology*, 142, pp. 96–110. Available at: https://doi.org/10.1016/j.bcp.2017.07.005.

Wootten, D., Christopoulos, A., Marti-Solano, M., Babu, M.M. and Sexton, P.M. (2018) 'Mechanisms of signalling and biased agonism in G protein-coupled receptors', *Nature Reviews Molecular Cell Biology*, 19(10), pp. 638–653. Available at: https://doi.org/10.1038/s41580-018-0049-3.

Wootten, D., Lindmark, H., Kadmiel, M., Willcockson, H., Caron, K.M., Barwell, J., Drmota, T. and Poyner, D.R. (2013) 'Receptor activity modifying proteins (RAMPs) interact with the VPAC2 receptor and CRF1 receptors and modulate their function', *British Journal of Pharmacology*, 168(4), pp. 822–834. Available at: https://doi.org/10.1111/j.1476-5381.2012.02202.x.

Wootten, D., Miller, L.J., Koole, C., Christopoulos, A. and Sexton, P.M. (2017) 'Allostery and Biased Agonism at Class B G Protein-Coupled Receptors', *Chemical Reviews*, 117(1), pp. 111–138. Available at: https://doi.org/10.1021/acs.chemrev.6b00049.

Wright, K.L., Duncan, M. and Sharkey, K.A. (2008) 'Cannabinoid CB2 receptors in the gastrointestinal tract: a regulatory system in states of inflammation', *British Journal of Pharmacology*, 153(2), pp. 263–270. Available at: https://doi.org/10.1038/sj.bjp.0707486.

Wright, S.C., Kozielewicz, P., Kowalski-Jahn, M., Petersen, J., Bowin, C.-F., Slodkowicz, G., Marti-Solano, M., Rodríguez, D., Hot, B., Okashah, N., Strakova, K., Valnohova, J., Babu, M.M., Lambert, N.A., Carlsson, J. and Schulte, G. (2019) 'A conserved molecular switch in Class F receptors regulates receptor activation and pathway selection', *Nature Communications*, 10, p. 667. Available at: https://doi.org/10.1038/s41467-019-08630-2.

Wu, M. and Lander, G.C. (2020) 'Present and Emerging Methodologies in Cryo-EM Single-Particle Analysis', *Biophysical Journal*, 119(7), pp. 1281–1289. Available at: https://doi.org/10.1016/j.bpj.2020.08.027.

wwPDB consortium (2019) 'Protein Data Bank: the single global archive for 3D macromolecular structure data', *Nucleic Acids Research*, 47(D1), pp. D520–D528. Available at: https://doi.org/10.1093/nar/gky949.

Xiang, Y. and Kobilka, B.K. (2003) 'Myocyte Adrenoceptor Signaling Pathways', *Science*, 300(5625), pp. 1530–1532. Available at: https://doi.org/10.1126/science.1079206.

Xiong, Y., Guo, J., Candelore, M.R., Liang, R., Miller, C., Dallas-Yang, Q., Jiang, G., McCann, P.E., Qureshi, S.A., Tong, X., Xu, S.S., Shang, J., Vincent, S.H., Tota, L.M., Wright, M.J., Yang, X., Zhang, B.B., Tata, J.R. and Parmee, E.R. (2012) 'Discovery of a Novel Glucagon Receptor Antagonist N-[(4-{(1S)-1-[3-(3, 5-Dichlorophenyl)-5-(6-methoxynaphthalen-2-yl)-1H-pyrazol-1-yl]ethyl}phenyl)carbonyl]-β-alanine (MK-0893) for the Treatment of Type II Diabetes', *Journal of Medicinal Chemistry*, 55(13), pp. 6137–6148. Available at: https://doi.org/10.1021/jm300579z.

Xu, B., Chakraborty, R., Eilers, M., Dakshinamurti, S., O'Neil, J.D., Smith, S.O., Bhullar, R.P. and Chelikani, P. (2013) 'High-Level Expression, Purification and Characterization of a Constitutively Active Thromboxane A2 Receptor Polymorphic Variant', *PLOS ONE*, 8(9), p. e76481. Available at: https://doi.org/10.1371/journal.pone.0076481.

Xu, F., Wu, H., Katritch, V., Han, G.W., Jacobson, K.A., Gao, Z.-G., Cherezov, V. and Stevens, R.C. (2011) 'Structure of an Agonist-Bound Human A2A Adenosine Receptor', *Science*, 332(6027), pp. 322–327. Available at: https://doi.org/10.1126/science.1202793.

Xue, L., Rovira, X., Scholler, P., Zhao, H., Liu, J., Pin, J.-P. and Rondard, P. (2015) 'Major ligand-induced rearrangement of the heptahelical domain interface in a GPCR dimer', *Nature Chemical Biology*, 11(2), pp. 134–140. Available at: https://doi.org/10.1038/nchembio.1711.

Yamashita, S., Lukacik, P., Barnard, T.J., Noinaj, N., Felek, S., Tsang, T.M., Krukonis, E.S., Hinnebusch, B.J. and Buchanan, S.K. (2011) 'Structural insights into Ail-mediated adhesion in Yersinia pestis', *Structure (London, England: 1993)*, 19(11), pp. 1672–1682. Available at: https://doi.org/10.1016/j.str.2011.08.010.

Yang, F., Ling, S., Zhou, Y., Zhang, Y., Lv, P., Liu, S., Fang, W., Sun, W., Hu, L.A., Zhang, L., Shi, P. and Tian, C. (2021) 'Different conformational responses of the β2-adrenergic receptor-Gs complex upon binding of the partial agonist salbutamol or the full agonist isoprenaline', *National Science Review*, 8(9), p. nwaa284. Available at: https://doi.org/10.1093/nsr/nwaa284.

Yang, G., Rosen, D.G., Liu, G., Yang, F., Guo, X., Xiao, X., Xue, F., Mercado-Uribe, I., Huang, J., Lin, S.-H., Mills, G.B. and Liu, J. (2010) 'CXCR2 Promotes Ovarian Cancer Growth through Dysregulated Cell Cycle, Diminished Apoptosis, and Enhanced Angiogenesis', *Clinical Cancer Research*, 16(15), pp. 3875–3886. Available at: https://doi.org/10.1158/1078-0432.CCR-10-0483.

Yang, L., Liu, D. and Wüthrich, K. (2022) 'GPCR structural characterization by NMR spectroscopy in solution', *Acta Biochimica et Biophysica Sinica*, 54(9), pp. 1207–1212. Available at: https://doi.org/10.3724/abbs.2022106.

Yang, N.J. and Hinner, M.J. (2015) 'Getting Across the Cell Membrane: An Overview for Small Molecules, Peptides, and Proteins', in A. Gautier and M.J. Hinner (eds) *Site-Specific Protein*

Labeling. New York, NY: Springer New York (Methods in Molecular Biology), pp. 29–53. Available at: https://doi.org/10.1007/978-1-4939-2272-7_3.

Yao, Y., Fujimoto, L.M., Hirshman, N., Bobkov, A.A., Antignani, A., Youle, R.J. and Marassi, F.M. (2015) 'Conformation of BCL-XL upon membrane-integration', *Journal of molecular biology*, 427(13), pp. 2262–2270. Available at: https://doi.org/10.1016/j.jmb.2015.02.019.

Yarwood, R.E., Imlach, W.L., Lieu, T., Veldhuis, N.A., Jensen, D.D., Klein Herenbrink, C., Aurelio, L., Cai, Z., Christie, M.J., Poole, D.P., Porter, C.J.H., McLean, P., Hicks, G.A., Geppetti, P., Halls, M.L., Canals, M. and Bunnett, N.W. (2017) 'Endosomal signaling of the receptor for calcitonin gene-related peptide mediates pain transmission', *Proceedings of the National Academy of Sciences of the United States of America*, 114(46), pp. 12309–12314. Available at: https://doi.org/10.1073/pnas.1706656114.

Yin, J., Babaoglu, K., Brautigam, C.A., Clark, L., Shao, Z., Scheuermann, T.H., Harrell, C.M., Gotter, A.L., Roecker, A.J., Winrow, C.J., Renger, J.J., Coleman, P.J. and Rosenbaum, D.M. (2016) 'Structure and ligand-binding mechanism of the human OX1 and OX2 orexin receptors', *Nature Structural & Molecular Biology*, 23(4), pp. 293–299. Available at: https://doi.org/10.1038/nsmb.3183.

Yin, J., Mobarec, J.C., Kolb, P. and Rosenbaum, D.M. (2015) 'Crystal structure of the human OX2 orexin receptor bound to the insomnia drug suvorexant', *Nature*, 519(7542), pp. 247–250. Available at: https://doi.org/10.1038/nature14035.

Yin, Y., de Waal, P.W., He, Y., Zhao, L.-H., Yang, D., Cai, X., Jiang, Y., Melcher, K., Wang, M.-W. and Xu, H.E. (2017) 'Rearrangement of a polar core provides a conserved mechanism for constitutive activation of class B G protein-coupled receptors', *The Journal of Biological Chemistry*, 292(24), pp. 9865–9881. Available at: https://doi.org/10.1074/jbc.M117.782987.

Youvan, D.C., Silva, C.M., Bylina, E.J., Coleman, W.J., Dilworth, M.R. and Yang, M.M. (no date) 'Calibration of Fluorescence Resonance Energy Transfer in Microscopy Using Genetically Engineered GFP Derivatives on Nickel Chelating Beads', p. 19.

Yu, J.-L. and Liao, H.-Y. (2021) 'Piezo-type mechanosensitive ion channel component 1 (Piezo1) in human cancer', *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie*, 140, p. 111692. Available at: https://doi.org/10.1016/j.biopha.2021.111692.

Yuzlenko, O. and Kieć-Kononowicz, K. (2009) 'Molecular modeling of A1 and A2A adenosine receptors: comparison of rhodopsin- and beta2-adrenergic-based homology models through the docking studies', *Journal of Computational Chemistry*, 30(1), pp. 14–32. Available at: https://doi.org/10.1002/jcc.21001.

Zelditch, M., Swiderski, D., Sheets, H.D. and Fink, W. (2004) 'Geometric Morphometrics for Biologists: A Primer', *A Primer* [Preprint]. Available at: https://doi.org/10.1016/B978-012778460-1/50003-X.

Zhang, A. (2009) *Protein Interaction Networks: Computational Analysis*. Cambridge University Press.

Zhang, C., Srinivasan, Y., Arlow, D.H., Fung, J.J., Palmer, D., Zheng, Y., Green, H.F., Pandey, A., Dror, R.O., Shaw, D.E., Weis, W.I., Coughlin, S.R. and Kobilka, B.K. (2012) 'High-resolution crystal structure of human protease-activated receptor 1', *Nature*, 492(7429), pp. 387–392. Available at: https://doi.org/10.1038/nature11701.

Zhang, C., Zhang, T., Zou, J., Miller, C.L., Gorkhali, R., Yang, J.-Y., Schilmiller, A., Wang, S., Huang, K., Brown, E.M., Moremen, K.W., Hu, J. and Yang, J.J. (2016) 'Structural basis for regulation of human calcium-sensing receptor by magnesium ions and an unexpected tryptophan derivative co-agonist', *Science Advances*, 2(5), p. e1600241. Available at: https://doi.org/10.1126/sciadv.1600241.

Zhang, H., Unal, H., Gati, C., Han, G.W., Liu, W., Zatsepin, N.A., James, D., Wang, D., Nelson, G., Weierstall, U., Sawaya, M.R., Xu, Q., Messerschmidt, M., Williams, G.J., Boutet, S., Yefanov, O.M., White, T.A., Wang, C., Ishchenko, A., Tirupula, K.C., Desnoyer, R., Coe, J., Conrad, C.E., Fromme, P., Stevens, R.C., Katritch, V., Karnik, S.S. and Cherezov, V. (2015) 'Structure of the Angiotensin receptor revealed by serial femtosecond crystallography', *Cell*, 161(4), pp. 833–844. Available at: https://doi.org/10.1016/j.cell.2015.04.011.

Zhang, R. and Xie, X. (2012) 'Tools for GPCR drug discovery', *Acta Pharmacologica Sinica*, 33(3), pp. 372–384. Available at: https://doi.org/10.1038/aps.2011.173.

Zhang, X., Johnson, R.M., Drulyte, I., Yu, L., Kotecha, A., Danev, R., Wootten, D., Sexton, P.M. and Belousoff, M.J. (2021) 'Evolving cryo-EM structural approaches for GPCR drug discovery', *Structure (London, England: 1993)*, 29(9), pp. 963-974.e6. Available at: https://doi.org/10.1016/j.str.2021.04.008.

Zhang, Ying, Li, Q., Dong, M. and Han, X. (2020) 'Effect of cholesterol on the fluidity of supported lipid bilayers', *Colloids and Surfaces B: Biointerfaces*, 196, p. 111353. Available at: https://doi.org/10.1016/j.colsurfb.2020.111353.

Zhang, Y. and Schepartz, L.A. (2021) 'Three-dimensional geometric morphometric studies of modern human occipital variation', *PLOS ONE*, 16(1), p. e0245445. Available at: https://doi.org/10.1371/journal.pone.0245445.

Zhang, Y., Sun, B., Feng, D., Hu, H., Chu, M., Qu, Q., Tarrasch, J.T., Li, S., Kobilka, T.S., Kobilka, B.K. and Skiniotis, G. (2017) 'Cryo-EM structure of the activated GLP-1 receptor in complex with G protein', *Nature*, 546(7657), pp. 248–253. Available at: https://doi.org/10.1038/nature22394.

Zhang, Yanan, Yang, F., Ling, S., Lv, P., Zhou, Y., Fang, W., Sun, W., Zhang, L., Shi, P. and Tian, C. (2020) 'Single-particle cryo-EM structural studies of the β2AR–Gs complex bound with a full agonist formoterol', *Cell Discovery*, 6(1), pp. 1–5. Available at: https://doi.org/10.1038/s41421-020-0176-9.

Zhang, Z., Kuipers, G., Niemiec, Ł., Baumgarten, T., Slotboom, D.J., de Gier, J.-W. and Hjelm, A. (2015) 'High-level production of membrane proteins in E. coli BL21(DE3) by omitting the inducer IPTG', *Microbial Cell Factories*, 14(1), p. 142. Available at: https://doi.org/10.1186/s12934-015-0328-z.

Zhao, L.-H., Ma, S., Sutkeviciute, I., Shen, D.-D., Zhou, X.E., de Waal, P.W., Li, C.-Y., Kang, Y., Clark, L.J., Jean-Alphonse, F.G., White, A.D., Yang, D., Dai, A., Cai, X., Chen, J., Li, C., Jiang, Y., Watanabe, T., Gardella, T.J., Melcher, K., Wang, M.-W., Vilardaga, J.-P., Xu, H.E. and Zhang, Y. (2019) 'Structure and dynamics of the active human parathyroid hormone receptor-1', *Science (New York, N.Y.)*, 364(6436), pp. 148–153. Available at: https://doi.org/10.1126/science.aav7942.

Zhao, P., Liang, Y.-L., Belousoff, M.J., Deganutti, G., Fletcher, M.M., Willard, F.S., Bell, M.G., Christe, M.E., Sloop, K.W., Inoue, A., Truong, T.T., Clydesdale, L., Furness, S.G.B., Christopoulos, A., Wang, M.-W., Miller, L.J., Reynolds, C.A., Danev, R., Sexton, P.M. and

Wootten, D. (2020) ‘Activation of the GLP-1 receptor by a non-peptidic agonist’, *Nature*, 577(7790), pp. 432–436. Available at: https://doi.org/10.1038/s41586-019-1902-z.

Zhou, H.-X. and Cross, T.A. (2013) ‘Influences of Membrane Mimetic Environments on Membrane Protein Structures’, *Annual review of biophysics*, 42, pp. 361–392. Available at: https://doi.org/10.1146/annurev-biophys-083012-130326.

Zhou, Q., Yang, D., Wu, M., Guo, Y., Guo, W., Zhong, L., Cai, X., Dai, A., Jang, W., Shakhnovich, E.I., Liu, Z.-J., Stevens, R.C., Lambert, N.A., Babu, M.M., Wang, M.-W. and Zhao, S. (2019) ‘Common activation mechanism of class A GPCRs’, *eLife*. Edited by Y. Shan and O. Boudker, 8, p. e50279. Available at: https://doi.org/10.7554/eLife.50279.

Zhou, W. (2012) ‘The new face of anaphylatoxins in immune regulation’, *Immunobiology*, 217(2), pp. 225–234. Available at: https://doi.org/10.1016/j.imbio.2011.07.016.

Zhou, X.E., Melcher, K. and Xu, H.E. (2012) ‘Structure and activation of rhodopsin’, *Acta Pharmacologica Sinica*, 33(3), pp. 291–299. Available at: https://doi.org/10.1038/aps.2011.171.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E. and Scheres, S.H. (2018) ‘New tools for automated high-resolution cryo-EM structure determination in RELION-3’, *eLife*. Edited by E.H. Egelman and J. Kuriyan, 7, p. e42166. Available at: https://doi.org/10.7554/eLife.42166.

Zoonens, M. and Popot, J.-L. (2014) ‘Amphipols for each season’, *The Journal of Membrane Biology*, 247(9–10), pp. 759–796. Available at: https://doi.org/10.1007/s00232-014-9666-8.

Zou, Y., Weis, W.I. and Kobilka, B.K. (2012) ‘N-terminal T4 lysozyme fusion facilitates crystallization of a G protein coupled receptor’, *PloS One*, 7(10), p. e46039. Available at: https://doi.org/10.1371/journal.pone.0046039.

Zucchelli, S., Patrucco, L., Persichetti, F., Gustincich, S. and Cotella, D. (2016) ‘Engineering Translation in Mammalian Cell Factories to Increase Protein Yield: The Unexpected Use of Long Non-Coding SINEUP RNAs’, *Computational and Structural Biotechnology Journal*, 14, pp. 404–410. Available at: https://doi.org/10.1016/j.csbj.2016.10.004.

Zuo, J., Currin, A., Griffin, B.D., Shannon-Lowe, C., Thomas, W.A., Ressing, M.E., Wiertz, E.J.H.J. and Rowe, M. (2009) ‘The Epstein-Barr Virus G-Protein-Coupled Receptor Contributes to Immune Evasion by Targeting MHC Class I Molecules for Degradation’, *PLOS Pathogens*, 5(1), p. e1000255. Available at: https://doi.org/10.1371/journal.ppat.1000255.

Appendices

Appendix A – Comparisons Between Designed Constructs and Sanger Sequencing

Appendix A1: Sequencing Data for CTR-YFP

A comparison of the amino acid sequences of the designed CTR-YFP construct in pcDNA3.1- (Query) and Sanger sequencing data (Sbjct). Dots indicate identical residues, red indicates a mismatch or blank, and sequencing was 97% identical.

Score	Expect	Method	Identities	Positives	Gaps
1496 bits(3873)	0.0	Compositional matrix adjust.	719/743(97%)	720/743(96%)	16/743(2%)

```
Query  7    VTALLLPLALLLHAARPDYASYPYDVPDYAEFSNQTYPTIEPKPFLYVVGRKKMMDAQYK  66
Sbjct  4    .X.................................N.........................  63

Query  67   CYDRMQQLPAYQGEGPYCNRTWDGWLCWDDTPAGVLSYQFCPDYFPDFDPSEKVTKYCDE  126
Sbjct  64   ............................................................  123

Query  127  KGVWFKHPENNRTWSNYTMCNAFTPEKLKNAYVLYYLAIVGHSLSIFTLVISLGIFVFFR  186
Sbjct  124  ............................................................  183

Query  187  KLTTIFPLNWKYRKALSLGCQRVTLHKNMFLTYILNSMIIIIHLVEVVPNGELVRRDPVS  246
Sbjct  184  ----------------............................................  227

Query  247  CKILHFFHQYMMACNYFWMLCEGIYLHTLIVVAVFTEKQRLRWYYLLGWGFPLVPTTIHA  306
Sbjct  228  ............................................................  287

Query  307  ITRAVYFNDNCWLSVETHLLYIIHGPVMAALVVNFFFLLNIVRVLVTKMRETHEAESHMY  366
Sbjct  288  ....M.........................................................  347

Query  367  LKAVKATMILVPLLGIQFVVFPWRPSNKMLGKIYDYVMHSLIHFQGFFVATIYCFCNNEV  426
Sbjct  348  ............................................................  407

Query  427  QTTVKRQWAQFKIQWNQRWGRRPSNRSARAAAAAAAEAGDIPIYICHQEPRNEPANNQGEE  486
Sbjct  408  .....XX.............X.......................L...........  467

Query  487  SAEIIPLNIIEQESSAGGSGGSGSVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGD  546
Sbjct  468  ............................................................  527

Query  547  ATYGKLTLKFICTTGKLPVPWPTLVTTFGYGLQCFARYPDHMKQHDFFKSAMPEGYVQER  606
Sbjct  528  ............................................................  587

Query  607  TIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADK  666
Sbjct  588  ............................................................  647

Query  667  QKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSYQSALSKDPNEKRD  726
Sbjct  648  ............................................................  707

Query  727  HMVLLEFVTAAGITLGMDELYKK  749
Sbjct  708  .....................X.  730
```

Appendix A2: Sequencing Data for RAMP1-CFP

A comparison of the amino acid sequences of the designed RAMP1-CFP construct in pcDNA3.1$^-$ (Query) and Sanger sequencing data (Sbjct). Dots indicate identical residues, red indicates a mismatch or blank, and sequencing was 99% identical.

Score	Expect	Method	Identities	Positives	Gaps
783 bits(2022)	0.0	Compositional matrix adjust.	375/376(99%)	375/376(99%)	0/376(0%)

```
Query  1    EFMDYKDDDDKTTACQEANYGALLRELCLTQFQVDMEAVGETLWCDWGRTIRSYRELADC  60
Sbjct  11   ............................................................  70

Query  61   TWHMAEKLGCFWPNAEVDRFFLAVHGRYFRSCPISGRAVRDPPGSILYPFIVVPITVTLL  120
Sbjct  71   ............................................................  130

Query  121  VTALVVWQSKRTEGIVGGSGGSGSVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGD  180
Sbjct  131  ............................................................  190

Query  181  ATYGKLTLKFICTTGKLPVPWPTLVTTLTWGVQCFSRYPDHMKQHDFFKSAMPEGYVQER  240
Sbjct  191  ............................................................  250

Query  241  TIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYISHNVYITADK  300
Sbjct  251  ............................................................  310

Query  301  QKNGIKANFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRD  360
Sbjct  311  ............................................................  370

Query  361  HMVLLEFVTAAGITLG  376
Sbjct  371  ...........X....  386
```

Appendix A3: Sequencing Data for RAMP2-CFP

A comparison of the amino acid sequences of the designed RAMP2-CFP construct in pcDNA3.1$^-$ (Query) and Sanger sequencing data (Sbjct). Dots indicate identical residues, red indicates a mismatch or blank, and sequencing was 100% identical.

Score	Expect	Method	Identities	Positives	Gaps
819 bits(2116)	0.0	Compositional matrix adjust.	393/393(100%)	393/393(100%)	0/393(0%)

```
Query  1    EFMDYKDDDDKTQPLPTTGTPGSEGGTVKNYETAVQFCWNHYKDQMDPIEKDWCDWAMIS  60
Sbjct  1    ............................................................  60

Query  61   RPYSTLRDCLEHFAELFDLGFPNPLAERIIFETHQIHFANCSLVQPTFSDPPEDVLLAMI  120
Sbjct  61   ............................................................  120

Query  121  IAPICLIPFLITLVVWRSKDSEAQAGGSGGSGSVSKGEELFTGVVPILVELDGDVNGHKF  180
Sbjct  121  ............................................................  180

Query  181  SVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTWGVQCFSRYPDHMKQHDFFKSA  240
Sbjct  181  ............................................................  240

Query  241  MPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYIS  300
Sbjct  241  ............................................................  300

Query  301  HNVYITADKQKNGIKANFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSAL  360
Sbjct  301  ............................................................  360

Query  361  SKDPNEKRDHMVLLEFVTAAGITLGMDELYKKL  393
Sbjct  361  .................................  393
```

Appendix A4: Sequencing Data for RAMP3-CFP

A comparison of the amino acid sequences of the designed RAMP3-CFP construct in pcDNA3.1^{-} (Query) and Sanger sequencing data (Sbjct). Dots indicate identical residues, red indicates a mismatch or blank, and sequencing was 99% identical.

Score	Expect	Method	Identities	Positives	Gaps
754 bits(1947)	0.0	Compositional matrix adjust.	362/364(99%)	362/364(99%)	0/364(0%)

```
Query  1    EFMDYKDDDDKTGGCNETGMLERLPLCGKAFADMMGKVDVWKWCNLSEFIVYYESFTNCT  60
Sbjct  8    ............................................................  67

Query  61   EMEANVVGCYWPNPLAQGFITGIHRQFFSNCTVDRVHLEDPPDEVLIPLIVIPVVLTVAM  120
Sbjct  68   ............................................................  127

Query  121  AGLVVWRSKRTDTLLGGSGGSGSVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDA  180
Sbjct  128  ............................................................  187

Query  181  TYGKLTLKFICTTGKLPVPWPTLVTTLTWGVQCFSRYPDHMKQHDFFKSAMPEGYVQERT  240
Sbjct  188  ............................................................  247

Query  241  IFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYISHNVYITADKQ  300
Sbjct  248  ............................................................  307

Query  301  KNGIKANFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDH  360
Sbjct  308  ................................X.............X..........  367

Query  361  MVLL  364
Sbjct  368  ....  371
```

Appendix B – Lists of Structures Included in Geometric Morphometric Analyses

Appendix B details a list of all structures used in the geometric morphometric analyses of Chapters 4 and 5, including their PDB code, method of structure determination, resolution in Angstroms, activation state, class of bound ligand, and reference to their publication. All of these details were obtained from the GPCRdb.

Appendix B1: The Family A β2-Adrenergic Receptors Analysed in Section 4.3.1

5JQH was removed from analyses due to its extreme skewing of data.

PDB	Method	Resolution (Å)	State	Ligand	Reference
2R4R	X-ray	3.4	Inactive	Inverse agonist	-
2R4S	X-ray	3.4	Inactive	Inverse agonist	10.1038/NATURE06325
2RH1	X-ray	2.4	Inactive	Inverse agonist	10.1126/SCIENCE.1150577
3D4S	X-ray	2.8	Inactive	Inverse agonist	10.1016/J.STR.2008.05.001
3KJ6	X-ray	3.4	Inactive	Inverse agonist	10.1038/NATURE08650
3NY8	X-ray	2.8	Inactive	Inverse agonist	-
3NY9	X-ray	2.8	Inactive	Inverse agonist	10.1021/JA105108Q
3NYA	X-ray	3.2	Inactive	Antagonist	-
3P0G	X-ray	3.5	Active	Agonist	10.1038/NATURE09648
3PDS	X-ray	3.5	Inactive	Agonist	10.1038/NATURE09665
3SN6	X-ray	3.2	Active	Agonist	10.1038/NATURE10361
4GBR	X-ray	4	Inactive	Inverse agonist	10.1371/JOURNAL.PONE.0046039
4LDE	X-ray	2.8	Active	Agonist	10.1038/NATURE12572
4LDL	X-ray	3.1	Active	Agonist	10.1038/NATURE12572
4LDO	X-ray	3.2	Active	Agonist	10.1038/NATURE12572
4QKX	X-ray	3.3	Active	Agonist	10.1073/PNAS.1410415111
5D5A	X-ray	2.5	Inactive	Inverse agonist	10.1107/S2059798315021683
5D5B	X-ray	3.8	Inactive	Inverse agonist	10.1107/S2059798315021683
5D6L	X-ray	3.2	Inactive	Inverse agonist	10.1038/NPROT.2017.057
5JQH	X-ray	3.2	Inactive	Inverse agonist	10.1038/NATURE18636
5X7D	X-ray	2.7	Inactive	Antagonist	10.1038/NATURE23652
6E67	X-ray	3.7	Active	Agonist	10.1016/J.CELL.2019.04.021
6MXT	X-ray	3	Active	Agonist	10.1038/s41589-018-0145-x
6N48	X-ray	3.2	Active	Agonist PAM	10.1126/SCIENCE.AAW8981
6NI3	Cryo-EM	3.8	Active	Agonist	10.1038/S41594-019-0330-Y

6OBA	X-ray	3.1	Inactive	Antagonist NAM	10.1038/S41589-020-0549-2
6PRZ	X-ray	2.8	Inactive	Antagonist	10.1107/S2052252519013137
6PS0	X-ray	3.4	Inactive	Antagonist	10.1107/S2052252519013137
6PS1	X-ray	3.2	Inactive	Antagonist	10.1107/S2052252519013137
6PS2	X-ray	2.4	Inactive	Antagonist	10.1107/S2052252519013137
6PS3	X-ray	2.5	Inactive	Antagonist	10.1107/S2052252519013137
6PS4	X-ray	2.6	Inactive	Antagonist	10.1107/S2052252519013137
6PS5	X-ray	2.9	Inactive	Antagonist	10.1107/S2052252519013137
6PS6	X-ray	2.7	Inactive	Antagonist	10.1107/S2052252519013137
7BZ2	Cryo-EM	3.8	Active	Agonist	10.1038/S41421-020-0176-9
7DHI	Cryo-EM	3.3	Active	Agonist	10.1093/NSR/NWAA284
7DHR	Cryo-EM	3.8	Active	Agonist	10.1093/NSR/NWAA284

Appendix B2: The Family A Adenosine A2A Receptors Analysed in Section 4.3.2

PDB	Method	Resolution (Å)	State	Ligand	Reference
2YDO	X-ray	3	Active	Agonist	10.1038/NATURE10136
2YDV	X-ray	2.6	Active	Agonist	-
3EML	X-ray	2.6	Inactive	Antagonist	10.1126/SCIENCE.1164772
3PWH	X-ray	3.3	Inactive	Antagonist	10.1016/J.STR.2011.06.014
3QAK	X-ray	2.7	Active	Agonist	10.1126/SCIENCE.1202793
3REY	X-ray	3.3	Inactive	Antagonist	-
3RFM	X-ray	3.6	Inactive	Antagonist	-
3UZA	X-ray	3.3	Inactive	Antagonist	10.1021/jm201376w
3UZC	X-ray	3.3	Inactive	Antagonist	10.1021/jm201376w
3VG9	X-ray	2.7	Inactive	Antagonist	10.1038/NATURE10750
3VGA	X-ray	3.1	Inactive	Antagonist	10.1038/NATURE10750
4EIY	X-ray	1.8	Inactive	Antagonist	10.1126/SCIENCE.1219218
4UG2	X-ray	2.6	Active	Agonist	10.1124/mol.114.097360
4UHR	X-ray	2.6	Active	Agonist	10.1124/mol.114.097360
5G53	X-ray	3.4	Active	Agonist	10.1038/NATURE18966
5IU4	X-ray	1.7	Inactive	Antagonist	10.1021/acs.jmedchem.6b00653
5IU7	X-ray	1.9	Inactive	Antagonist	10.1021/acs.jmedchem.6b00653
5IU8	X-ray	2	Inactive	Antagonist	10.1021/acs.jmedchem.6b00653
5IUA	X-ray	2.2	Inactive	Antagonist	10.1021/acs.jmedchem.6b00653
5IUB	X-ray	2.1	Inactive	Antagonist	10.1021/acs.jmedchem.6b00653
5JTB	X-ray	2.8	Inactive	Antagonist	10.1126/SCIADV.1602952
5K2A	X-ray	2.5	Inactive	Antagonist	10.1126/SCIADV.1600292
5K2B	X-ray	2.5	Inactive	Antagonist	10.1126/SCIADV.1600292
5K2C	X-ray	1.9	Inactive	Antagonist	10.1126/SCIADV.1600292
5K2D	X-ray	1.9	Inactive	Antagonist	10.1126/SCIADV.1600292
5MZJ	X-ray	2	Inactive	Antagonist	10.1016/j.str.2017.06.012

5MZP	X-ray	2.1	Inactive	Antagonist	10.1016/j.str.2017.06.012
5N2R	X-ray	2.8	Inactive	Antagonist	10.1016/j.str.2017.06.012
5NLX	X-ray	2.1	Inactive	Antagonist	-
5NM2	X-ray	2	Inactive	Antagonist	10.1038/S41467-017-00630-4
5NM4	X-ray	1.7	Inactive	Antagonist	-
5OLG	X-ray	1.9	Inactive	Antagonist	10.1038/S41598-017-18570-W
5OLH	X-ray	2.6	Inactive	Antagonist	-
5OLO	X-ray	3.1	Inactive	Antagonist	10.1038/S41598-017-18570-W
5OLV	X-ray	2	Inactive	Antagonist	10.1038/S41598-017-18570-W
5OLZ	X-ray	1.9	Inactive	Antagonist	10.1038/S41598-017-18570-W
5OM1	X-ray	2.1	Inactive	Antagonist	10.1038/S41598-017-18570-W
5OM4	X-ray	2	Inactive	Antagonist	10.1038/S41598-017-18570-W
5UIG	X-ray	3.5	Inactive	Antagonist	10.1073/PNAS.1621423114
5UVI	X-ray	3.2	Inactive	Antagonist	10.1107/S205225251700570X
5VRA	X-ray	2.4	Inactive	Antagonist	10.1038/NPROT.2017.135
5WF5	X-ray	2.6	Active	Agonist	10.1016/J.STR.2017.12.013
5WF6	X-ray	2.9	Active	Agonist	10.1016/J.STR.2017.12.013
6AQF	X-ray	2.5	Inactive	Antagonist	10.1016/J.CELL.2017.12.004
6GDG	Cryo-EM	4.1	Active	Agonist	10.7554/ELIFE.35946
6GT3	X-ray	2	Inactive	Antagonist	-
6JZH	X-ray	2.3	Inactive	Antagonist	10.1107/S1600576719012846
6LPJ	X-ray	1.8	Inactive	Antagonist	10.1038/S41598-020-76277-X
6LPK	X-ray	1.8	Inactive	Antagonist	10.1038/S41598-020-76277-X
6LPL	X-ray	2	Inactive	Antagonist	10.1038/S41598-020-76277-X
6MH8	X-ray	4.2	Inactive	Antagonist	10.1107/S205225251900263X
6PS7	X-ray	1.9	Inactive	Antagonist	10.1107/S2052252519013137
6S0L	X-ray	2.7	Inactive	Antagonist	10.1107/S2052252520011379
6S0Q	X-ray	2.7	Inactive	Antagonist	10.1107/S2052252520011379
6WQA	X-ray	2	Inactive	Antagonist	10.1107/S2052252520012701
6ZDR	X-ray	1.9	Inactive	Antagonist	10.1002/ANIE.202003788
6ZDV	X-ray	2.1	Inactive	Antagonist	10.1002/ANIE.202003788
7ARO	X-ray	3.1	Inactive	Agonist (partial)	10.1021/acs.jmedchem.0c01856

Appendix B3: The Family B1 Secretin-Like Receptors Analysed in Section 4.3.3

PDB	Method	Resolution (Å)	State	Ligand	Reference
4K5Y	X-ray	3	Inactive	Antagonist	10.1038/NATURE12357
4L6R	X-ray	3.3	Inactive	Agonist	10.1038/nature12393
4Z9G	X-ray	3.2	Inactive	Antagonist	10.2174/1874467210666170110114727
5EE7	X-ray	2.5	Inactive	Antagonist NAM	10.1038/NATURE17414
5UZ7	Cryo-EM	4.1	Active	Agonist	10.1038/NATURE22327

5VEW	X-ray	2.7	Inactive	NAM	10.1038/NATURE22378
5VEX	X-ray	3	Inactive	NAM	10.1038/NATURE22378
5XEZ	X-ray	3	Inactive	NAM	10.1038/NATURE22363
5XF1	X-ray	3.2	Inactive	NAM	10.1038/NATURE22363
5YQZ	X-ray	3	Inactive	Agonist (partial)	10.1038/NATURE25153
6B3J	Cryo-EM	3.3	Active	Agonist	10.1038/NATURE25773
6E3Y	Cryo-EM	3.3	Active	Agonist	10.1038/S41586-018-0535-Y
6FJ3	X-ray	2.5	Inactive	Agonist (partial)	10.1038/S41594-018-0151-4
6KJV	X-ray	2.8	Inactive	NAM	10.1107/S2052252519013496
6KK1	X-ray	2.8	Inactive	NAM	10.1107/S2052252519013496
6KK7	X-ray	3.1	Inactive	NAM	10.1107/S2052252519013496
6LMK	Cryo-EM	3.7	Active	Agonist	10.1126/SCIENCE.AAZ5346
6LML	Cryo-EM	3.9	Active	Agonist	10.1126/SCIENCE.AAZ5346
6LN2	X-ray	3.2	Inactive	NAM	10.1038/S41467-020-14934-5
6LPB	Cryo-EM	3.9	Active	Agonist	10.1038/S41594-020-0386-8
6M1H	Cryo-EM	3.6	Active	Agonist	10.1038/S41422-020-0280-2
6M1I	Cryo-EM	3.5	Active	Agonist	10.1038/S41422-020-0280-2
6NBF	Cryo-EM	3	Active	Agonist	10.1126/SCIENCE.AAV7942
6NBH	Cryo-EM	3.5	Active	Agonist	10.1126/SCIENCE.AAV7942
6NBI	Cryo-EM	4	Active	Agonist	10.1126/SCIENCE.AAV7942
6NIY	Cryo-EM	3.3	Active	Agonist	10.1021/ACSPTSCI.8B00056
6ORV	Cryo-EM	3	Active	Agonist	10.1038/s41586-019-1902-z
6P9X	Cryo-EM	2.9	Active	Agonist	10.1016/J.MOLCEL.2020.01.012
6P9Y	Cryo-EM	3	Active	Agonist	10.1016/J.MOLCEL.2020.01.012
6PB0	Cryo-EM	3	Active	Agonist	10.1016/J.MOLCEL.2020.01.013
6PB1	Cryo-EM	2.8	Active	Agonist	10.1016/J.MOLCEL.2020.01.013
6UUN	Cryo-EM	3	Active	Agonist	10.1021/ACSPTSCI.9B00080
6UUS	Cryo-EM	2.4	Active	Agonist	10.1021/ACSPTSCI.9B00080
6UVA	Cryo-EM	2.3	Active	Agonist	10.1021/ACSPTSCI.9B00080
6VCB	Cryo-EM	3.3	Active	Agonist PAM	10.1038/S41589-020-0589-7
6VN7	Cryo-EM	3.2	Active	Agonist	10.1038/S41467-020-17933-8
6WHC	Cryo-EM	3.4	Active	Agonist	10.1074/JBC.RA120.013793

6WI9	Cryo-EM	4.3	Active	Agonist	10.1038/S41467-020-17791-4
6WPW	Cryo-EM	3.1	Active	Agonist	10.1126/SCIENCE.ABA3373
6WZG	Cryo-EM	2.3	Active	Agonist	10.1038/S41467-020-17791-4
6X18	Cryo-EM	2.1	Active	Agonist	10.1016/J.MOLCEL.2020.09.020
6X19	Cryo-EM	2.1	Active	Agonist	10.1016/J.MOLCEL.2020.09.020
6X1A	Cryo-EM	2.5	Active	Agonist	10.1016/J.MOLCEL.2020.09.020
6XOX	Cryo-EM	3.1	Active	Agonist	10.1073/PNAS.2014879117
7C2E	Cryo-EM	4.2	Active	Agonist	10.1038/S41422-020-0384-8
7CZ5	Cryo-EM	2.6	Active	Agonist	10.1038/S41467-020-18945-0
7D3S	Cryo-EM	2.9	Active	Agonist	10.1016/J.BBRC.2020.08.042
7D68	Cryo-EM	3	Active	Agonist	10.1038/S41422-020-00442-0
7DTY	Cryo-EM	3	Active	Agonist	10.7554/ELIFE.68719
7DUQ	Cryo-EM	2.5	Active	Ago-PAM Agonist	10.1038/S41467-021-24058-Z
7DUR	Cryo-EM	3.3	Active	Ago-PAM	10.1038/S41467-021-24058-Z
7E14	Cryo-EM	2.9	Active	Ago-PAM Agonist	10.1038/S41467-021-24058-Z
7EVM	Cryo-EM	2.5	Active	Ago-PAM	10.1038/S41467-021-24058-Z
7F16	Cryo-EM	2.8	Active	Agonist	10.1073/PNAS.2101279118
7KI0	Cryo-EM	2.5	Active	Agonist	10.1016/J.CELREP.2021.109374
7KI1	Cryo-EM	2.5	Active	Agonist	-
7KNT	Cryo-EM	3.2	Inactive	Apo	10.1126/SCIENCE.ABF7258
7KNU	Cryo-EM	3.5	Inactive	Agonist	10.1126/SCIENCE.ABF7258
7LCI	Cryo-EM	2.9	Active	Agonist	10.1016/J.STR.2021.04.008
7LCJ	Cryo-EM	2.8	Active	Agonist	10.1016/J.STR.2021.04.008
7LCK	Cryo-EM	3.2	Active	Agonist	10.1016/J.STR.2021.04.008

Appendix B4: The Family C Calcium-Sensing Receptors Analysed in Section 4.3.4

PDB	Method	Resolution (Å)	State	Ligand	Reference
7DD5	Cryo-EM	3.2	Inactive	Agonist NAM PAM	10.1126/SCIADV.ABG1483
7DD6	Cryo-EM	3.2	Active	Agonist PAM	10.1126/SCIADV.ABG1483
7DD7	Cryo-EM	3.2	Active	Agonist PAM PAM	10.1126/SCIADV.ABG1483
7DTT	Cryo-EM	3.8	Active	Agonist	10.1038/S41422-021-00474-0
7DTU	Cryo-EM	4.4	Intermediate	PAM	10.1038/S41422-021-00474-0
7DTV	Cryo-EM	3.5	Active	Agonist PAM	10.1038/S41422-021-00474-0
7DTW	Cryo-EM	4.5	Inactive	Apo	10.1038/S41422-021-00474-0
7M3E	Cryo-EM	3.2	Inactive	Agonist NAM PAM	10.1038/S41586-021-03691-0
7M3F	Cryo-EM	2.8	Active	Agonist PAM PAM	10.1038/S41586-021-03691-0
7M3G	Cryo-EM	2.5	Active	Agonist PAM PAM	10.1038/S41586-021-03691-0
7M3J	Cryo-EM	4.1	Inactive	NAM	10.1038/S41586-021-03691-0

Appendix B5: The G Protein-Coupled Structures Analysed in Section 5.2

PDB	Method	Resolution (Å)	State	Ligand	Reference
1F88	X-ray	2.8	Inactive	Inverse agonist	10.1126/SCIENCE.289.5480.739
1GZM	X-ray	2.7	Inactive	Inverse agonist	10.1016/J.JMB.2004.08.090
1HZX	X-ray	2.8	Inactive	Inverse agonist	10.1021/BI0155091
1L9H	X-ray	2.6	Inactive	Inverse agonist	10.1073/PNAS.082666399
1U19	X-ray	2.2	Inactive	Inverse agonist	10.1016/J.JMB.2004.07.044
2G87	X-ray	2.6	Inactive	Agonist	10.1002/ANIE.200600595
2HPY	X-ray	2.8	Inactive	Agonist	10.1073/PNAS.0601765103
2I35	X-ray	3.8	Inactive	Inverse agonist	10.1073/PNAS.0608022103
2I36	X-ray	4.1	Inactive	Apo	10.1073/PNAS.0608022103
2I37	X-ray	4.2	Inactive	Apo	10.1073/PNAS.0608022103
2J4Y	X-ray	3.4	Inactive	Inverse agonist	10.1016/J.JMB.2007.03.007
2PED	X-ray	3	Inactive	Inverse agonist	10.1529/BIOPHYSJ.107.108225
2R4R	X-ray	3.4	Inactive	-	10.1038/NATURE06325
2R4S	X-ray	3.4	Inactive	-	-

2RH1	X-ray	2.4	Inactive	Inverse agonist	10.1126/SCIENCE.1150577
2VT4	X-ray	2.7	Inactive	Antagonist	10.1038/NATURE07101
2Y00	X-ray	2.5	Inactive	Agonist (partial)	10.1038/NATURE09746
2Y01	X-ray	2.6	Inactive	Agonist (partial)	10.1038/NATURE09746
2Y02	X-ray	2.6	Inactive	Agonist	10.1038/NATURE09746
2Y03	X-ray	2.9	Inactive	Agonist	10.1038/NATURE09746
2Y04	X-ray	3.1	Inactive	Agonist (partial)	10.1038/NATURE09746
2YCW	X-ray	3	Inactive	Antagonist	10.1073/PNAS.1100185108
2YCX	X-ray	3.3	Inactive	Antagonist	10.1073/PNAS.1100185108
2YCY	X-ray	3.2	Inactive	Antagonist	10.1073/PNAS.1100185108
2YCZ	X-ray	3.7	Inactive	Antagonist	10.1073/PNAS.1100185108
2Z73	X-ray	2.5	Inactive	Inverse agonist	10.1038/NATURE06925
2ZIY	X-ray	3.7	Inactive	Inverse agonist	10.1074/JBC.C800040200
3AYM	X-ray	2.8	Inactive	Agonist	10.1016/J.JMB.2011.08.044
3AYN	X-ray	2.7	Inactive	Inverse agonist	10.1016/J.JMB.2011.08.044
3C9L	X-ray	2.7	Inactive	Inverse agonist	10.1107/S0907444908017162
3C9M	X-ray	3.4	Inactive	Inverse agonist	10.1107/S0907444908017162
3D4S	X-ray	2.8	Inactive	Inverse agonist	10.1016/J.STR.2008.05.001
3EML	X-ray	2.6	Inactive	Antagonist	10.1126/SCIENCE.1164772
3KJ6	X-ray	3.4	Inactive	-	10.1038/NATURE08650
3NY8	X-ray	2.8	Inactive	Inverse agonist	10.1021/JA105108Q
3NY9	X-ray	2.8	Inactive	Inverse agonist	10.1021/JA105108Q
3NYA	X-ray	3.2	Inactive	Antagonist	10.1021/JA105108Q
3OAX	X-ray	2.6	Inactive	Inverse agonist	10.1016/J.BPJ.2010.08.003
3ODU	X-ray	2.5	Inactive	Antagonist	10.1126/SCIENCE.1194396
3OE0	X-ray	2.9	Inactive	Antagonist	10.1126/SCIENCE.1194396
3OE6	X-ray	3.2	Inactive	Antagonist	10.1126/SCIENCE.1194396
3OE8	X-ray	3.1	Inactive	Antagonist	10.1126/SCIENCE.1194396
3OE9	X-ray	3.1	Inactive	Antagonist	10.1126/SCIENCE.1194396
3PBL	X-ray	2.9	Inactive	Antagonist	10.1126/SCIENCE.1197410
3PDS	X-ray	3.5	Inactive	Agonist	10.1038/NATURE09665
3PWH	X-ray	3.3	Inactive	Antagonist	10.1016/J.STR.2011.06.014
3REY	X-ray	3.3	Inactive	Antagonist	10.1016/J.STR.2011.06.014
3RFM	X-ray	3.6	Inactive	Antagonist	10.1016/J.STR.2011.06.014
3RZE	X-ray	3.1	Inactive	Antagonist	10.1038/NATURE10236
3UON	X-ray	3	Inactive	Antagonist	10.1038/NATURE10753
3UZA	X-ray	3.3	Inactive	Antagonist	10.1021/jm201376w
3UZC	X-ray	3.3	Inactive	Antagonist	10.1021/jm201376w
3VG9	X-ray	2.7	Inactive	Antagonist	10.1038/NATURE10750

3VGA	X-ray	3.1	Inactive	Antagonist	10.1038/NATURE10750
3ZPQ	X-ray	2.8	Inactive	Antagonist	10.1021/jm400140q
3ZPR	X-ray	2.7	Inactive	Antagonist	10.1021/jm400140q
4AMI	X-ray	3.2	Inactive	Agonist	10.1016/J.STR.2012.03.014
4AMJ	X-ray	2.3	Inactive	Inverse agonist	10.1016/J.STR.2012.03.014
4BVN	X-ray	2.1	Inactive	Antagonist	10.1371/JOURNAL.PONE.0092727
4DAJ	X-ray	3.4	Inactive	Antagonist	10.1038/NATURE10867
4EIY	X-ray	1.8	Inactive	Antagonist	10.1126/SCIENCE.1219218
4GBR	X-ray	4	Inactive	Inverse agonist	10.1371/JOURNAL.PONE.0046039
4GPO	X-ray	3.5	Inactive	Apo	10.1038/NSMB.2504
4RWS	X-ray	3.1	Inactive	Antagonist	10.1126/SCIENCE.1261064
4U14	X-ray	3.6	Inactive	Antagonist	10.1016/J.STR.2014.08.022
4U15	X-ray	2.8	Inactive	Antagonist	10.1016/J.STR.2014.08.022
4U16	X-ray	3.7	Inactive	Antagonist	10.1016/J.STR.2014.08.022
4WW3	X-ray	2.8	Inactive	Agonist	10.1371/JOURNAL.PONE.0126970
5A8E	X-ray	2.4	Inactive	Inverse agonist	10.1124/MOL.115.101030
5D5A	X-ray	2.5	Inactive	Inverse agonist	10.1107/S2059798315021683
5D5B	X-ray	3.8	Inactive	Inverse agonist	10.1107/S2059798315021683
5D6L	X-ray	3.2	Inactive	Inverse agonist	10.1038/NPROT.2017.057
5F8U	X-ray	3.4	Inactive	Antagonist	10.1038/NSMB.3130
5IU4	X-ray	1.7	Inactive	Antagonist	10.1021/acs.jmedchem.6b00653
5IU7	X-ray	1.9	Inactive	Antagonist	10.1021/acs.jmedchem.6b00653
5IU8	X-ray	2	Inactive	Antagonist	10.1021/acs.jmedchem.6b00653
5IUA	X-ray	2.2	Inactive	Antagonist	10.1021/acs.jmedchem.6b00653
5IUB	X-ray	2.1	Inactive	Antagonist	10.1021/acs.jmedchem.6b00653
5JQH	X-ray	3.2	Inactive	Inverse agonist	10.1038/NATURE18636
5JTB	X-ray	2.8	Inactive	Antagonist	10.1126/SCIADV.1602952
5K2A	X-ray	2.5	Inactive	Antagonist	10.1126/SCIADV.1600292
5K2B	X-ray	2.5	Inactive	Antagonist	10.1126/SCIADV.1600292
5K2C	X-ray	1.9	Inactive	Antagonist	10.1126/SCIADV.1600292
5K2D	X-ray	1.9	Inactive	Antagonist	10.1126/SCIADV.1600292
5MZJ	X-ray	2	Inactive	Antagonist	10.1016/j.str.2017.06.012
5MZP	X-ray	2.1	Inactive	Antagonist	10.1016/j.str.2017.06.012
5N2R	X-ray	2.8	Inactive	Antagonist	10.1016/j.str.2017.06.012
5NLX	X-ray	2.1	Inactive	Antagonist	10.1038/S41467-017-00630-4
5NM2	X-ray	2	Inactive	Antagonist	10.1038/S41467-017-00630-4
5NM4	X-ray	1.7	Inactive	Antagonist	10.1038/S41467-017-00630-4
5OLG	X-ray	1.9	Inactive	Antagonist	10.1038/S41598-017-18570-W
5OLH	X-ray	2.6	Inactive	Antagonist	10.1038/S41598-017-18570-W
5OLO	X-ray	3.1	Inactive	Antagonist	10.1038/S41598-017-18570-W
5OLV	X-ray	2	Inactive	Antagonist	10.1038/S41598-017-18570-W

5OLZ	X-ray	1.9	Inactive	Antagonist	10.1038/S41598-017-18570-W
5OM1	X-ray	2.1	Inactive	Antagonist	10.1038/S41598-017-18570-W
5OM4	X-ray	2	Inactive	Antagonist	10.1038/S41598-017-18570-W
5TE5	X-ray	4	Inactive	Agonist	10.1073/PNAS.1617446114
5UIG	X-ray	3.5	Inactive	Antagonist	10.1073/PNAS.1621423114
5UVI	X-ray	3.2	Inactive	Antagonist	10.1107/S205225251700570X
5VRA	X-ray	2.4	Inactive	Antagonist	10.1038/NPROT.2017.135
5X7D	X-ray	2.7	Inactive	Allosteric antagonist	10.1038/NATURE23652
5YC8	X-ray	2.5	Inactive	Antagonist	10.1038/S41589-018-0152-Y
5ZHP	X-ray	3.1	Inactive	Antagonist	10.1073/PNAS.1813988115
5ZK3	X-ray	2.6	Inactive	Antagonist	10.1038/S41589-018-0152-Y
5ZK8	X-ray	3	Inactive	Antagonist	10.1038/S41589-018-0152-Y
5ZKB	X-ray	3	Inactive	Antagonist	10.1038/S41589-018-0152-Y
5ZKC	X-ray	2.3	Inactive	Antagonist	10.1038/S41589-018-0152-Y
6AQF	X-ray	2.5	Inactive	Antagonist	10.1016/J.CELL.2017.12.004
6GT3	X-ray	2	Inactive	Antagonist	-
6JZH	X-ray	2.3	Inactive	Antagonist	10.1107/S1600576719012846
6LPJ	X-ray	1.8	Inactive	Antagonist	10.1038/S41598-020-76277-X
6LPK	X-ray	1.8	Inactive	Antagonist	10.1038/S41598-020-76277-X
6LPL	X-ray	2	Inactive	Antagonist	10.1038/S41598-020-76277-X
6MH8	X-ray	4.2	Inactive	Antagonist	10.1107/S205225251900263X
6OBA	X-ray	3.1	Inactive	Antagonist	10.1038/S41589-020-0549-2
6OFJ	cryo-EM	4.5	Inactive	Apo	10.1074/JBC.RA119.010089
6PRZ	X-ray	2.8	Inactive	Antagonist	10.1107/S2052252519013137
6PS0	X-ray	3.4	Inactive	Antagonist	10.1107/S2052252519013137
6PS1	X-ray	3.2	Inactive	Antagonist	10.1107/S2052252519013137
6PS2	X-ray	2.4	Inactive	Antagonist	10.1107/S2052252519013137
6PS3	X-ray	2.5	Inactive	Antagonist	10.1107/S2052252519013137
6PS4	X-ray	2.6	Inactive	Antagonist	10.1107/S2052252519013137
6PS5	X-ray	2.9	Inactive	Antagonist	10.1107/S2052252519013137
6PS6	X-ray	2.7	Inactive	Antagonist	10.1107/S2052252519013137
6PS7	X-ray	1.9	Inactive	Antagonist	10.1107/S2052252519013137
6S0L	X-ray	2.7	Inactive	Antagonist	10.1107/S2052252520011379
6S0Q	X-ray	2.7	Inactive	Antagonist	-
6WQA	X-ray	2	Inactive	Antagonist	10.1107/S2052252520012701
6ZDR	X-ray	1.9	Inactive	Antagonist	10.1002/ANIE.202003788
6ZDV	X-ray	2.1	Inactive	Antagonist	10.1002/ANIE.202003788
7ARO	X-ray	3.1	Inactive	Agonist (partial)	10.1021/ACS.JMEDCHEM.0C01856
7BVQ	X-ray	2.5	Inactive	Inverse agonist	10.1038/S41422-020-00424-2

Appendix B6: The Non-/Thermostabilised Receptor Structures Analysed in Section 5.3

PDB	Method	Resolution (Å)	State	Reference
2R4R	X-ray	3.4	Non-Thermo	-
2R4S	X-ray	3.4	Non-Thermo	10.1038/NATURE06325
2RH1	X-ray	2.4	Non-Thermo	10.1126/SCIENCE.1150577
2VT4	X-ray	2.7	Thermo	10.1038/NATURE07101
2Y00	X-ray	2.5	Thermo	10.1038/NATURE09746
2Y01	X-ray	2.6	Thermo	10.1038/NATURE09746
2Y02	X-ray	2.6	Thermo	10.1038/NATURE09746
2Y03	X-ray	2.9	Thermo	10.1038/NATURE09746
2Y04	X-ray	3.1	Thermo	10.1038/NATURE09746
2YCW	X-ray	3	Thermo	10.1073/PNAS.1100185108
2YCX	X-ray	3.3	Thermo	-
2YCY	X-ray	3.2	Thermo	10.1073/PNAS.1100185108
2YCZ	X-ray	3.7	Thermo	10.1073/PNAS.1100185108
2YDO	X-ray	3	Thermo	10.1038/NATURE10136
2YDV	X-ray	2.6	Thermo	-
3D4S	X-ray	2.8	Thermo	10.1016/J.STR.2008.05.001
3EML	X-ray	2.6	Non-Thermo	10.1126/SCIENCE.1164772
3KJ6	X-ray	3.4	Non-Thermo	10.1038/NATURE08650
3NY8	X-ray	2.8	Thermo	-
3NY9	X-ray	2.8	Thermo	10.1021/JA105108Q
3NYA	X-ray	3.2	Thermo	-
3P0G	X-ray	3.5	Non-Thermo	10.1038/NATURE09648
3PDS	X-ray	3.5	Thermo	10.1038/NATURE09665
3PWH	X-ray	3.3	Thermo	10.1016/J.STR.2011.06.014
3QAK	X-ray	2.7	Non-Thermo	10.1126/SCIENCE.1202793
3REY	X-ray	3.3	Thermo	-
3RFM	X-ray	3.6	Thermo	-
3SN6	X-ray	3.2	Non-Thermo	10.1038/NATURE10361
3UZA	X-ray	3.3	Thermo	10.1021/jm201376w
3UZC	X-ray	3.3	Thermo	10.1021/jm201376w
3VG9	X-ray	2.7	Non-Thermo	10.1038/NATURE10750
3VGA	X-ray	3.1	Non-Thermo	10.1038/NATURE10750
3ZPQ	X-ray	2.8	Thermo	10.1021/jm400140q
3ZPR	X-ray	2.7	Thermo	10.1021/jm400140q
4AMI	X-ray	3.2	Thermo	10.1016/J.STR.2012.03.014
4AMJ	X-ray	2.3	Thermo	10.1016/J.STR.2012.03.014
4BVN	X-ray	2.1	Thermo	10.1371/JOURNAL.PONE.0092727
4EIY	X-ray	1.8	Non-Thermo	10.1126/SCIENCE.1219218
4GBR	X-ray	4	Non-Thermo	10.1371/JOURNAL.PONE.0046039
4GPO	X-ray	3.5	Thermo	10.1038/NSMB.2504
4LDE	X-ray	2.8	Non-Thermo	10.1038/NATURE12572

4LDL	X-ray	3.1	Non-Thermo	10.1038/NATURE12572
4LDO	X-ray	3.2	Non-Thermo	10.1038/NATURE12572
4QKX	X-ray	3.3	Thermo	10.1073/PNAS.1410415111
4UG2	X-ray	2.6	Thermo	10.1124/mol.114.097360
4UHR	X-ray	2.6	Thermo	10.1124/mol.114.097360
5A8E	X-ray	2.4	Thermo	10.1124/MOL.115.101030
5D5A	X-ray	2.5	Non-Thermo	10.1107/S2059798315021683
5D5B	X-ray	3.8	Non-Thermo	10.1107/S2059798315021683
5D6L	X-ray	3.2	Non-Thermo	10.1038/NPROT.2017.057
5F8U	X-ray	3.4	Thermo	10.1038/NSMB.3130
5G53	X-ray	3.4	Non-Thermo	10.1038/NATURE18966
5IU4	X-ray	1.7	Thermo	10.1021/acs.jmedchem.6b00653
5IU7	X-ray	1.9	Thermo	10.1021/acs.jmedchem.6b00653
5IU8	X-ray	2	Thermo	10.1021/acs.jmedchem.6b00653
5IUA	X-ray	2.2	Thermo	10.1021/acs.jmedchem.6b00653
5IUB	X-ray	2.1	Thermo	10.1021/acs.jmedchem.6b00653
5JTB	X-ray	2.8	Non-Thermo	10.1126/SCIADV.1602952
5K2A	X-ray	2.5	Non-Thermo	10.1126/SCIADV.1600292
5K2B	X-ray	2.5	Non-Thermo	10.1126/SCIADV.1600292
5K2C	X-ray	1.9	Non-Thermo	10.1126/SCIADV.1600292
5K2D	X-ray	1.9	Non-Thermo	10.1126/SCIADV.1600292
5MZJ	X-ray	2	Thermo	10.1016/j.str.2017.06.012
5MZP	X-ray	2.1	Thermo	10.1016/j.str.2017.06.012
5N2R	X-ray	2.8	Thermo	10.1016/j.str.2017.06.012
5NLX	X-ray	2.1	Thermo	-
5NM2	X-ray	2	Thermo	10.1038/S41467-017-00630-4
5NM4	X-ray	1.7	Thermo	-
5OLG	X-ray	1.9	Thermo	10.1038/S41598-017-18570-W
5OLH	X-ray	2.6	Thermo	-
5OLO	X-ray	3.1	Thermo	10.1038/S41598-017-18570-W
5OLV	X-ray	2	Thermo	10.1038/S41598-017-18570-W
5OLZ	X-ray	1.9	Thermo	10.1038/S41598-017-18570-W
5OM1	X-ray	2.1	Thermo	10.1038/S41598-017-18570-W
5OM4	X-ray	2	Thermo	10.1038/S41598-017-18570-W
5UIG	X-ray	3.5	Non-Thermo	10.1073/PNAS.1621423114
5UVI	X-ray	3.2	Non-Thermo	10.1107/S205225251700570X
5VRA	X-ray	2.4	Non-Thermo	10.1038/NPROT.2017.135
5WF5	X-ray	2.6	Thermo	10.1016/J.STR.2017.12.013
5WF6	X-ray	2.9	Thermo	10.1016/J.STR.2017.12.013
5X7D	X-ray	2.7	Non-Thermo	10.1038/NATURE23652
6AQF	X-ray	2.5	Non-Thermo	10.1016/J.CELL.2017.12.004
6E67	X-ray	3.7	Non-Thermo	10.1016/J.CELL.2019.04.021
6GDG	Cryo-EM	4.1	Non-Thermo	10.7554/ELIFE.35946

6GT3	X-ray	2	Non-Thermo	-
6H7J	X-ray	2.8	Thermo	10.1126/SCIENCE.AAU5595
6H7L	X-ray	2.7	Thermo	10.1126/SCIENCE.AAU5595
6H7M	X-ray	2.8	Thermo	10.1126/SCIENCE.AAU5595
6H7N	X-ray	2.5	Thermo	10.1101/436212
6H7O	X-ray	2.8	Thermo	10.1126/SCIENCE.AAU5595
6IBL	X-ray	2.7	Thermo	10.1038/S41586-020-2419-1
6JZH	X-ray	2.3	Non-Thermo	10.1107/S1600576719012846
6LPJ	X-ray	1.8	Non-Thermo	10.1038/S41598-020-76277-X
6LPK	X-ray	1.8	Non-Thermo	10.1038/S41598-020-76277-X
6LPL	X-ray	2	Non-Thermo	10.1038/S41598-020-76277-X
6MH8	X-ray	4.2	Non-Thermo	10.1107/S205225251900263X
6MXT	X-ray	3	Non-Thermo	10.1038/s41589-018-0145-x
6N48	X-ray	3.2	Non-Thermo	10.1126/SCIENCE.AAW8981
6NI3	Cryo-EM	3.8	Non-Thermo	10.1038/S41594-019-0330-Y
6OBA	X-ray	3.1	Non-Thermo	10.1038/S41589-020-0549-2
6PRZ	X-ray	2.8	Non-Thermo	10.1107/S2052252519013137
6PS0	X-ray	3.4	Non-Thermo	10.1107/S2052252519013137
6PS1	X-ray	3.2	Non-Thermo	10.1107/S2052252519013137
6PS2	X-ray	2.4	Non-Thermo	10.1107/S2052252519013137
6PS3	X-ray	2.5	Non-Thermo	10.1107/S2052252519013137
6PS4	X-ray	2.6	Non-Thermo	10.1107/S2052252519013137
6PS5	X-ray	2.9	Non-Thermo	10.1107/S2052252519013137
6PS6	X-ray	2.7	Non-Thermo	10.1107/S2052252519013137
6PS7	X-ray	1.9	Non-Thermo	10.1107/S2052252519013137
6S0L	X-ray	2.7	Non-Thermo	10.1107/S2052252520011379
6S0Q	X-ray	2.7	Non-Thermo	10.1107/S2052252520011379
6TKO	Cryo-EM	3.3	Thermo	10.1038/S41586-020-2419-1
6WQA	X-ray	2	Non-Thermo	10.1107/S2052252520012701
6ZDR	X-ray	1.9	Non-Thermo	10.1002/ANIE.202003788
6ZDV	X-ray	2.1	Non-Thermo	10.1002/ANIE.202003788
7ARO	X-ray	3.1	Non-Thermo	10.1021/acs.jmedchem.0c01856
7BTS	X-ray	3.1	Non-Thermo	10.1038/S41422-020-00424-2
7BU6	X-ray	2.7	Non-Thermo	10.1038/S41422-020-00424-2
7BU7	X-ray	2.6	Non-Thermo	10.1038/S41422-020-00424-2
7BVQ	X-ray	2.5	Non-Thermo	10.1038/S41422-020-00424-2
7BZ2	Cryo-EM	3.8	Non-Thermo	10.1038/S41421-020-0176-9
7DHI	Cryo-EM	3.3	Non-Thermo	10.1093/NSR/NWAA284
7DHR	Cryo-EM	3.8	Non-Thermo	10.1093/NSR/NWAA284
7JJO	Cryo-EM	2.6	Non-Thermo	10.1016/J.MOLCEL.2020.08.001

Appendix B7: The Non-/Fused Orexin Receptor Structures Analysed in Section 5.4

PDB	Method	Resolution (Å)	Fusion	Reference
4ZJ8	X-ray	2.8	GlgA glycogen synthase	10.1038/NSMB.3183
4S0V	X-ray	2.5	Glycogen synthase	10.1038/NATURE14035
4ZJC	X-ray	2.8	GlgA glycogen synthase	10.1038/NSMB.3183
5WQC	X-ray	2	Glycogen synthase	10.1016/J.STR.2017.11.005
5WS3	X-ray	2.3	Glycogen synthase	10.1016/J.STR.2017.11.005
6TO7	X-ray	2.3	-	10.1021/acs.jmedchem.9b01787
6TOD	X-ray	2.1	-	10.1021/acs.jmedchem.9b01787
6TOS	X-ray	2.1	-	10.1021/acs.jmedchem.9b01787
6TOT	X-ray	2.2	-	10.1021/acs.jmedchem.9b01787
6TP3	X-ray	3	-	10.1021/acs.jmedchem.9b01787
6TP4	X-ray	3	-	10.1021/acs.jmedchem.9b01787
6TP6	X-ray	2.3	-	10.1021/acs.jmedchem.9b01787
6TPG	X-ray	2.7	GlgA glycogen synthase	10.1021/acs.jmedchem.9b01787
6TPJ	X-ray	2.7	GlgA glycogen synthase	10.1021/acs.jmedchem.9b01787
6TPN	X-ray	2.6	GlgA glycogen synthase	10.1021/acs.jmedchem.9b01787
6TQ4	X-ray	2.3	-	10.1021/acs.jmedchem.9b01787
6TQ6	X-ray	2.6	-	10.1021/acs.jmedchem.9b01787
6TQ7	X-ray	2.7	-	10.1021/acs.jmedchem.9b01787
6TQ9	X-ray	2.7	-	10.1021/acs.jmedchem.9b01787
6V9S	X-ray	3.5	GlgA glycogen synthase	10.1073/PNAS.2002704117
7L1U	Cryo-EM	3.2	-	10.1038/S41467-021-21087-6
7L1V	Cryo-EM	3	-	10.1038/S41467-021-21087-6

Appendix B8: The AlphaFold and Experimentally Resolved Structures Analysed in Section 5.5

Code	Method	Resolution (Å)	State	Ligand	Reference
2R4R	X-ray	3.4	Inactive	Inverse agonist	-
2R4S	X-ray	3.4	Inactive	Inverse agonist	10.1038/NATURE06325
2RH1	X-ray	2.4	Inactive	Inverse agonist	10.1126/SCIENCE.1150577
3D4S	X-ray	2.8	Inactive	Inverse agonist	10.1016/J.STR.2008.05.001
3KJ6	X-ray	3.4	Inactive	Inverse agonist	10.1038/NATURE08650
3NY8	X-ray	2.8	Inactive	Inverse agonist	-
3NY9	X-ray	2.8	Inactive	Inverse agonist	10.1021/JA105108Q
3NYA	X-ray	3.2	Inactive	Antagonist	-
3P0G	X-ray	3.5	Active	Agonist	10.1038/NATURE09648
3PDS	X-ray	3.5	Inactive	Agonist	10.1038/NATURE09665
3SN6	X-ray	3.2	Active	Agonist	10.1038/NATURE10361

4GBR	X-ray	4	Inactive	Inverse agonist	10.1371/JOURNAL.PONE.0046039
4LDE	X-ray	2.8	Active	Agonist	10.1038/NATURE12572
4LDL	X-ray	3.1	Active	Agonist	10.1038/NATURE12572
4LDO	X-ray	3.2	Active	Agonist	10.1038/NATURE12572
4QKX	X-ray	3.3	Active	Agonist	10.1073/PNAS.1410415111
5D5A	X-ray	2.5	Inactive	Inverse agonist	10.1107/S2059798315021683
5D5B	X-ray	3.8	Inactive	Inverse agonist	10.1107/S2059798315021683
5D6L	X-ray	3.2	Inactive	Inverse agonist	10.1038/NPROT.2017.057
5JQH	X-ray	3.2	Inactive	Inverse agonist	10.1038/NATURE18636
5X7D	X-ray	2.7	Inactive	Antagonist	10.1038/NATURE23652
6E67	X-ray	3.7	Active	Agonist	10.1016/J.CELL.2019.04.021
6MXT	X-ray	3	Active	Agonist	10.1038/s41589-018-0145-x
6N48	X-ray	3.2	Active	Agonist PAM	10.1126/SCIENCE.AAW8981
6NI3	Cryo-EM	3.8	Active	Agonist	10.1038/S41594-019-0330-Y
6OBA	X-ray	3.1	Inactive	Antagonist NAM	10.1038/S41589-020-0549-2
6PRZ	X-ray	2.8	Inactive	Antagonist	10.1107/S2052252519013137
6PS0	X-ray	3.4	Inactive	Antagonist	10.1107/S2052252519013137
6PS1	X-ray	3.2	Inactive	Antagonist	10.1107/S2052252519013137
6PS2	X-ray	2.4	Inactive	Antagonist	10.1107/S2052252519013137
6PS3	X-ray	2.5	Inactive	Antagonist	10.1107/S2052252519013137
6PS4	X-ray	2.6	Inactive	Antagonist	10.1107/S2052252519013137
6PS5	X-ray	2.9	Inactive	Antagonist	10.1107/S2052252519013137
6PS6	X-ray	2.7	Inactive	Antagonist	10.1107/S2052252519013137
7BZ2	Cryo-EM	3.8	Active	Agonist	10.1038/S41421-020-0176-9
7DHI	Cryo-EM	3.3	Active	Agonist	10.1093/NSR/NWAA284
7DHR	Cryo-EM	3.8	Active	Agonist	10.1093/NSR/NWAA284
P07550	Alpha Fold	-	-	-	https://alphafold.ebi.ac.uk/entry/P07550