Molecular modelling of the GLP-1 receptor – a prototypic class B GPCR

Carla Gómez Santiago

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

School of Biomedical Sciences

Faculty of Biological Sciences

July 2021

Intellectual Property and Publication Statements

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The work in Chapter 3 of the thesis has appeared in publication as follows:
Gómez Santiago, Carla, Emanuele Paci, and Dan Donnelly. 2018. 'A mechanism for agonist activation of the glucagon-like peptide-1 (GLP-1) receptor through modelling & molecular dynamics', *Biochem Biophys Res Commun*, 498: 359-65. http://pubmed.ncbi.nlm.nih.gov/29397068/

The following aspects of the publication re attributable to me: build of the inactive GLP-1R, simulation of the GLP-1R model and data analysis.

The co-authors reviewed all aspects of the publication and provided comments which I acted upon to develop the paper.

Acknowledgements

I would like to thank my supervisor Dr. Dan Donnelly for his patience and continual support and guidance during this PhD and the motivation he provides for running. I would also like to thank Dr. Emanuele Paci for his invaluable advice, not only academical but also personal.

My thanks go to my family Carlota and Hugo and my husband Gordon for their support. I also want to thank my friend Dr. Safi Masandi, I enjoyed our talks and little book club. A special mention is also owed to Timber for keeping me sane during lockdown.

Finally, I would like to thank CONACyT for their financial sponsoring.

Abstract

The main aim of this project was to gain an understanding of GLP-1R's molecular mechanism of action and the key interactions leading to agonist-mediated activation. At the start of this project there were no experimentally determined GLP-1R structures available, the initial approach was to use experimentally determined structures of other class B GPCRs for GLP-1R homology modelling. Soon different GLP-1R structures became available (Jazayeri et al. 2017; Zhang et al. 2017; Song et al. 2017) providing valuable information. However, none of these structures showed GLP-1R in complex with GLP-1, the endogenous peptide; based on the experimental available structures, GLP-1R -GLP-1 complex models were built and simulated using different methods of molecular dynamics simulations. After defining interactions in the active and inactive state, the second aim was to identify conformational changes leading to activation. Since Class B share structural characteristics and possible an activation mechanism, PTH_1 receptor model was built.

Simulation of GLP-1R and PTH_1 receptor in complex with their respectively ligands show ligands disrupting the TM1-TM2 interface and relying on hydrophobic residues to facilitate the movement into the binding pocket. At the extracellular end, the movement of the ligand into the binding pocket shift TM7 towards TM6, contributing to TM6 rearrangement, in addition interaction between charged residues in the N-terminal and TM6 allowed TM6 flexibility and outward movement while maintaining the stability of the receptor. Interaction between charged residues in the N-terminal and $Arg^{2.60b}$ contribute to the binding and stabilisation of the ligand as translation of conformational changes in TM3-TM2 middle region, where $Asn^{3.43b}$ removes spatial restraints from the core of the receptor. The removal of spatial restrains is seen as TM3-TM2 packing allowing $Leu^{3.47b}$ rotamer change pushing a hydrophobic residue in TM6 resulting in the outward movement stabilised by changes in hydrophobic residues rotamers.

Contents

List of Figures and Tables

List of figures

List of tables

Abbreviations

$\vec{F}$	force
$\vec{r}$	acceleration
ϕ	EM density
ξ	scaling factor
β2AR	beta 2 adrenergic receptor
2D	two-dimensional
3D	three-dimensional
Å	angstrom
$Å^2$	square Angstrom
AC	adenylyl cyclase
AGS	activators of G protein signalling
AKT	threonine kinase
AMBER	Assisted Model Building with Energy Refinement
AP2	adaptor protein 2
AT1R	angiotensin II type 1 receptor
atm	atmosphere
C-terminal	carboxyl-terminal
cAMP	3',5'-cyclic adenosine monophosphate
CHARMM	Chemistry at Harvard Macromolecular Mechanics
CRF	corticotropin-releasing factor
CRF1R	corticotropin-releasing factor receptor type 1
CRHR1	corticotropin-releasing hormone 1
CRHR2	corticotropin-releasing hormone 2
CRLR	calcitonin-like receptor
cryo-EM	cryo-electron microscopy
CT	calcitonin
CTR	calcitonin receptor
DPP-4	dipeptidyl peptidase -IV
ECD	extracellular domain

ECL	extracellular loop
Ex-4	exendin-4
FF	force fields
FRET	Föster Resonance Energy Transfer
fs	femtosecond
G protein	guanine nucleotide-binding proteins
GCG	Glucagon
GCGR	Glucagon receptor
GDP	guanosine-5'-diphosphate
GGRPR	gene-related peptide receptor
GHRHR	growth-hormone-releasing hormone receptor
GIP	gastric inhibitory polypeptide / glucose-dependent insulinotropic polypeptide
GIPR	gastric inhibitory polypeptide receptor
GLP-1	Glucagon-like peptide-1
GLP-1R	Glucagon-like peptide-1 receptor
GLP-1RAs	GLP-1R agonists
GLP-2	Glucagon-like peptide-2
GPCRs	G protein-coupled receptors
GRF	growth hormone-releasing factor
GRKs	GPCRs kinases
GRPD	glicentin-related pancreatic polypeptide
GTP	guanosine-5'-triphosphate
ICL	intracellular loop
IDF	International Diabetes Federation
IP-1	intervening peptide-1
IP-2	intervening peptide-2
ISG	insulin-secretory granule
K	Kelvin
Kcal	kilocalorie
K_d	equilibrium dissociation constant
kDa	kilodalton
LDL	low-density lipoprotein
m	mass

MAPKs	mitogen-activated protein kinases
MDFF	Molecular Dynamics Flexible Fitting
MDs	Molecular dynamics
MI	myocardial infarction
MOR	μ-opioid receptor
N-terminal	amino-terminal
NaCl	Sodium chloride
NAM	negative allosteric modulator
NAMD	Nanoscale Molecular Dynamics
Nb	nanobody
NEP24.11	neutral endopeptidase
NICE	National Institute for Health and Care Excellence
NPAT	constant pressure, area, and temperature
ns	nanosecond
NTD	N-terminus domain
NVT	constant volume and temperature
OPM	orientation protein membrane
PAC1/ADCYAP1R1	adenylate cyclase activating polypeptide receptor
PAM	positive allosteric modulator
PC	prohormone convertase
PDB	Protein data bank
PGS	*Pyrococcus abyssi* glycogen synthase
PI3K	phosphoinositide 3-kinase
PLC	phospholipase C
PME	particle mesh Ewald
POPC	phosphatidylcholine 1-palmitoyl-2-oleoyl-sn-glycero-3-phophochoine
POPE	Phosphatidylethanolamine
POPI	Phosphatidylinositol
POPS	Phosphatidylserine

ps	picosecond
PSM	sphingomyelin
PTH	parathyroid hormone
PTH_1	parathyroid hormone 1
PTH_2	parathyroid hormone 2
PTHrP	parathyroid hormone-related protein
R	inactive receptor conformation
R*	active receptor conformation
RAMP	receptor activity-modifying protein
RMSD	root mean square deviation
SASA	solvent accessible surface area
SC	subcutaneous
SCT	secretin
SCTR	secretin receptor
SMS-1	sphingomyelin synthase-1
SNAC	salcaprozate sodium
T2DM	type 2 diabetes mellitus
Targeted MDs	Targeted Molecular Dynamics
TEM	transmission electron microscope
TIP-39	Tuberoinfundibular peptide 39
TM	transmembrane
TMD	transmembrane domain
U_{EM}	cryo-EM electronic density map gradient
U_{MD}	MD potential energy
U_{ss}	potential of secondary structures
VIP	vasoactive intestinal peptide
VIP1	vasoactive intestinal peptide receptor 1
VIP2	vasoactive intestinal peptide receptor 2
VMD	Visual Molecular Dynamics software
VPAC1R	vasoactive intestinal polypeptide type1
VPAC2R	vasoactive intestinal polypeptide type2

Chapter 1: Introduction

G protein-coupled receptors (GPCRs) are the largest family of membrane proteins and regulate multiple physiological processes with clinical relevance, making them important drug targets (reviewed in Kobilka 2007; Hauser et al. 2018; Sriram and Insel 2018). Glucagon-like peptide-1 (GLP-1) and Parathyroid hormone (PTH) are examples of peptide ligands that bind and activate members of the Class B subfamily of GPCRs. GLP-1 and its receptor (GLP-1R) are involved in glucose homeostasis. Several GLP-1 analogues have been approved for the treatment of diabetes (Fehse et al. 2005; Limited 2020; Bond 2006; Leon et al. 2017; Christensen and Knop 2010; Green et al. 2007; Jackson et al. 2010; Cummings et al. 2010; Skrivanek et al. 2012; Pratley et al. 2021). New therapies involving the use of multiple co-agonist combinations aiming to enhance GLP-1R efficacy by using synergistic response have demonstrated clinical efficacy in the treatment of metabolic diseases such diabetes and obesity (reviewed in Baggio and Drucker 2021). The most common co-agonist combinations incorporate glucose-dependent insulinotropic polypeptide receptor (GIPR) agonists and glucagon receptor (GCGR) agonists. So far, the dual agonist tirzepatide, a GIPR and GLP-1R agonist showed to be superior to GLP-1R mono-agonist semaglutide with respect to the mean change in the glycated haemoglobin level from baseline to 40 weeks (Frías et al. 2021). Recent studies of the tri-agonist Peptide 20, a GIPR, GLP-1R and GCGR agonists, currently in phase 1 clinical trial, suggest a superior efficacy over GLP-1R mono-agonists (Alexiadou K., Anyiam O, and Tan T. 2019). Similarly, PTH-based peptide agonists are used to treat osteoporosis (Fan et al. 2020; Reginster et al. 2019; Lewiecki 2006) due to the role of PTH in calcium homeostasis via PTH1R. Despite the important clinical relevance, at the start of this PhD project there were no experimentally determined structures of either peptide-receptor complex, or any reliable structure-based understanding of how ligand binding results in receptor activation. The aim of this PhD project was initially to utilise homology modelling and molecular dynamics to understand peptide activation of these Class B GPCRs. However, as the project progressed, several new structures became available, enabling an increasing sophisticated approach to be adopted to understand peptides bind and activate these GPCRs. In order to set the scene, the thesis will first cover the basic background science, before establishing the research questions and details of the work carried out to answer them.

1.1. Proteins

Proteins are complex and dynamic molecules involved in all biological processes. It is the diversity and abundance of proteins that allows them to fulfil a wide range of functions. Protein function depends on structure; therefore, comprehension of protein architecture and dynamics allows a better understanding of protein functions, with multiple applications, such drug discovery.

1.1.1. Protein structure

Although environmental factors affect proteins conformation, the three-dimensional shape or tertiary structure of a protein is ultimately determined by the amino acid sequence, denominated primary structure (Sela, Anfinsen, and Harrington 1957). This primary structure, expressed at the genome level, defines the spatial distribution of individual amino acids; any change or mutation in the amino acid sequence potentially alter shape and function of the protein. Defined by the primary structure, individual residues are arranged into α-helices and β-sheets, denominated secondary structure, which shape the protein. Linkers formed by individual amino acids confer flexibility and communication between proteins regions or domains (Dobson 1990) while electrostatic forces, disulphide bridges, hydrophobic or hydrophilic interactions stabilise the protein architecture. Stable structures or subunits form complexes or quaternary structures (reviewed in Berg JM 2002; Zaretsky and Wreschner 2008; Figure 1.1).

Advances in experimental methods have increased the number of known primary and tertiary structures. However, not all the available experimental methods are suitable for the resolution of all proteins, resulting in an increasing gap between known sequences and three-dimensional structures. For cases in which experimental methods are not suitable for the protein, a structure-based approach is used, taking advantage of the information in the primary structure needed to shape a functional tertiary structure (Sela, Anfinsen, and Harrington 1957). Proteins tend to form stable structures at the smallest cost of energy, meaning that mutations on the primary sequence occur at individual residues without gross changes of the protein architecture (Lenstra, Hofsteenge, and Beintema 1977). Therefore, homologous

proteins may show different sequences but a similar architecture, a common ancestor and similar active or hydrophobic sites (Muirhead et al. 1967; M Bajaj and Blundell 1984).

1.1.2. Protein kinetics and dynamics

In contrast with the static images above, proteins are in constant movement undergoing conformational changes to find a stable energy-favouring conformation enabling them to execute their function. Protein dynamics aims to understand the relationship between conformational and functional state by studying time-dependent changes in atomic coordinates (Henzler-Wildman and Kern 2007; Dinner et al. 2000; Austin et al. 1975). Usually, the native state represents a functional protein at its lowest energy in a particular environment, eventually achieving energy equilibrium to perform their function (Austin et al. 1975; Frauenfelder, Sligar, and Wolynes 1991). Proteins are continually exploring different energy-landscapes because of thermal energy and crossing energy barriers (Elber and Karplus 1987). The probability of a protein to go from a conformational state (thermodynamics) to another depends on the energy difference between states; while the kinetics of the process depends on the height of the energy barrier, and the width of the energy well describes conformational flexibility (reviewed in Henzler-Wildman and Kern 2007). In addition to temperature, pressure and solvent conditions, ligand binding changes the energy well and barriers of the receptor, stabilising a different state with different architecture and function; or favouring a different conformation.

1.1.3. Membrane proteins

The environment affects protein structure, function and dynamics (Singer and Nicolson 1972). Of particular interest is the influence of lipid bilayer to membrane proteins, which account for almost 30% of the human proteome (Fagerberg et al. 2010; Hedger and Sansom 2016; Jackowski 1994; van Meer, Voelker, and Feigenson 2008).

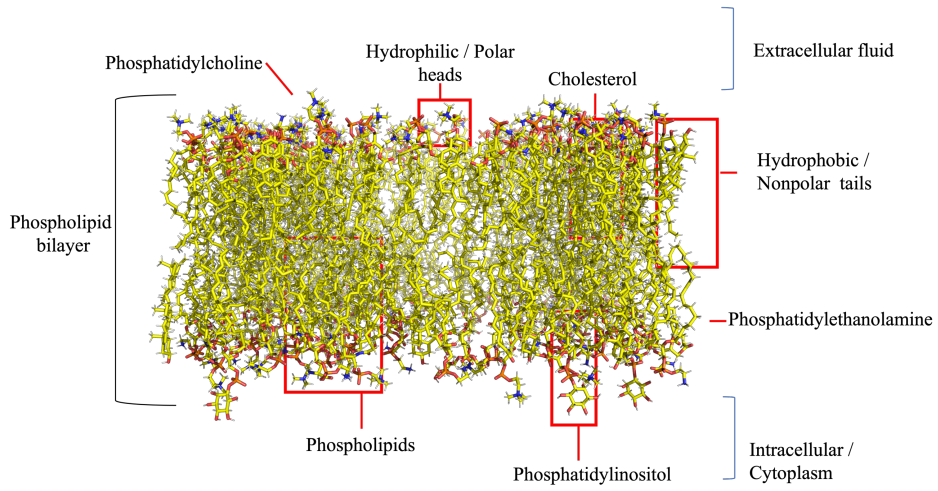
1.1.3.1. Physiological role of lipid bilayer

Lipid bilayers membranes are dynamic and asymmetric heterogeneous structures composed of cholesterol and phospholipids (Bloch 1991; van Meer and de Kroon 2011; Caillon, Lequin, and Khemtémourian 2013; Hallberg 1984b; Jackowski 1996; Singer and Nicolson 1972). They perform two main tasks: 1) act as a semipermeable barrier between different compartments and 2) are a highly organised place where biochemical functions take place (Sackmann et al. 1984; Hackenbrock C.R. 1980; Singer and Nicolson 1972).

Lipid composition affects protein function through chemical interactions (Sackmann et al. 1984; Helfrich 1973). In gross terms, lipid-lipid interactions modulate physiochemical properties of the membrane, such fluidity, membrane thickness, elasticity, curvature, surface tension order and lipid rafts; while protein-lipid interactions influence protein architecture, orientation and location (reviewed in Crane and Tamm 2004; Zocher et al. 2012).

The unique characteristic of lipid bilayers allows the exposure of the hydrophilic lipid head to the aqueous solution while the acyl chains inside the hydrophobic core influence transmembrane proteins (Spector and Yorek 1985; Sackmann et al. 1984; Figure 1.2).

Figure 1.1 Average membrane composition



1.1.3.2. Membrane composition/ heterogeneity

Cell membranes are composed of more than 200 different lipids depending on the tissue and function (Myher, Kuksis, and Pind 1989).

Cholesterol is an essential component of eukaryotic membranes, making up between 25 to 50% of lipid component in mammals cells (Bloch 1991). Cholesterol provides structural support to the membrane while maintaining the fluidity of the membrane (Crane and Tamm 2004; Zocher et al. 2012; Cherezov et al. 2007; Hanson et al. 2008; Liu, Chun, et al. 2012; Wu, Wang, et al. 2014). Identification of cholesterol-binding sites at crystal structures of transmembrane GPCRs has shown that cholesterol-protein interactions improve conformational stability (Zocher et al. 2012; Saxena and Chattopadhyay 2012; Hanson et al. 2008; Lyman et al. 2009; Manna et al. 2016; Guixà-González et al. 2017).

Phospholipids are the most abundant lipids in membranes (Alberts B 2002). Phosphatidylcholine (POPC) is the most abundant phospholipid in mammalian membranes; it is mainly located in the extracellular leaflet (Caillon, Lequin, and Khemtémourian 2013; Hallberg 1984b; Jackowski 1996). Phosphatidylethanolamine (POPE) comprises about 15-25% in mammals' cells and it is mainly found in the inner membrane (Patel and Witt 2017; Vance and Vance 2008; MacDonald et al. 2015; Koldso et al. 2014; van Meer and de Kroon 2011; Devaux and Morris 2004). Phosphatidylserine (POPS) accounts for 5-10% of cellular phospholipids, usually also found at the inner membrane (Devaux and Morris 2004; van Meer and de Kroon 2011; MacDonald et al. 2015; Vance and Vance 2008; Koldso et al. 2014). Phosphatidylinositol (POPI) accounts for 5% of lipids in the membrane, it is mainly found in the cytosolic leaflet (Vance and Vance 2008; Koldso et al. 2014).

Sphingolipids are a small group of lipids in mammals, they are mainly found on the outer plasma membrane (Devaux and Morris 2004; van Meer and de Kroon 2011; Verkleij et al. 1973).

1.2. Methods used for the study of protein structure

1.2.1. Experimental methods used to obtain protein structure

Different methods are used with the aim to determine the functional structure of proteins, especially of GPCRs. Within the most common used methods are X-ray crystallography and single particle cryogenic electron microscopy (cryo-EM). However, the determination of protein structure in stable functional state is challenging, and in most cases, it requires modifications or the addition of stabilising proteins. Therefore, the resulting structures are not always an accurate description of the functional state but the static representation of the protein in a moment of time.

1.2.1.1. X-ray crystallography

X-ray crystallography method obtains the position of atoms by striking a beam of X-rays into a crystal made up from purified and highly concentrate protein. The strike diffracts the waves into a predictable pattern, based on the crystal lattice structure, which can be used to generate an electron density map from which a molecular structure is built and refined based on the protein sequence (reviewed in Smyth and Martin 2000).

Crystallisation of GPCRs-complexes is challenging, and in order to obtain GPCRs in the active state C-terminal peptide of G proteins (Scheerer et al. 2008), nanobodies (Rasmussen, Choi, et al. 2011a; Rasmussen, DeVree, et al. 2011), or the addition of mini-G proteins (Carpenter et al. 2016) have been used to stabilise GPCRs complexes. However, obtaining high quality crystals is challenging, especially flexible and disordered areas such as loops. Large and symmetrical complexes crystallise well, but not flexible or membrane proteins. Many factors potentially alter the quality and architecture of the protein i) the x-ray radiation can damage the protein; ii) the addition of mutations, fusion proteins or engineer residues used to stabilise the protein during the crystallisation process or iii) the use of detergents used to isolate the protein can potentially alter the protein's architecture and surface, as well as lipid-protein interactions (Kunji et al. 2008; Bill et al. 2011).

1.2.1.2. Single-particle Cryo-Electron Microscopy

Advances in electron microscopes, electron detectors (McMullan, Faruqi, and Henderson 2016) and software allowed the development of single-particle cryo-electron microscopy (cryo-EM). Since cryo-EM rapidly freezes molecules in vitreous (non-crystalline) ice, it allows the study of the sample in almost their native environment without crystallography challenges. Cryo-EM uses a transmission electron microscope (TEM), to record thousands of randomly orientated high-resolution 2D images which are later processed to construct a 3D map (reviewed in Danev, Yanagisawa, and Kikkawa 2019).

Although less changes are applied to the sample, it still undergoes preparation to avoid dehydration caused by the vacuum and radiation of microscope. Usually the vitrification process is not immediate, (~1 s) allowing protein particles to collide with the air-water interface, altering the stability of the GPCR-complex and preferential orientation (reviewed in Noble et al. 2018). Since the quality of the image depends on i) the particles distribution and orientation, ii) the analysis of 2D images, removal of artifacts, empty fields, or invalid particles, iii) sample size as in smaller samples the beam-induced movement of the microscope's carbon film supporting the sample grid can produce a lower signal-to-noise ratio (Zhang et al. 2017; Cheng et al. 2015) determination of high-resolution GPCRs structures is still challenging. Even though cryo-EM has been used in the determination of many structures, there is not an objective quality criterion used to determine the accuracy of the map (reviewed in Cheng et al. 2015).

1.2.2. Computational approach for protein structure prediction

Not all experimental methods are suitable to all proteins. This is reflected as the number of 3D structures deposited in the PBD bank (160796 entries in the PDB bank (Bank 2020) remains significantly lower than the number of known protein sequences (172 million - Institute 2019).

Computational methods are an alternative to experimental methods; they rely on experimental data and computational power to create computational models for the study of proteins structure, interactions and behaviour.

1.2.2.1. Homology Modelling

Homology modelling is a tool for the prediction of the structure from the amino acid sequence of a target protein. Homology modelling relies on the availability of the structure of one or more proteins with similar sequences (templates) Its reliability depends strongly on the similarity between the target and templates.

Homology modelling follows a common procedure: 1) Identification and selection of templates, during this step templates that will be used as targets are selected based on their availability in the PDB bank, structural homology, sequence identity, root mean square deviation (RMSD) of Cα atoms to the target structure, state, resolution and experimental methods used; 2) Alignment of target and template; 3) Model building, where a protein is built using the templates and homology modelling servers or software; 4) Model optimization, where the newly built model is submitted to energy minimisation to avoid steric clashes or poor bond torsion geometrics (reviewed in Jabeen, Mohamedali, and Ranganathan 2019; Robinson, Afzal, and Leader 2014; Guex, Peitsch, and Schwede 2009; Oldham and Hamm 2008; Schrödinger 2015).

1.2.3. Molecular Dynamics simulations

1.2.3.1. Foundations of molecular dynamics methods

Molecular dynamics (MD) simulations is a computational approach that takes advantage of available structural and experimental data to study the movement of atoms over a period of time. MD simulations are based on Newton's equations to describe the trajectory of atoms into time steps. In accordance with Newton's second law of motion:

$$\vec{F}_i = m_i * \vec{r}_i$$

where $\vec{F}_i$ is the force of the atom $_i$, m_i is the mass and $\vec{r}_i$ is the acceleration; therefore, by knowing the position of all atoms it is possible to calculate the force exerted by each atom on others through time (Adcock and McCammon 2006; Schlick 2010; Karplus and Kuriyan 2005). Models known as force fields and based on experimental data, are used to calculate forces on MDs, such electrostatic interactions, covalent bonds and interaction between atoms. (Ackermann 1990; Elber and Karplus 1987; Karplus and Kuriyan 2005; McCammon, Gelin, and Karplus 1977).

1.2.3.2. Molecular dynamics in biology

Since 1977, when the first MD simulation (9.2 ps) of the bovine pancreatic trypsin inhibitor (BPTI) in vacuum took place (McCammon, Gelin, and Karplus 1977) and with the increasing computational power, the size and complexity of the systems simulated has increased. Since MDs relies on experimental data and has the ability to control characteristics of the system studied at an atomic level making it possible to study at timesteps or conditions that would be impossible to study through macroscopic experiments. Membrane proteins are the perfect example as they no longer need to isolate it from its native environment for study.

1.2.3.3. All-atoms classic molecular dynamics

All-atoms MDs simulations applies MD foundations to all the atoms in a system, as it takes into account the potential energy, acceleration, velocity and position of atoms at a moment in time and uses it to compute or calculate the velocity and position in the next movement, resulting in a trajectory or the description of the evolution of a system over time, while taking into account the effect of atoms in the surrounding environment (water, membrane).

Because an enormous computational effort would be needed to do this for every atom, standardized empirically derived force fields (FF) are used, these

designated constants or equations are based on tested properties and experimental data (MacKerell 2000) used to maintain and reproduce molecular geometries of bond lengths, bond angles, bond torsions, non-bonding or improper and electrostatic interactions. FF subdivide potential function into the sum of: Bonded interactions, also known as local contributions, which include covalent bond-stretching (Σ_{bonds}), angle-bending (Σ_{angles}), torsion potential ($\Sigma_{torsion}$) and improper torsion potential ($\Sigma_{improper}$); and non-bonded interactions which include Lennard-Jones (Σ_{LJ}) (short-range Pauli repulsion and long-range Van der Waals attraction) potential and Coulomb electrostatics (Σ_{elec}) (Guvench and MacKerell 2008; MacKerell 2000).

$$U = \underbrace{\sum_{\text{bonds}} \frac{1}{2}k_b(r - r_0)^2 + \sum_{\text{angles}} \frac{1}{2}k_a(\theta - \theta_0)^2 + \sum_{\text{torsions}} \frac{V_n}{2}[1 + \cos(n\phi - \delta)]}_{\text{Bonded interactions, local contributions or intramolecular contributions to the total energy}} + \underbrace{\sum_{\text{improper}} V_{imp} + \sum_{\text{LJ}} 4\epsilon_{ij}\left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}}\right) + \sum_{\text{elec}} \frac{q_i q_j}{r_{ij}}}_{\text{Repulsive, Van der Waals and the Coulombic contributions to the total energy}}$$

Where:
r_0 is obtained from X-ray diffraction experiments
k correspond to the spring constant that may be estimated from infrared or Raman spectra
ϕ is the torsional angle
δ is the phase
n defines the number of minima or maxima between 0
V_n determines the height of the potential barrier
ij interaction between two atoms, i, j
12-6 correspond to the Lennard-Jones (LJ) potential
ϵ_{ij} is the LJ potential corresponding depth of the well for the interaction between two atoms, i, j
σ Value at which the potential becomes zero

In classical MD simulations, covalent bonds do not form or break and to favour efficiency, force fields are computed for neighbour interactions within a delimited cut-off range and Coulomb interactions using Particle mesh Ewald (PME). PME calculate electrostatic interactions by splitting the summation into short- and long-range parts (Essmann et al. 1995).

MD timescale is usually divided into femtoseconds (10^{-15} s), meaning that at each timestep the forces acting on each atom will be computed, and the position and velocity of each atom will be updated, resulting in an accurate description of the movement of atoms or a trajectory. MD uses a computer software to calculate the algorithms for propagation (Table 1.1). To avoid surface artifacts periodic boundary conditions are added; this avoids the interaction of the same molecule with itself (Guvench and MacKerell 2008). Thus, by knowing 1) the initial coordinates of each atom, information usually obtained from experimental data or a built-model based on

experimental data, 2) the potential obtained from a force field, and 3) the computer software that applies the algorithms for propagation, then it is possible to know the forces and their effects acting on each atom and its surroundings at almost atomic resolution and timescales (Brooks et al. 2009; Case et al. 2005; Hess et al. 2008; Phillips et al. 2005; Schlick 2010; Karplus and Kuriyan 2005; Table 1.1).

Table 1.1. Common force fields and computer software

Force fields	Software/propagation algorithm
CHARMM (Brooks et al. 2009)	CHARMM (Brooks et al. 2009)
AMBER (Case et al. 2005)	AMBER (Case et al. 2005)
GROMACS (Hess et al. 2008)	GROMACS (Hess et al. 2008)
	NAMD (Phillips et al. 2005)

1.2.3.4. Molecular Dynamics Flexible Fitting simulations

Improvements in cryo-EM atomic resolution also improved flexible fitting techniques (Trabuco et al. 2010). Flexible fitting methods consider the degrees of freedom to fit and allow atomic structure to undergo conformational changes to improve its correspondence to the density map (Trabuco et al. 2008; Trabuco et al. 2009). However, the disadvantage of flexible fitting is that interaction between subunits or different conformations cannot be determined (reviewed in Trabuco et al. 2008; Wriggers and Chacón 2001).

Molecular Dynamics Flexible Fitting (MDFF) is a method based on MD simulation and 3D cryo-EM single-particle reconstruction used to fit atomic structures into cryo-EM density maps (Trabuco et al. 2008). In the MDFF method external forces proportional to the gradient of the density map are added to a molecular dynamic simulation to drive the atoms into high-density regions while MD force fields and harmonic restraints avoid structural distortion keeping the integrity of the structure (Bernardi et al. 2016; Trabuco et al. 2009; Trabuco et al. 2008; Chapter 2). Therefore, MDFF takes into account all the information contained in the map avoiding the use of reduced representations or a single PDB structure, to fit components even when the structure is not available and because MDFF fitting is performed locally, it

considers the information of the map through a potential and only uses global measures of the fit to assess convergence but not to drive the fit, all this independently of the system size (Trabuco et al. 2008).

1.2.3.5. Targeted molecular dynamics simulations

Targeted Molecular Dynamics (Targeted MD) is a molecular dynamics technique used to calculate the transition pathways between two known structures (Schlitter, Engels, and Kruger 1994; Diaz et al. 1997). Targeted MD is a MD simulation with the addition of a constraint force or biasing force that induce conformation changes, therefore, the equation of motion is solved, and individual atoms are free to move in accordance to the force field and environmental parameters (Schlitter, Engels, and Kruger 1994).

The aim of Targeted MD is to find a representative pathway of the transition from an initial to a final conformation (x_i - x_F) at a given temperature, pressure and environment (Schlitter, Engels, and Kruger 1994). This method requires two sets of coordinates, 1) the reference structure x_i which represent the starting geometry and 2) the target structure, x_F, which provide the directing constraints (Schlitter, Engels, and Kruger 1994). Therefore, for any conformation x, the distance to the target structure is x-x_F, as at each timestep, the target structure is aligned to the current coordinates and the RMSD distance is computed between the current coordinates and the target structure (Bernardi et al. 2016). This distance and the expected distance to the target structure p, at a given time define a time-dependent holonomic constraint (that depends on the coordinates and time and not on velocities or other) $\phi(x) \equiv |x - x_F|^2 - p^2 = 0$ (Schlegel 1982; Diaz et al. 1997). After each timestep, p decreases linearly and continuously towards the target conformation x_F (Diaz et al. 1997; Schlegel 1982).

Since activation or large conformational changes occur on the timescale of microseconds or longer, conformational transitions could be difficult to observe with classic MD nanosecond time scale. Targeted MD allow to guide a subset of atoms in the simulation towards the final structure, therefore the restriction of this group coordinates will not affect the mobility and allow enough flexibility to explore

plausible pathways with respect to energy barriers (Schlitter, Engels, and Kruger 1994; Bernardi et al. 2016).

1.3. G protein-coupled receptors

G protein-coupled receptors are transmembrane receptors whose main function is to mediate cellular responses by translating the signal of extracellular messengers into intracellular signals.

In humans, GPCRs form the largest group of transmembrane proteins (Fredriksson et al. 2003; Fagerberg et al. 2010). Due to their role in physiological processes, clinical importance and accessible location in the cell membrane, GPCRs are effective drug targets; currently being the target for 35% of approved drugs (Kobilka 2007; Sriram and Insel 2018).

1.3.1. GPCRs classification

Members of GPCRs superfamily have been grouped into two main classifications: 1) A-F classification, group GPCRs into six classes based on their sequence homology and functional similarities (Kolakowski 1994; Attwood and Findlay 1994); 2) GRAFS classification, which group GPCRs based on their genomic origin into five sub-families: Glutamate (G), Rhodopsin-like (R), Adhesion-like (A, also referred as B2), Frizzled/Taste2 (F), and Secretin-like (S, also referred as B1) (Fredriksson et al. 2003; Schiöth and Fredriksson 2005) (Table 1.2).

The main difference between A-F and GRAFS classification is the division of class B into adhesion (A / B2) and secretin (S /B1) receptors due to their different evolutionary history (Kolakowski 1994; Fredriksson et al. 2003).

Table 1.2. GPCR classification

Sequence homology (A-F)	Class A rhodopsin like	Class B secretin receptor		Class C metabotropic glutamate	Class D fungal mating pheromone	Class E cAMP receptors	Class F frizzled or smoothened
Genomic origin (GRAFS)	Rhodopsin (R)	Adhesion (A / B2)	Secretin (S / B1)	Glutamate (G)	Not found in vertebrates		Frizzled / Taste 2 (F)

Kolakowski 1994; Fredriksson et al. 2003

1.3.2. GPCRs topology

Despite differences in sequence and genomic origin, GPCRs share a common topology characterised by seven transmembrane (7TM) α-helices arranged as an anti-clockwise bundle as viewed from the extracellular side, connected by less conserved three extracellular loops (ECLs), and three intracellular loops (ICLs), an extracellular N-terminus of varied length and an intracellular C-terminus (Kolakowski 1994; Attwood and Findlay 1994; Archbold et al. 2011; Bang and Choi 2015; Che et al. 2018; Cherezov et al. 2007; Choe et al. 2011; Hanson et al. 2008; Hollenstein et al. 2013; Kang et al. 2015; Kumar et al. 2011; Li et al. 2019; Miller-Gallacher et al. 2014; Palczewski et al. 2000; Park et al. 2008; Rasmussen, Choi, et al. 2011a; Rasmussen et al. 2007; Rasmussen, Choi, et al. 2011b; Rasmussen, DeVree, et al. 2011; Runge et al. 2008; Scheerer et al. 2008; Shao et al. 2016; ter Haar et al. 2010; Underwood et al. 2010; Table 1.3).

Table 1.3. GPCR Characteristics

	Class A Rhodopsin (R)	Class B Adhesion (A / B2)	Class B Secretin (S / B1)	Class C Glutamate (G)	Class F Frizzled / Taste 2 (F)
Members in humans	~700	33	15	22	Frizzled: 11 Taste 2: 25
N-termini	Short (majority)	Largest N-terminal 200 to 2800 amino acids	Long N-terminal 60 to 80 amino acids	280 to 580 amino acids	Taste: short Frizzled: 200 amino acids with conserved Cys
Type of Ligand	photons, organic odorants, amines, nucleotides, nucleosides, peptides, proteins, lipids	Antibodies, toxins, proteins	Peptide hormones Neuropeptides	Neurotransmitter and Ions (Ca^{2+}, Mg^{2+})	Wnt glycoproteins
Characteristic	- Heterogeneous TMD - TM3-ICL2 ionic lock DRY/ D(E)RY motif - NSxxNPxxY$x_{5,6}$F motif in TM7 - ~80% of GPCRs - palmitoylated cysteine at C-terminus	- Rich in glycosylation sites and proline residues (mucin-like stalks) - Adhesion-like motif located in NTD - EGF-like repeats, mucin-like regions, conserved cysteine-rich motifs	- Conserved Cys bridges important for ligand binding - Activate adenylyl cyclase and phosphatidyl-inositol-calcium pathway	- "Venus fly trap" - Cys rich ECD - Stimulate inositol phosphate/Ca^{2+} intracellular pathway - Variable C-terminal	- Wnt signalling pathway - Frizzled proteins control planar cell polarity, embryonic development, cell proliferation and formation of neural synapses

Harding et al. 2018; Chun, Zhang, and Liu 2012; Salon, Lodowski, and Palczewski 2011; Bjarnadottir et al. 2004; Paavola and Hall 2012; Dror et al. 2009; Fredriksson et al. 2003; Lagerstrom and Schioth 2008; Salon, Lodowski, and Palczewski 2011; Hu, Mai, and Chen 2017

1.3.3. GPCRs Ligands

GPCRs are activated though a wide range of ligands: small molecules, covalent chromophores, ions, peptides or proteins (Ji, Grossmann, and Ji 1998; Kobilka 2007; Gether 2000; Zhao et al. 2020). These ligands differ in structure, type and binding mechanisms according to the type and function of the receptor. Although there are exceptions, class A are usually characterised by a short N-terminus and often smaller agonist molecules bind in the transmembrane region. In contrast, Class B have longer N-terminus involved in the binding of the C-terminal part of their larger peptide ligands, although the N-terminus of the ligand interacts with the receptor's transmembrane domain (two-site binding model). Class C have a very large N-terminal domain which contains the ligand binding site for a small molecule or ion. Some ligands have the ability to bias the signal pathway depending on which conformational state they stabilise (Rankovic, Brust, and Bohn 2016; Deupi and Kobilka 2007) (Table 1.3).

1.3.3.1. Ligand-receptor interaction

In biochemistry a ligand is any molecule or atom that binds to a protein molecule or receptor. These can occur naturally and made of organic or inorganic molecules or can be made synthetically.

1.3.3.1.1. Affinity

In order to produce a response, ligands should first bind the receptor. Affinity defines the strength of the binding interaction between the ligand and receptor at a given concentration; it is characterised by the equilibrium dissociation constant (K_d) which describes the concentration of a drug required to occupy 50% of the target receptors at equilibrium (Ariens 1954; Black and Leff 1983; Stephenson 1956).

1.3.3.1.2. Efficacy

Efficacy describe the ability of ligand or drug to produce an effect on the receptor (Ariens 1954; Black and Leff 1983; Stephenson 1956). It is determined by measuring the concentration of the ligand needed to produce a response (concentration/response) and is often used as a measure of potency. The effective concentration (EC_{50}) refers to the concentration of drug that produces half maximal effect. While the transducer τ constant represents the inverse fraction of receptors occupied by the agonist to obtain half-maximal response (Black and Leff 1983).

1.3.3.1.3. Agonists - Antagonists

Based on the affinity and efficacy, ligands are classified as agonists and antagonists, with different states between. Full agonists are ligands that bind, stabilise, and fully activate a receptor, antagonists are ligands with no basal activity but are able to bind and competitively block the access to other ligands. Partial agonists are able to induce a submaximal activation at saturating concentrations,

meaning that they are less effective to shift the equilibrium towards the active state. Inverse agonists are able to bind and stabilise a receptor but without producing a response.

1.3.3.2. Allosteric ligands

Allosteric ligands are compounds able to bind to a site different from the orthosteric site, usually simultaneously to the orthosteric ligand, with the capacity to change the affinity and binding kinetics or shift the equilibrium states of the receptor (Nygaard et al. 2009; De Lean, Stadel, and Lefkowitz 1980; Kenakin 2004). Based on the duration and effect of the orthosteric ligand; allosteric ligands are classified into positive allosteric modulator (PAM) if they increase the effect, or negative allosteric modulator (NAM) if they decrease it.

Allosteric compounds are able to alter the dynamics and structure of GPCR with clinical relevancy. For example, Novo Nordisk compound 2 and other quinoxalines (Knudsen et al. 2007), flavonoids, and the Eli Lilly pyrimidine-based compounds - compound B (Koole et al. 2010; Schann et al. 2009) induce GLP-1R-mediated cAMP signalling and insulin secretion in a dose-dependent manner. This suggests PAM in GLP-1R can be used for the treatment of type 2 diabetes and obesity.

1.3.4. GPCRs activation

Usually, receptor activation starts with the binding of an agonist ligand to the receptor. This transmits the extracellular signal into an intracellular response. First it was thought that GPCRs were only able to signal through heterotrimeric G proteins and hence the name; however, β-arrestins had shown to activate intracellular G protein independent pathways (Luttrell et al. 2018; Gurevich and Gurevich 2019; Rasmussen, Choi, et al. 2011a; Mahoney and Sunahara 2016; Deupi and Kobilka 2007; Yao et al. 2009; Kim et al. 2013; Manglik et al. 2015).

1.3.4.1. Two-states model theory

The two-states model, also known as minimal two-states, propose the existence of receptors in two states: inactive (R) and active (R*) states coexisting in equilibrium (Samama et al. 1993a; Leff 1995; Del Castillo and Katz 1954). Binding of the ligand shifts the equilibrium towards the state it has a higher affinity to; depending on the ligand, this can activate (agonists) or block (antagonists) intracellular responses (Mahoney and Sunahara 2016; Deupi and Kobilka 2007; Manglik et al. 2015; Deupi and Kobilka 2010).

According to this theory, GPCRs exist in: i) Active state with high affinity to agonists and couple to an intracellular protein; or ii) inactive state with low affinity for agonists and the absence of intracellular protein. It is important to note that GPCRs are dynamic structures constant exploring energy landscapes which correspond to diverse or intermediate conformational states (Gether 2000; Rasmussen, Choi, et al. 2011b; Rasmussen, DeVree, et al. 2011).

1.3.4.2. Ternary model

The ternary complex model state that the complex formed by the receptor, agonist and G protein is responsible of translating the external stimulus, generated by the ligand binding to the receptor, into response by the intracellular protein and that receptors can spontaneously exist in the active or inactive state even if they are bound or not to a ligand (Samama et al. 1993b; Pierce, Premont, and Lefkowitz 2002; Kenakin T. 2000; Andre De Lean 1980).

1.3.4.3. Common activation features within the transmembrane domain

GPCRs activation trigger conformational changes to transmit the extracellular signal into an intracellular response. Despite differences between GPCRs families, common features during conformational changes within GPCRs transmembrane

domain and cytoplasmic end suggest similar activation mechanisms (reviewed in Kobilka 2007; Venkatakrishnan et al. 2016).

Conformational changes are seen from the extracellular side to the intracellular side towards cytoplasmic proteins (Ji, Murdoch, and Ji 1995). At the extracellular end, the inward movement of TM5, TM6 and TM7 extracellular end appear to be responsible for the binding of endogenous ligands. In the transmembrane domain, "microswitches" rotamer changes in highly conserved side chains together with hydrogen-bond networks between conserved polar residues and structural water molecules guide the movement of transmembrane helices (Nygaard et al. 2009). The characteristic outward movement of the intracellular half of TM6 that creates a cavity that accommodate intracellular proteins; however, the outward movement of TM5 and TM7 intracellular end are also observed within activation (Liang et al. 2017; Liang et al. 2018; Farrens et al. 1996). Rearrangement and rotation of TM3 contribute to GPCRs activation, as mutation in TM3 E/DRY motif in class A impairs activation. At the intracellular side, changes at ICL2, ICL3 and the C-terminal, appear to be involved in the coupling of intracellular proteins (reviewed in Gether 2000; Bockaert and Pin 1999b; Wess 1997b; Liu, Horst, et al. 2012). Nevertheless, specific activation traits in each family are expected.

1.3.5. GPCRs signalling

When activated, GPCRs transduce the outer stimulus (e.g., agonist binding) to intracellular response. Contrary to the initial idea in which GPCRs signalling was exclusively mediated by heterotrimeric G proteins, and hence the name, GPCRs are also able signal through β-arrestins. GPCRs are constantly exploring different energy landscapes, and sometimes they are able to signal in a ligand-free state, however in most cases binding of the ligand triggers conformational changes in the receptor which activate different signalling pathways: G protein or β-arrestin (Oakley et al. 2000; Oakley et al. 2001; Key et al. 2003; Pan, Gurevich, and Gurevich 2003; Shenoy and Lefkowitz 2003; Shenoy et al. 2001; Kumari et al. 2017; Cahill et al. 2017).

Superimposition of GPCRs-G protein complex with GPCRs-arrestin crystal complex structures demonstrate that G proteins and arrestins bind in the same

intracellular cavity of the receptor (Rasmussen et al. 2007; Shukla et al. 2014; Scheerer et al. 2008; Szczepek et al. 2014), however, there are differences in the receptor-intracellular protein interface (Qiao et al. 2020) and intracellular response (Carpenter et al. 2016; Kang et al. 2015; Liang et al. 2017; Rasmussen, Choi, et al. 2011b; Zhang et al. 2017). Usually G proteins activate second messengers, while β-arrestins regulate MAP kinases, receptor desensitisation and internalisation. Therefore, understanding GPCRs kinetics and characterisation of the different conformational states of the receptor raise the possibility for development therapeutics drugs able to target a specific response via a particular signalling pathway.

1.3.5.1. G protein pathway

Heterotrimeric G proteins (G proteins) are intracellular proteins composed by α, β and γ subunits. They are classified according to their α-subunit into four families: $G\alpha_{i/o}$, $G\alpha_{s}$, $G\alpha_{q/11}$ and $G\alpha_{12/13}$ (Milligan and Kostenis 2006; Neubig 1994; Koehl et al. 2018; Rasmussen, DeVree, et al. 2011).

The inactive state is characterised by the Gα-subunit bound to guanosine-5'-diphosphate (GDP) and βγ-dimer. Activation of the receptor triggers conformational changes at the Gα subunit, such that the movement of the C-terminal helix α5 away from the nucleotide binding site (Rasmussen, DeVree, et al. 2011) which disrupts the binding site and facilitates i) the exchange of GDP for guanosine-5'-triphosphate (GTP) facilitated by the higher intracellular concentration of GTP, and ii) the dissociation of the Gα subunit from Gβγ dimer to independently modulate different intracellular effectors, such as enzymes and ion channels depending on the activated Gα subunit (Table 1.4). However, the existence of mechanisms that do not require direct interaction between the receptor and the G protein, such activators of G protein signalling (AGS), suggest unknown functions of G proteins involved in the selectivity of intracellular signalling over G protein subunits (Blumer et al. 2005). To keep physiological process in equilibrium, the G protein inactive conformation needs to be restored, which terminates or attenuates GPCR signalling. The intrinsic GTPase activity of Gα, enhanced by regulator of G proteins signalling (RGS) (Vries et al. 2000) hydrolyse GTP to GDP, returning its affinity for βγ dimer, restoring the inactive

conformation and ends G protein signalling; therefore, GDP/GTP exchange rate determinates the overall rate of reaction in G protein activation (Gether 2000; Wess 1997a; Bockaert and Pin 1999a; Kleuss et al. 1994).

Besides determining the signalling pathway, the Gα subunit is involved in binding specificity to GPCRs (Conklin et al. 1993; Okashah et al. 2019a). Comparison between families A and B show that even though Class B display a bigger outward movement of helix 6, both families are able to bound G_s proteins (Lebon 2020) and contrary to the initial "barcode" of amino acids defining G protein coupling, it is the interaction between the C-terminal α-helix of the G protein a-subunit and TM3, TM5, TM6 and ICLs what defines G protein coupling specificity (García-Nafría and Tate 2019). Therefore, differences in the C-terminal end of the G protein α5 helix and interaction surface area with the receptor rather than the position of the receptor TM6 are key factors determining G protein selectivity (Hilger et al. 2019; Okashah et al. 2019b; Lebon 2020; García-Nafría and Tate 2019).

Table 1.4. Heterotrimeric G proteins and signal transduction

G protein subunit	α-subunit	Signal transduction
$G\alpha_s$	α_s, α_{olf}	Stimulates adenylyl cyclase (AC) and AC-induced PKA activation
$G\alpha_i$	α_i α_o α_t, α_z	Inhibits adenylyl cyclase Inhibition of AC and activation of MEK/ERK pathway to mediate cell cycle progression Activation of phosphodiesterase 6
$G\alpha_q$	α_q, α_{11}, α_{14}, α_{15}, α_{16}	Activates phospholipase C-β and IP3 for calcium mobilisation
$G\alpha_{12/13}$	$G\alpha_{12}$, $G\alpha_{13}$	Activates Rho kinase signalling pathways Activate small GTPase families
Gβ	-	Activate phosphatidylinositol 3-kinase
Gγ	-	Activate kinases, protein kinase D
Gβγ	-	Directly gating ion channels, activation of second messengers

reviewed in Salon, Lodowski, and Palczewski 2011; Kiselyov, Shin, and Muallem 2003; Weis and Kobilka 2018

1.3.5.2. β-arrestin pathway

Arrestins were first believed to terminate or "arrest" GPCR signalling, however they have shown they can re-direct GPCR signalling through G protein-independent pathways.

In mammals four types of arrestins have been identified and classified into visual or non-visual arrestins; i) arrestin-1 and arrestin-4, are located in the

photoreceptors, thus, only found in the of visual system, ii) arrestin-2 also called β-arrestin1, and arrestin-3 or β-arrestin2, are ubiquitously expressed and able to interact with GPCRs (Pfister et al. 1985; Murakami et al. 1993; Lohse et al. 1990; Benovic et al. 1987; Table 1.5).

Table 1.5. Mammalian arrestins

	Arrestin	Other names
Visual	arrestin-1	S antigen
	arrestin-4	X-arrestin, cone arrestin, C-arrestin
Non-visual	β-arrestin1	arrestin-2
	β-arrestin2	arrestin-3

Two models have been proposed for arrestin binding to receptors. In the first model, arrestins detect phosphorylated and activated receptors. Where it is suggested that after the dissociation of the G protein from the active receptor and phosphorylation of serine or threonine residues at C-terminal or ICLs, favoured by GPCRs kinases (GRKs), arrestins bind the activated receptor. In the second model, receptor-attached phosphates release conformational constraints in arrestin as the active receptor conformation provides a binding surface (Bologna et al. 2017; Perry and Lefkowitz 2002; Pitcher, Freedman, and Lefkowitz 1998; Krupnick and Benovic 1998). However, binding of arrestin phosphorylation-independent mutants to inactive phosphor-receptor suggest arrestin-binding elements are accessible even in the inactive receptor suggesting a more complex mechanism (reviewed in Gurevich and Gurevich 2004).

Depending on the type of receptor, arrestin and phosphorylation pattern, arrestins are able to i) signal through mitogen-activated protein kinases (MAPKs), serine/threonine kinase AKT, they tyrosine kinase SRC, nuclear factor- κB (NF-κB) and phosphoinositide 3-kinase (PI3K) pathways (Luttrell et al. 2001); or ii) internalise the receptor via clathrin heavy chain, adaptor protein 2 (AP2) complex or E3 ubiquitin ligase Mdm2; with different results (Lefkowitz 1998; Ferguson 2001; Laporte et al. 1999; Shenoy et al. 2007; Ranjan, Gupta, and Shukla 2016; Table 1.6). Experimental data show that arrestins are able to signal in G protein-independent ways (Alvarez-Curto et al. 2016; Grundmann et al. 2018).

Table 1.6. Comparison of arrestin effect in Class A and B receptors

GPCR class	Arrestin affinity	Receptor-arrestin complex during internalisation
A	Higher affinity towards β-arrestin2 than β-arrestin1 but no interaction with visual arrestin	Rapid dissociation and are quickly recycled to the cell membrane
B	Similar affinity towards β-arrestins and visual arrestins	Stable complex
Switching the c-terminal tails of Class A and B receptors reverse the affinity of Class A and B for β-arrestins and visual arrestins (Oakley et al. 2000). This shows that GPCRs signalling depends on the cellular complement		

1.3.5.3. Biased agonism

Receptors are constantly exploring different energy landscapes and adopting different conformational states, usually, binding of a ligand to the receptor shift the equilibrium and stabilise a conformational state leading to an intracellular response. Since different molecular conformations can lead to different intracellular responses according to the adopted conformation, 'biased agonism' is used to describe the ability of a ligand to activate a different signalling pathway in the same receptor, for example G protein or β-arrestins pathways (reviewed in Jarpe et al. 1998; Buchwald 2019; Shenoy and Lefkowitz 2005; Rakesh et al. 2010; Ryba et al. 2017; Luttrell et al. 1999; Jorgensen et al. 2007; Liu, Horst, et al. 2012; Lefkowitz 1998; Sprang 1997; Deupi and Kobilka 2007). It is important to notice that biased agonism is different from system bias, as system bias depends on the tissue.

Molecular mechanism of biased signalling is not fully understood, but experimental data show that receptor conformation stabilised by G protein-biased agonist differ from those stabilised by β-arrestin-biased agonists (reviewed in Nobles et al. 2011; Rankovic, Brust, and Bohn 2016) (Table 1.7 and 1.8). Since a single receptor can activate different signalling pathways, understanding the conformation preferred by the G protein or β-arrestin pathway opens the possibility of targeted modulation to select the desired effect and avoid side effects. An example of biased agonism successfully used in clinical treatment is carvedilol, a β-biased β-AR agonist that through the β-arrestin pathway increases cell survival in heart failure patients and avoids the G protein effect of catecholamines (Poole-Wilson et al. 2003).

Table 1.7. Biased agonism

Agonist	Common effect on GPCR
G protein-biased	Effect on TM6 and ICL3 movement TM5 and TM6 primarily interaction with Gα subunit
β-arrestin-biased	TM7 and helix 8 conformation required for β-arrestin recruitment

Westfield et al. 2011; Rasmussen, DeVree, et al. 2011; Rahmeh et al. 2012; Rasmussen, DeVree, et al. 2011; Westfield et al. 2011

Table 1.8. Examples of different signalling pathways

GPCR	G protein pathway	β-arrestin pathway
Histamine H4 receptor (H_4R)	Immune and inflammatory processes	Inhibition of pruritus and skin inflammation
Angiotensin II type 1 receptor (AT1R)	Potent vasoconstriction and increase blood pressure	Modified peptide angiotensin II (SII) Cytoprotection and antiapoptotic effects
μ-opioid receptor (MOR)	Analgesic effect	Constipation, respiratory suppression, tolerance and dependence

Leurs et al. 2009; Thompson et al. 2015; Ryba et al. 2017; Ahn et al. 2009; Wang et al. 2017

1.4. Class B GPCRs

Secretin-like family, also called Class B, is made of 15 members (Figure 1.2) which are endogenously activated by peptide hormones. Class B receptors are attractive drug targets in human diseases; they regulate exocrine and endocrine secretion, metabolism, growth, feeding behaviour and neuro- and immune modulation as they are activated by peptide hormones (Table 1.9). Understanding of Class B receptors activation mechanisms could allow the design of synthetic drugs to be applied to modulate Class B GPCRs activity.

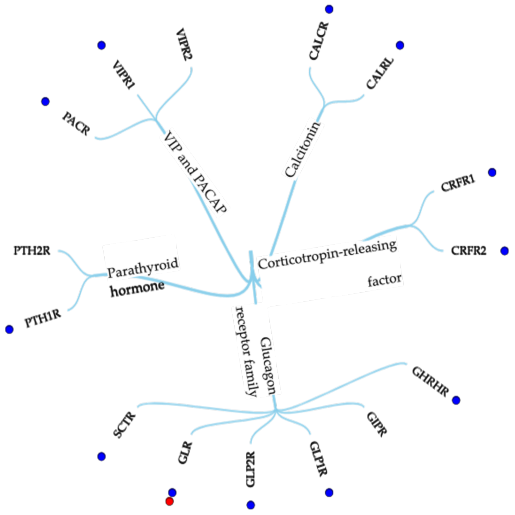


Figure 1.2. Phylogeny of Class B receptors
Image from GPCRdb in 2021: integrating GPCR sequence, structure and function. Kooistra AJ, Mordalski S, Pándy-Szekeres G, Esguerra M, Mamyrbekov A, Munk C, Keserű GM, Gloriam DE *Nucleic Acid Reserarch,* 2020, 49:D335-D343.

TABLE 1.9. CLASS B RECEPTORS, ENDOGENOUS PEPTIDES AND PATHOLOGIES INVOLVED

Secretin family receptors	**Peptide agonist**	**Pathology**
Calcitonin (CTR)	Calcitonin (CT)	Osteoporosis
Calcitonin-like (CALCRL)	Calcitonin gene-related peptide	Migraine
Corticotropin-releasing hormone (CRHR1, CRHR2)	Corticotropin-releasing factor (CRF)	Chronic stress, depression, anxiety
Glucagon receptor (GCGR)	Glucagon (GCG)	Diabetes, obesity
Gastric inhibitory polypeptide receptor (GIPR)	Gastric inhibitory polypeptide (GIP)/ glucose-dependent insulinotropic polypeptide (GIP)	Diabetes, obesity
Glucagon-like peptide receptors (GLP-1R, GLP-2R)	Glucagon-like peptide 1 & 2 (GLP-1, GLP-2)	Diabetes, obesity and GLP-2 for short bowel syndrome
Growth-hormone-releasing hormone receptor (GHRHR)	Growth hormone-releasing factor (GRF)	Dwarfism
Adenylate cyclase activating polypeptide receptor (PAC1/ADCYAP1R1)	Pituitary adenylate cyclase-activating peptide (PACAP21)	Neurodegenerative disease
Parathyroid hormone receptors (PTHR1, PTHR2)	Parathyroid hormone (PTH)	Osteoporosis
Secretin receptor (SCTR)	Secretin (SCT)	
Vasoactive intestinal peptide receptors (VIPR1, VIPR2)	Vasoactive intestinal peptide (VIP)	Inflammatory diseases

Harmar 2001; Archbold et al. 2011; Laburthe et al. 1996; Gomariz et al. 2001; Mulder, Kolatkar, and LeBoff 2006; Brenneman 2007; Campbell, Bongers, and Felix 1995; Gilligan and Li 2004; Mulder, Kolatkar, and LeBoff 2006; Lars and Karin 2013; Jeppesen 2006; Joo et al. 2017

1.4.1. Structural features of Class B receptor

Before 2017, only isolated domains of Class B were available. However, by 2021, structures of twelve different class B GPCRs (full-length or transmembrane) became available (Table 1.10).

Even though all receptors belong to the same family and share an overall architecture there are differences between states and receptors, such as the position of the N-terminal relative to the TMD, loops conformation the position of the ligand and individual side chains.

Since most Class B structures have been solved with a ligand agonist, activation mechanisms leading to activation may differ, as physical studies suggesting that partial agonist stabilise distinct states different to the one stabilised by the native ligand (Gregorio et al. 2017).

Table 1.10. Available Class B receptors (full-length or transmembrane domain)

Receptor	State	PDB code	Structure	Resolution (Å)	Method	year
GLP-1R	active	5NX2	Thermostabilised full-length GLP-1R in complex with a peptide agonist	3.7	X-ray diffraction	2017
GLP-1R	active	5VAI	Full-length rabbit peptide-activated GLP-1R-Gs complex	4.1	cryo-EM	2017
GLP-1R	active	6B3J	Biased agonist-bound human GLP-1R-Gs complex	3.3	cryo-EM	2018
GLP-1R	inactive	5VEW	Human GLP-1R TMD in complex with NAM (PF-06372222)	2.7	X-ray diffraction	2017
GLP-1R	inactive	5VEX	Human GLP-1R TMD in complex with NAM (NNC0640)	3.0	X-ray diffraction	2017
GLP-1R	active	6ORV	Non-peptide agonist (TT-OAD2) bound to GLP-1R	3.0	Cryo-EM	2020
CGRPR	active	6E3Y	Human CGRPR-CGRP-Gs complex and RAMP1	3.3	cryo-EM	2018
GCGR	inactive	5YQZ	Full-length human GCGR complex with GCG partial agonist analogue (NNC1702)	3.0	X-ray diffraction	2018
GCGR	inactive	5XEZ	Full-length GCGR in complex with NAM and antigen-binding fragment of an inhibitory antibody	3.0	X-ray diffraction	2017
GCGR	inactive	5XF1	Full-length GCGR in complex with NAM (NNC0640) and antigen-binding fragment of an inhibitory antibody	3.19	X-ray diffraction	2017
GCGR	inactive	4L6R	Transmembrane domain of a GCGR	3.4	X-ray diffraction	2013
PTH1R	inactive	6FJ3	Human PTH1R in complex with a peptide agonist	2.5	X-ray diffraction	2018
PTH1R	active	6NBF	Human PTH1R bound to long-acting PTH analogue and Gs complex	3.0	cryo-EM	2019
PTH1R	active	6NBH	Human PTH1R bound to long-acting PTH analogue and Gs complex	3.5	cryo-EM	2019
PTH1R	active	6NBI	Human PTH1R bound to long-acting PTH analogue and Gs complex	4.0	cryo-EM	2019
CTR	active	5UZ7	Full-length CTR -peptide ligand complex with Gs	4.1	cryo-EM	2017
CTR	active	6NIY	Alternative modelling of original data in 5UZ7	3.34	cryo-EM	2019
CRF1R	inactive	4K5Y	TMD of human CRF1R in complex with antagonist (CP-376395)	2.98	X-ray diffraction	2013
CRF1R	inactive	4Z9G	TMD of human CRF1R in complex with antagonist (CP-376395)	3.183	X-ray diffraction	2015
CRF1R	active	6PB0	Urocortin 1-bound CRF1 receptor in complex with Gs protein and Nb35	3.00	cryo-EM	2020
CGRP	inactive	7KNU	CGRP receptor with bound CGRP peptide in a detergent micelle	3.49	X-ray diffraction	2013
CGRP	inactive	7KNT	apo-CGRP receptor in a detergent micelle	3.15	cryo-EM	2021
GLP-1R	active	7LCK	PF 06882961 bound GLP-1R	3.24	cryo-EM	2021
GLP-1R	active	7LCJ	PF 06882961 bound to GLP-1R:Gs complex	2.82	cryo-EM	2021
GLP-1R	active	7LCI	PF 06882961 bound to GLP-1R:Gs complex	2.90	cryo-EM	2021
GLP-2R	active	7D68	Human GLP-2 receptor-Gs protein complex	3.00	cryo-EM	2020
GHRH	active	7CZ5	Human growth hormone-releasing hormone receptor-Gs protein complex	2.60	cryo-EM	2020
GLP-1R	active	6XOX	Human GLP-1R bound to non-peptide agonist LY3502970	3.10	cryo-EM	2020
SECR	active	7D3S	Human SECR in complex with an engineered Gs heterotrimer	2.90	cryo-EM	2020
GLP-1R	active	6X18	GLP-1 peptide hormone bound to GLP-1R	2.10	cryo-EM	2020
GLP-1R	active	6X19	Non peptide agonist CHU-128 bound to GLP-1R	2.10	cryo-EM	2020
GLP-1R	active	6X1A	Non peptide agonist PF-06882961 bound to GLP-1R	2.50	cryo-EM	2020
VIP1	active	6VN7	VIP1 receptor-G protein complex	3.20	cryo-EM	2020
GLP-1R	active	7C2E	GLP-1R-Gs complex structure with a small molecule full agonist	4.20	cryo-EM	2020
GCGR	active	6WPW	GCGR-Gs signalling complex bound to a designed glucagon derivative	3.10	cryo-EM	2020
SECR	active	6WZG	Human secretin receptor Gs complex	2.30	cryo-EM	2020
SECR	active	6WI9	Human secretin receptor Gs complex	4.30	cryo-EM	2020
SECR	active	7DS3	Human secretin receptor coupled to an engineered heterotrimeric G protein	2.90	cryo-EM	2020
SECR	Active	6WI9	Human secretin receptor Gs complex	4.90	cryo-EM	2020
GLP-1R	active	6VCB	GLP-1R in complex with G protein GLP-1 peptide and a PAM	3.30	cryo-EM	2020
GCGR	active	6WHC	GCGR with a dual-agonist peptide	3.40	cryo-EM	2020
AM2R	active	6UUS	Adrenomedullin 2 receptor G protein complex with adrenomedullin peptide	2.40	cryo-EM	2020
AM2R	active	6UVA	Adrenomedullin 2 receptor G protein complex with adrenomedullin 2 peptide	2.30	cryo-EM	2020
GCGR	active	6LML	Human GCGR in complex with Gi1	3.90	cryo-EM	2020
GCGR	active	6LMK	Human GCGR in complex with Gs	3.70	cryo-EM	2020
AM1R	active	6UUN	Adrenomedullin 1 receptor G protein complex with adrenomedullin peptide	3.00	cryo-EM	2020
GLP-1R	active	6LN2	Full-length human GLP1R in complex with Fab fragment (Fab7F38)	3.20	X-ray diffraction	2020
PAC1R	active	6LPB	Human PAC1 receptor coupled to an engineered heterotrimeric G protein	3.90	cryo-EM	2020
PAC1R	active	6M1I	Human PAC1 receptor in complex with PACAP38	3.50	cryo-EM	2020
PAC1R	active	6M1H	Human PAC1 receptor in complex with maxadilan	3.60	cryo-EM	2020
CRF1R	active	6PB1	Urocortin 1-bound Corticotropin-releasing factor 2 receptor in complex with Gs protein and Nb35	2.80	cryo-EM	2020
CRF1R	active	6P9X	CRF1 Receptor Gs GPCR protein complex with CRF1 peptide	2.91	cryo-EM	2020
CRF1R	active	6P9Y	PAC1 GPCR Receptor complex	3.01	cryo-EM	2020
GLP-1R	inactive	6KJV	Thermal-stabilised(M9) human GLP-1R TMD	2.80	X-ray diffraction	2019
GLP-1R	inactive	6KK1	Thermal-stabilised(M8) human GLP-1R TMD	2.80	X-ray diffraction	2019
GLP-1R	inactive	6KK7	Thermal-stabilised(M6) human GLP-1R TMD	3.10	X-ray diffraction	2019
GCGR	inactive	5EE7	Human GCGR in complex with the antagonist MK-0893	2.50	X-ray diffraction	2016

1.4.1.1. NTD and stalk, extracellular structures

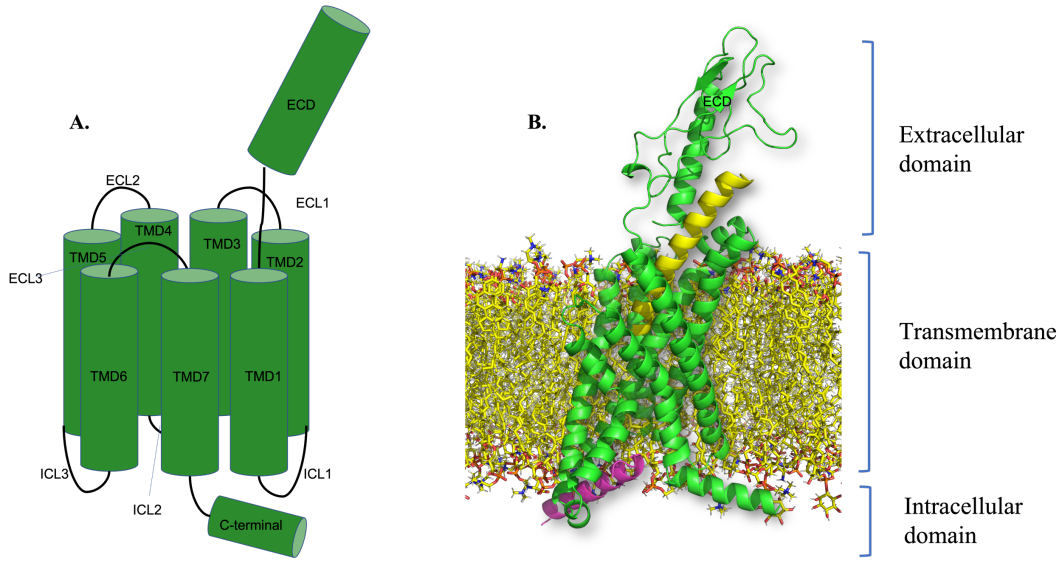
The N-terminal domain (NTD) is a conserved structure in Class B located in the extracellular area where it participates in ligand binding (Figure 1.3). Available full-length Class B structures in complex with different ligands show a similar orientation of the NTD relative to the TMD (Hilger et al. 2019; Zhang et al. 2018) supporting the role of the NTD in the two-steps activation model as it is driven by interactions between the ligand's N-terminal and the receptor. Although PTH_1, CRF_1, $VPAC_1$ and GCG receptors (Cegla et al. 2017) can interact with receptor activity-modifying proteins (RAMPs) influencing trafficking and signalling, only the calcitonin (CT) receptor and calcitonin-like (CLR) affects ligand binding (Hay and Pioszak 2016), as the RAMP-calcitonin gene-related peptide receptor (GGRPR) shows a different orientation between the ECD and peptides not compatible with the rest of Class B receptors (Liang et al. 2018; Liang et al. 2020).

The conserved NTD is formed by an α-β-β-α motif: N-terminal α-helix, two anti-parallel β-sheets (β1-β2 and β3-β4) and an α-helical stabilised by three disulphide bonds and six conserved residues (Asp, Trp, Pro, Arg/Lys, Gly, Trp) (Pioszak and Xu 2008; Runge et al. 2008; Underwood et al. 2010; Kumar et al. 2011; ter Haar et al. 2010; Sun et al. 2007). The first disulphide bond links the N-terminal α-helix with β1-β2 sheet, the second connects the two β-sheets with each other and the third connects the β3-β4 sheet with the C-terminal (Pioszak and Xu 2008; Runge et al. 2008; Underwood et al. 2010; Kumar et al. 2011; ter Haar et al. 2010; Sun et al. 2007). Conserved residues play an important role in ligand binding and receptor-ligand contact sites (Pioszak and Xu 2008; Perrin et al. 2007). Loop 2 in the NTD is stabilised by Asp-Trp interaction and facilitates ligand binding; mutation of Asp destabilises the structure impairing GHRH function causing transmitted dwarfism in mice (Lin et al. 1993; Perrin et al. 2007), while mutation of Asp and Arg in PAC1 receptor alters ligand-binding properties (Sun et al. 2007). In Class B, a fourth disulphide bond between ECL2-TM3 (C^{ECL2}-$C^{3.29b}$, class B GPCR Wootten numbering will be used; Wootten, Simms, et al. 2013) is suggested to stabilise the receptor transmembrane fold (Siu et al. 2013).

An extension of TM1 over the extracellular membrane denominated stalk links the NTD with the TMD. The exact role of the stalk is debated as it is not always observable in Class B structures. However, due to its flexibility and dynamism it is

suggested to act as a pivot coordinating the dynamics between ECD and TMD and to participate in ligand binding (Zhang et al. 2018; Sun et al. 2007; Yang et al. 2015b). In the glucagon receptor it is seen as a α-helical helix but its flexibility suggest it probably adopts the shape of a loop in structures where it hasn't been modelled (Jazayeri et al. 2017; Qiao et al. 2020; Zhang et al. 2018; Hilger, Kumar, Hu, Pedersen, O'Brien, Giehm, Jennings, Eskici, Inoue, and Lerch 2020). Comparison between active CTR and GLP-1R structures with inactive GCGR suggest the stalk facilitate TM1 shifts towards TM7 (Jazayeri et al. 2017; Liang et al. 2017; Zhang et al. 2017).

Figure 1.3. **A.** Diagram representation of a Class B GPCR showing the seven helices transmembrane domain and loops **B.** Cartoon representation of a Class B GPCR (green) in complex with ligand (yellow) and G protein *α*5 helix (pink) in the active state embedded into a lipid bilayer (stick representation).



1.4.1.2. Transmembrane domain

Conformational changes translate extracellular signal through the seven transmembrane domain (TMD) to produce an intracellular response (reviewed in de Graaf et al. 2016). Despite sharing a conserved gross seven-helix architecture, there are differences between families' TMD. For example, in order to accommodate

peptide ligands, Class B show a wider and deeper binding pocket when compared to class A (Siu et al. 2013).

Class B conserved clusters include the polar HETx motif ($H^{2.50b}$, $E^{3.50b}$, $T^{6.42b}$, $Y^{7.57b}$), intracellular polar network ($R^{2.46b}$, $R/K^{6.37b}$, $N^{7.61b}$, $E^{8.41b}$) (Liang et al. 2018; Wootten et al. 2016), GWGxP in TM4 (Hollenstein et al. 2013), $FQG^{7.50}$, $Y^{7.57}CF$, and the $P^{6.47}xxG^{6.50}$ in helix 6 (reviewed in de Graaf et al. 2017; Siu et al. 2013) and intrahelical $N^{5.50b}$, $P^{6.47b}$, $G^{6.50b}$, $H^{6.52b}$, and $Q^{7.49b}$ motif. Mutagenesis shows that residues belonging to these motifs and clusters are involved in conformational stabilisation, conservation of the inactive state and activation ($T^{6.42b}$ and $H^{2.50b}$) (Zhao et al. 2019; Yin et al. 2017; Edward Zhou, Melcher, and Eric Xu 2019; Schipani, Kruse, and Juppner 1995) and signalling (Qiao et al. 2020; Hollenstein et al. 2013).

Structural data show conformational changes in the TMD are ligand-dependent (Zhang et al. 2018); as the ligand side chains have the ability to modify the orthosteric binding pocket, and it appears to be that the occupancy of the N-terminal end of the ligand in the TMD pocket potentially provides structural basis for a higher potency resulting in a sustained response (Zhao et al. 2019). Within activation the most evident conformational change is the outward movement of the intracellular end of helix 6. Structural data show residues before the N-terminal helical end of Class B ligands (GLP-1, CGRP or PTH analogue LA-PTH) pointing towards TM6 and potentially pushing the C-terminal half of TM6 contributing to TM6 unwinding (Zhao et al. 2019).

Class B is characterised by the formation of a sharp kink in the middle region of TM6 with the partial unwinding of the α-helix at the intracellular end (Hilger et al. 2019; Zhang, Qiao, et al. 2017; Liang et al. 2017; Liang et al. 2018; Zhao et al. 2019; Pedersen et al. 2006). TM6 outward movement, along with a more discrete outward displacement of TM5 and TM7, create an inner cavity for intracellular protein binding (G protein or arrestins) (Hilger et al. 2019; Zhang et al. 2017; Jazayeri et al. 2017; Liang et al. 2018). This intracellular cavity formed by polar residues is involved in G protein signalling (Qiao et al. 2020; Hollenstein et al. 2013). Compared with Class A, Class B can remain activated for longer, probably due to higher energy barriers between states (Dror et al. 2011; Hilger et al. 2019; Qiao et al. 2020; Zhao et al. 2019).

Unlike Class A, Class B show a larger TM6 outward movement creating a sharper kink in its middle (Lebon 2020; Hilger, Kumar, Hu, Pedersen, O'Brien, Giehm, Jennings, Eskici, Inoue, and Lerch 2020; Qiao et al. 2020; Rose et al. 2014).

It was proposed that the resulting position from TM6 displacement was a factor determining the specificity and selectivity of the receptor towards G protein (Krishna Kumar et al. 2019; Kang et al. 2018; Koehl et al. 2018; García-Nafría and Tate 2019; Maeda et al. 2019; García-Nafría et al. 2018), however, these appear to be independent from TM6 position. G protein specificity and selectivity appear to be dependent on the contact area surface within the intracellular pocket; i) comparison between families A and B show different TM6 positions but still both families are able to bind G_s protein (Qiao et al. 2020; Rose et al. 2014); ii) intracellular cavities in Class B glucagon receptor have the same specificity for Gαs and $G\alpha_I$ but with different efficiencies according to the contact surface; iii) reducing the size of the intracellular pocket affects the binding of the G_s α5 C-terminal (Hilger et al. 2019; Qiao et al. 2020).

Despite differences in the orthosteric pocket and TMD conformations, a common ligand-dependent activation mechanism is seen in the Class B TMD region. These include: i) TM7 extracellular end movement towards TM6 as TM7 intracellular end rotates away from the intracellular end of TM6 (TM7 extracellular end) (Liang, et al. 2018; Zhang et al. 2017; Hilger et al. 2019), ii) TM5 shift towards TM6 followed by iii) helix 3 upward movement which remove steric restraints from residues in helix 3 that allows iv) the rearrangement of $P^{6.47}xxG^{6.50}$ motif in helix 6 stabilised by neighbouring conserved residues $H^{6.52b}$, $Q^{7.49}$ and $N^{5.50b}$ leading to v) the rotation of $T^{6.42b}$ and formation of TM6 sharp kink and outward movement (Hilger et al. 2019; Hilger et al. 2020; Qiao et al. 2020).

In receptors from the glucagon family, the starting movement of TM7 towards TM6 is the result from a polar interaction between the ligand N-terminus with $R^{7.35}$ (TM7 extracellular end), still more data is needed (Liang et al. 2018; Zhang et al. 2017; Hilger et al. 2019).

1.4.1.3. Loops

Loops are the most diverse structure in Class B1 architecture, still they play an important role in receptor-ligand interactions and coordination of conformational changes relative to TMD and intracellular proteins (Siu et al. 2013; Zhang, Qiao, et al. 2017; Yang et al. 2015a).

Function and architecture of each loop varies within Class B receptors, while CTRs receptor have the shortest ECL1, and PTH1 have the longest ECL1, which certainly have different effects in the ligand binding and activation. Mutagenesis data supports the role of ECL1 in glucagon receptor for ligand recognition and coordination of conformational changes (Siu et al. 2013; Zhang, Qiao, et al. 2017; Yang et al. 2015a), but it is not the same for GLP-1R (Dods and Donnelly 2015).

The loops conformation depends on the ligand; in the recent glucagon receptor structure (PDB 6WPW) ECL2 and ECL3 position and conformation changes according to the ligand bound (Hilger et al. 2019; Qiao et al. 2020). Mutagenesis data show that ECL2 is involved in peptide ligand binding and coordination of conformational changes (Siu et al. 2013; Zhang, Qiao, et al. 2017; Yang et al. 2015a). Which is not surprising as ECL2 connects extracellular ends of TM4-TM5, while ECL3 connects TM6-TM7, helices involved in activation and suggesting coordinated conformational changes.

Intracellular loops (ICLs) directly interact with intracellular proteins and due to their interaction interface ICLs are able to modulate these interactions (Bavec et al. 2003; Mathi et al. 1997; Conner et al. 2006; Iida-Klein et al. 1997; Cypess et al. 1999; Hilger et al. 2019; Qiao et al. 2020). The intracellular region appears to be more consistent within GPCR families. In families A and B, ICL3 and ICL2, selectively stimulates G_s or G_i (Hallbrink et al. 2001; Hilger et al. 2019). In the GCGR-G protein complex ICL2 of the receptor favours the binding in the same binding pocket of G_s over G_i due to a wider interaction with αN helix, β1 strand and α5 helix of G_s resulting in a lower efficacy of G_i coupling but also exposing the selective role of ICLs, as ICL1 and ICL3 favour G_i binding (Hilger et al. 2019; Lebon 2020). However, coupling mechanisms differ between families A and B; in GCGR ICL2 has weaker interactions with the hydrophobic pocket (Hilger, Kumar, Hu, Pedersen, O'Brien, Giehm, Jennings, Eskici, Inoue, and Lerch 2020).

1.4.2. Class B endogenous ligands

Class B native endogenous ligands are peptide hormones involved in physiological and pathological conditions (Table 1.9). Understanding of ligands

structure, binding and activation of the receptor would improve the synthesis of new and synthetic ligands with clinical significance.

Peptide ligands share a similar structure and binding mechanism. The ligand has an amino-terminal (N-terminal) end and a carboxyl-terminal (C-terminal) end. Mutagenesis studies show that residues at the N-terminal end are critical for efficacy and receptor activation, while the C-terminal is involved in binding and mutation of residues in the C-terminal can alter binding or potency (Montrose-Rafizadeh, Egan, and Roth 1994; Göke et al. 1993; Turner, Jones, and Bylund 1986). N-terminally truncated peptides have decreased biological activity, but they are able to bind their receptor and often used as antagonists; and vice versa, C-terminally truncated peptides have decreased affinity, which affect biological activity although maintaining some of its efficacy (Turner, Jones, and Bylund 1986; Ariens 1954; Couvineau et al. 1996). Changes in the biological activity of ligands as residues are removed or mutated appear to be the result of removal or lack of interaction with residues in the receptor, rather than to the destabilisation of the peptide.

Class B ligands are disordered structures in aqueous solution but adopt an α-helical conformation when binding to the receptor (Runge et al. 2008; Underwood et al. 2010; Parthier et al. 2009; Sun et al. 2007; Pioszak and Xu 2008; Grace et al. 2010; Inooka et al. 2001). The propensity of hormone ligands to adopt a helical conformation is critical for binding and activation of the receptor. The amphipathic nature of the ligand enables the first-step, in which hydrophobic contacts allows the binding of the ligand to the NTD at the same time, these confer affinity and selectivity towards the receptor ECD (Parthier et al. 2009). Then, the α-helical conformation facilitates the second-step as it shortens the length of the ligand while maintaining the stability of the ligand and positioning N-terminal residues of the ligand into the TMD active site (Parthier et al. 2009).

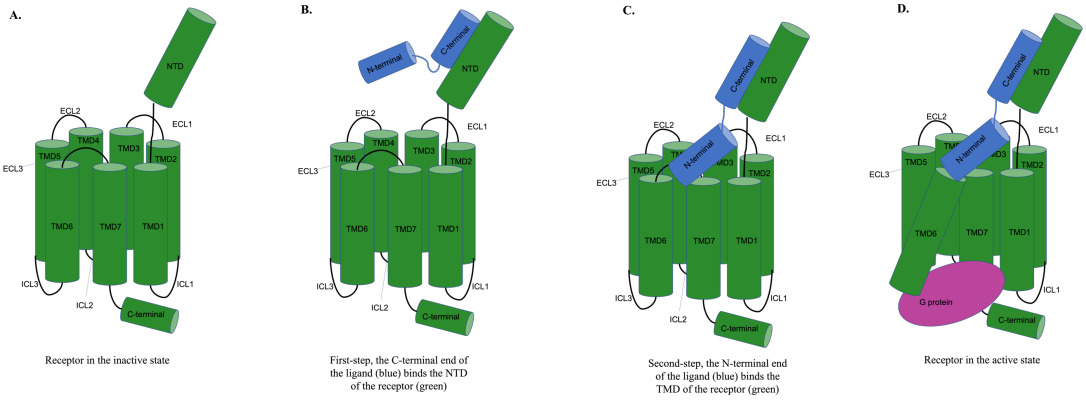
1.4.3. Class B activation

1.4.3.1. Two-steps activation model

The two-steps binding model has been used to describe peptide hormones binding to Class B receptors. This model has been supported by structural data, Föster

Resonance Energy Transfer (FRET) experiments and mutagenesis (Vilardaga et al. 2011; Bergwitz et al. 1996; Unson 2002; Grace et al. 2010; Pioszak and Xu 2008; Underwood et al. 2010; Donnelly 2012; Hoare 2005). During the first step the C-terminal end of the ligand engages with the N-terminal end of the receptor, allowing it to accommodate the N-terminus of the ligand towards the transmembrane area of the receptor (Figure 1.4). During the second step the N-terminal end of the ligand binds the TMD of the receptor leading to receptor activation (Bergwitz et al. 1996; Castro et al. 2005; Hoare, Gardella, and Usdin 2001). Chimera studies show the first step occurs quickly and is followed by a slower second step, which induce conformational changes and is essential for receptor activation and intracellular signalling (Vilardaga et al. 2011; Bergwitz et al. 1996; Unson 2002; Grace et al. 2010; Pioszak and Xu 2008; Underwood et al. 2010).

Figure 1.4. Diagram representation the two-steps activation model. **A.** Inactive class B receptor. **B.** During the first step the C-terminal end of the ligand (blue) binds the NTD of the receptor (green). **C.** During the second step, the N-terminal end of the ligand (blue) binds the TMD of the receptor (green). **D.** Binding of the ligand activates the receptor, creating an intracellular cavity for the G protein (pink).



1.4.3.2. Receptor activity-modifying proteins

Receptor Activity-Modifying Proteins (RAMPs) are proteins sharing 30% homology to the receptor; as they span the membrane, they have an extracellular N-terminus and a cytoplasmic C terminus which allows them to participate in ligand selectivity, binding, receptor trafficking and coupling to G proteins (Parameswaran and Spielman 2006; Hay and Pioszak 2016; Sexton et al. 2001; Wootten, Lindmark, et al. 2013; Christopoulos et al. 2003). RAMPs have been observed to participate in the activation of calcitonin receptor (CTR), calcitonin-like receptor (CRLR), corticotropin-releasing factor receptor type 1 (CRF1R), vasoactive intestinal

polypeptide type-1 and type-2 receptor (VPAC1R, VPAC2R), glucagon receptor (GCGR), parathyroid hormone 1 and 2 receptor (PTH_1R, PTH_2R).

1.5. Glucagon-like peptide-1 receptor and GLP-1 ligand

1.5.1. GLP-1 Ligand

Glucagon-like peptide-1 (GLP-1) is a key hormone in glucose homeostasis as it promotes insulin secretion while decreasing hepatic glucose production, among other metabolic benefits (Deacon et al. 1995; Michael J. Perley 1967).

1.5.1.1. Physiological role of GLP-1

GLP-1 is an incretin, a hormone secreted into circulation from intestinal L-cells in response to food intake (Orskov et al. 1986; Orskov, Wettergren, and Holst 1993; Creutzfeldt et al. 1996; Kreymann et al. 1987; Nauck et al. 1986; Michael J. Perley 1967).

In healthy individuals, GLP-1 is responsible for almost 70% of glucose-dependent insulin secretion from pancreatic β-cells while inhibiting glucagon secretion from pancreatic α-cell (Orskov et al. 1986; Orskov, Wettergren, and Holst 1993; Creutzfeldt et al. 1996; Kreymann et al. 1987).

GLP-1 hypoglycaemic effects are glucose-dependent, in physiological conditions GLP-1 effects only take place when glucose levels are above physiological values, avoiding hypoglycaemia. Besides, GLP-1 has protective effects towards pancreatic β-cells as it decreases β-cell apoptosis, maintaining mass and function (reviewed in Holz and Chepurny 2005). Extra-pancreatic effects of GLP-1 include increase satiety, slowing down gastric motility and weight loss (Baggio and Drucker 2014), cardiovascular (Drucker 2016; Liu et al. 2018; Iorga et al. 2020) and

renoprotection/improvement preventing the onset of macroalbuminuria and slowing glomerular filtration rate decline (Hare et al. 2010; Asmar, Holst, and Obesity 2010; Górriz et al. 2020) without increasing the risk for hypoglycemia. *In vitro* and *in vivo* studies have shown that GLP-1 restores normal cell functions by reducing endoplasmic reticulum stress, stimulating anti-inflammatory signalling, and influencing neuroprotective pathways (Rowlands et al. 2018). Experimentally, GLP-1 agonists have shown to reduce lung fibrosis (Fandiño et al. 2020) and mediate protective effect of dipeptidyl peptidase -IV (DPP-4) avoiding pulmonary hypertension (Wang et al. 2019).

1.5.1.2. GLP-1 production and half-life

GLP-1 peptide is produced in the intestinal L-cells after the proteolytic cleavage of the pre-proglucagon protein by prohormone convertase (PC) 1/3 (Holst et al. 1987; Zhu et al. 2002; Ugleholdt et al. 2004). This process results in different peptides including the 30 or 31 amino acids GLP-1 peptides: GLP-$1_{7\text{-}36amide}$ and GLP-$1_{7\text{-}37}$ respectively, both with similar insulinotropic activity in GLP-1R (Holst et al. 1987; Mojsov, Weir, and Habener 1987). Although both truncated forms may be in present in plasma in similar concentrations (Orskov et al. 1994), the amidated form GLP-$1_{7\text{-}36amide}$ possess an arginine at the C-terminal which makes it slightly more stable in plasma (Wettergren et al. 1998), explaining why it is the most abundant GLP-1 in humans (Orskov et al. 1994) and therefore, will be referred as GLP-1 ligand, except otherwise stated.

GLP-1 has short half-life (< 2 mins) as consequence of the cleavage action of DPP-4 (Deacon et al. 1995). DPP-4 is a circulating protease that cleaves the alanine residue from the GLP-1 N-terminus, resulting in GLP-1(9-36)NH_2, the so-called inactive metabolite of GLP-1(7-36)NH_2 found in circulation almost five times more than GLP-1(7-36)NH_2 (Egan et al. 2002).

It is unclear the insulinotropic action of GLP-1(9-36)NH_2. Truncation of H7* (* will be used to identify residues belonging to the ligand) and A8* create a partial agonist with 94-fold reduced affinity, but still able to produce, in lesser amount, cAMP accumulation (Montrose-Rafizadeh, Egan, and Roth 1994; Egan et al. 2002).

Contrary, some studies show no activation of pancreatic GLP-1R by GLP-1(9-36)NH_2 (Vahl et al. 2003), but with cytoprotective and anti-oxidant actions by GLP-1R independent-mechanism (reviewed in Tomas, Stanojevic, and Habener 2011; Deacon et al. 1995). In mechanisms not totally understood, the so called metabolically-inactive GLP-1(9-36)amide protects against diastolic dysfunction and reduce inflammation on post-myocardial infarction (MI), may cause of post-MI remodelling (Robinson et al. 2016).

Clearance of remaining GLP-1 is done in the kidney by neutral endopeptidase (NEP24.11). Interestingly, the C-terminal nonapeptide (FIAWLVKGR amide) derived from the cleavage of GLP-1 by NEP24.11 can suppress glucose production and oxidative stress with potential application in the treatment for fasting hyperglycaemia and metabolic syndrome (Hupe-Sodmann et al. 1995; Plamboeck et al. 2005).

Table 1.11. Products of Proglucagon

Proglucagon		GRPP	GCG	IP-1	GLP-1	IP-2	GLP-2
Tissue	Prohormone convertase (PC) hormone	Product (residue)					
Pancreas	PC2 (PCSK2)	GRPP (1-30)	GCG (33-61)	IP1 (64-72)	Major proglucagon fragment (72-158)		
GI and brain	PC1/3 (PCS12)	Glicentin (1-69)			GLP-1 (78-107/108)	IP-2 (111-123)	GLP-2 (126-158)
GI and brain	PC1/3 (PCS12)	-	Oxyntomodulin (33-69)		-		

GCG: glucagon, GI: gastrointestinal track, GRPD: glicentin-related pancreatic polypeptide, IP-1: intervening peptide 1, IP-2: intervening peptide 2

1.5.1.3. GLP-1 hormone structure

GLP-1 structure is medium-dependent (Andersen et al. 2002). In aqueous environment GLP-1 has a disordered structure that after binding and around lipid micelles, becomes helical. However this is not a continuous helix, as NMR structures show a linker region around residue Gly22* (Thornton and Gorenstein 1994; Parker

et al. 1998). The helicity and the medium are suggested to the ligand activity (Wang et al. 2011).

As a Class B peptide, GLP-1 is formed by a N-terminal end responsible for most of the peptide potency, and a C-terminal end involved in binding to the receptor. Indeed, removal of the first 14 residues produced a peptide with antagonist properties; conversely, removal of the last 11 residues at the C-terminal show no binding to the GLP-1R and therefore, no activity (Gallwitz et al. 1990).

Mutagenesis and alanine scans show special importance of residues His7*, Ala8*, Gly10*, Phe12*, Thr13* and Asp15* at the N-terminal and Phe28* and Ile29* at the C terminal (Adelhorst et al. 1994b; Montrose-Rafizadeh, Egan, and Roth 1994). The sidechains of these residues located at the N-terminal directly interact with the receptor's TMD and are critical for ligands potency; besides, being highly conserved in the growth hormone-releasing factor (GRF) superfamily (Adelhorst et al. 1994b). Residue at position 7 is critical in GLP-1 insulinotropic activity, the deletion of this residue, GLP-1(8-37), elicited no activity (Donnelly 2012; Mojsov 1992) while its mutation to alanine H7A had up to 374-fold affinity reduction (Gallwitz et al. 1994). However, analogues with aromatic tyrosine or phenylalanine instead of the native are able to activate the receptor and increase cAMP accumulation, although with a 10-fold affinity reduction in human GLP-1R (Xiao et al. 2001), suggesting a critical interaction and the importance of the aromatic ring. Although not critical, replacement of Glu9* to alanine reduced affinity and potency (Adelhorst et al. 1994b; Xiao et al. 2001). Similarly, to His7*, Gly10*, Asp15* and Ser17* are necessary for cAMP production and GLP-1 insulinotropic action (Adelhorst et al. 1994a; Watanabe et al. 1994; Siegel et al. 1999; Dods and Donnelly 2015; O'Harte et al. 1998; Mapelli et al. 2009). Towards the C-terminal, residues Phe28* and Ile29* maintain secondary structure and ligand conformation facilitating the recognition of GLP-1 by the receptor (Adelhorst et al. 1994b). In the C-terminal, truncation of residues Arg36* or Gly37* have no effect over GLP-1 potency (Donnelly 2012; Mojsov 1992), but replacement of the ending sequence "VKGR" with its equivalent from glucagon ligand reduced 475-fold affinity (Donnelly 2012; Runge et al. 2003).

1.5.1.4. Type 2 Diabetes Mellitus and Obesity

Type 2 diabetes mellitus (T2DM) is a life-long condition characterised by defective insulin secretion, insulin resistance and chronic hyperglycaemia. T2DM is linked to obesity; Obesity increases 5-times the risk for T2DM and other comorbidities such as heart disease, musculoskeletal disorders, cancer and premature death ("Guidance. Health matters: obesity and the food environment" 2017; Colditz et al. 1995; Chan et al. 1994; Skyler et al. 2017; Global Burden of Disease Study 2013, 2015).

Worldwide obesity and type 2 diabetes mellitus (T2DM) prevalence and incidence rates are rapidly increasing, turning into a considerable medical, social and financial burden (ADA 2018; CDC 2017). In 2017 the International Diabetes Federation (IDF) estimate 425 million people worldwide living with diabetes, from which 90% of the cases were T2DM (IDF 2017). In 2018, 63% of adults in England were classed as overweight or obese and in the United Kingdome more than 3.5 million people live with diabetes (Valabhji 2018).

1.5.1.5. Type 2 Diabetes Mellitus pathophysiology

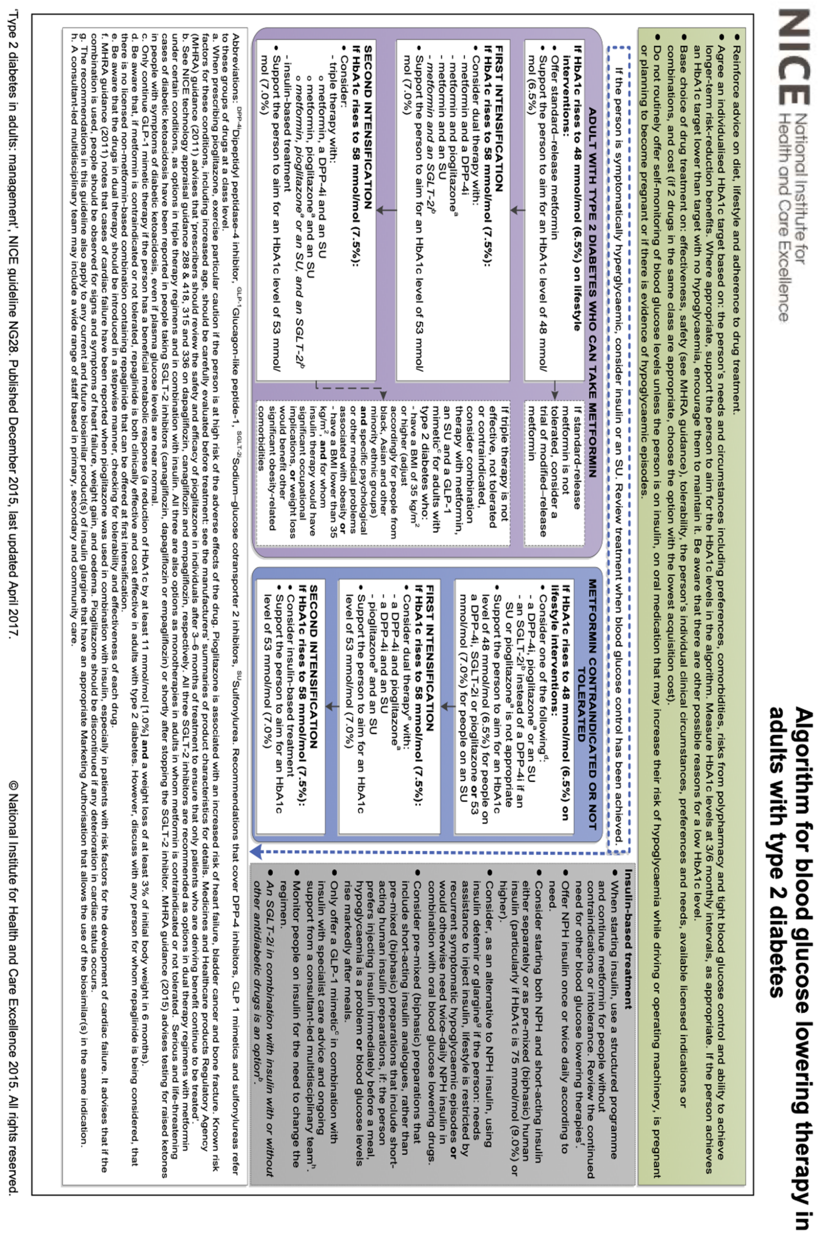
Under physiological conditions, insulin stimulates muscle uptake of glucose (Shulman et al. 1990), however, in insulin resistance states, such as in T2DM and obesity, muscle insulin-stimulated uptake of glucose is impaired, contributing to the persistance of a hyperglycaemic state. In an attempt to overcome peripheral insulin resistance, the pancreas secretes higher levels of insulin. The lack of response to insulin due to insulin resistance fail to supress glucagon production contributing to maintaining the hyperglycaemia (reviewed in van der Zijl et al. 2011). Over time, the chronic hyperinsulinemia and glucotoxicity deteriorates β-cell function and promotes apoptosis, while the sustained hyperglycaemia becomes the main cause for microvascular and macrovascular complication by inducing nerves and vessels damage (reviewed in Kahn, Cooper, and Del Prato 2014). Cardiovascular disease is the main cause of death in T2DM as consequence to the maintained blood vessels damage (Matheus et al. 2013; Orasanu and Plutzky 2009; Kemp et al. 2005; Laing et al. 2003); however, a good control of glycaemia, blood pressure and lipids can reduce

the risk of cardiovascular disease and complications. Therefore, treatment of T2DM is focused on glycemia management to prevent or delay long-term complication and maintain quality of life.

1.5.1.6. Type 2 Diabetes Mellitus treatment

Currently, lifestyle modifications and metformin are the first-line treatment for T2DM, however, within the first year of diagnosis half of the patients will require an additional antihyperglycemic (Raebel et al. 2014). It is important to individualise the approach to diabetes care and to consider safety, efficacy and associated comorbidities when choosing from available second-line options. The National Institute for Health and Care Excellence (NICE) guideline (NG28) published on December 2015 and updated in December 2020 recommend guideline algorithm for diabetes management in non-pregnant adults (Figure 1.5; NICE 2020).

Figure 1.5. NICE Algorithm for diabetes management



1.5.1.7. GLP-1R Agonists

Structural different and synthetic versions of GLP-1 resistant to DDP-4 degradation have been developed for T2DM treatment. These GLP-1R agonists (GLP-1RAs) have similar antihyperglycemic and insulinotropic effects as GLP-1 but with a lower risk for hypoglycaemia when compared to other T2DM medication.

Besides improving glycaemia control, GLP-1RAs had shown to be safe and improve diabetes associated comorbidities. Even though endogenous incretin effect is reduced in T2DM (Nauck et al. 1986), GLP-1 response is preserved in dose-dependent, allowing the use of GLP-1RAs in T2DM treatment (Holst 2019). In addition to glycaemic control, GLP-1RAs improve endothelial and cardiovascular function; promote weight loss and waist circumference reduction, and inflammation; reduce postprandial lipidemia, total cholesterol, low-density lipoprotein (LDL) cholesterol and triglycerides (reviewed in Tahrani, Barnett, and Bailey 2016; Ferrannini et al. 2015) and some, as liraglutide, offer renoprotection delaying the onset of macroalbuminuria and slowing glomerular filtration rate decline (Hare et al. 2010; Asmar et al. 2010; Górriz et al. 2020; Lim 2019; Nauck et al. 2021). It was reported that around 30-50% of patients taking GLP-1R at therapeutic doses will experience nausea and others gastrointestinal side effects, limiting their usage (Shyangdan et al. 2011). However, this was only seen after injection treatment or increasing the dose. Therefore, to overcome this issue, a dose-escalation approach is used to reduced side-effects, and allowing tolerance before exposing patients to higher doses (reviewed in Nauck et al. 2021).

In 2005, exenatide became the first GLP-1R agonist approved for the treatment of T2DM. This synthetic version of exendin-4 (Ex-4), a peptide isolated from the salivary secretion of Gila monster (*Heloderma suspectum*) had 53% sequence similarity to GLP-1 (Fehse et al. 2005) but is more potent that the native GLP-1 (Raufman 1996). Since then, more GLP-1RAs have become available, as short-acting or long-acting compounds to be used as monotherapy or add-on therapy (table 14). Due to their peptide nature, most available GLP-1RAs required to be injected; although improving glycaemic control in patients with T2DM the administration caused discomfort and low adherence to the treatment as most patients prefer an oral drug (Dibonaventura et al. 2010).

In 2020, semaglutide - Rybelsus® the first oral GLP-1R agonist to become available (Limited 2020). This drug is co-formulated with a purported gastric absorption enhancer, salcaprozate sodium (SNAC) which promote oral bioavailability (Griffith et al. 2020). Still the structural basis by which GLP-1 interacts and signal through GLP-1R remains unclear.

Table 1.12. GLP-1R analogues

GLP-1R analogue	Description	Dosing	Admin method	FDA Initial U.S. approval	Commercial name
Exenatide	Synthetic form of Exendin-4 (A8G substitution)	twice daily	SC	2005	Byetta
	Embedded exenatide within biodegradable polymeric microspheres (Drucker et al. 2008)	weekly	SC	2012	Bydureon
Lixisenatide	Synthetic form of Exendin-4 with six Lys residues added to c-terminus (Christensen et al. 2009)	daily	SC	2016	Adlyxin
Liraglutide	Modified GLP-1 linked with a 16-carbon fatty acid chain attaching K26 to albumin (Green et al. 2007)	daily	SC	2010	Victoza
Liraglutide	Higher dose approved for treatment of obesity in people without diabetes or overweight with one weight-related comorbid condition (hypertension, T2DM, dyslipidaemia)	daily	SC	2010	Saxenda
Albiglutide	Two copies of GLP-1 in series, with A8G substitution and fused to albumin (Christensen and Knop 2010) (withdraw from market in 2018)	weekly	SC	2014	Tanzeum
Dulaglutide	Two copies of GLP-1 analogue (substitutions at A8G, G22Q, R36G) covalently linked to Fc IgG4 (Jimenez-Solem et al. 2010)	weekly	SC	2014	Trulicity
Semaglutide	Modified GLP-1 linked with a fatty acid	weekly	SC	2017	Ozempic
Semaglutide	Oral GLP-1 analogue	daily	Oral	2017	Rybelsus

Adapted from (de Graaf et al. 2016) SC: subcutaneous injection

1.5.2. GLP-1R

Glucagon-like peptide-1 receptor (GLP-1R) is a class B GPCR and a key regulator in glucose homeostasis. This 64 kDa transmembrane protein is mainly expressed in pancreatic β-cells where activation by its endogenous ligand, GLP-1 produces insulin secretion. Besides the expression of GLP-1R in pancreatic α- and β-cells, GLP-1R are also expressed in brain tissue mainly around the circumventricular organs and neighbouring regions, epithelial-vascular structure that secretes cerebrospinal fluid, heart, liver and lungs (Ast et al. 2020; Baggio and Drucker 2014; Drucker 2016; Farkas et al. 2021; Yokomori and Ando 2020; Baggio et al. 2018); explaining GLP-1R extra pancreatic effects.

1.5.2.1. Structure and Activation

Up to July 2021 there are 14 full-length structures of GLP-1R in the active state (PDB code 5VAI, 5NX2, 6B3J, 6ORV, 6XOX, 7C2E, 6VCB, 6WHC, 7E14, 7DUQ, 7LCI, 6X18, 6X19 and 6X1A) bound to different ligands (Zhang et al. 2017; Jazayeri et al. 2017; Liang et al. 2018; Zhao et al. 2020; Ma, Huang, et al. 2020; Bueno et al. 2020; Chang et al. 2020; Cong et al. 2021; Zhang et al. ; Zhang et al. 2020b), an inactive full-length human GLP-1R without orthosteric ligand (PDB 6LN2) (Wu et al. 2020), and five TMD of GLP-1R in the inactive state (PDB 5VEW, 5VEX, 6KK1, 6KJV, 6KK7; Song et al. 2017; Xu et al. 2019). All these structures confirm GPCRs seven-transmembrane domain and providing insights into Class B structure and activation mechanisms. Engagement of the GLP-1 within the TMD core trigger conformational changes and the rearrangement of the NTD and TMD. Full activation mechanisms and order of these events is still unknown and despite expected differences between active and inactive states there are differences between activated GLP-1R structures. Activated receptors were obtained with the same G protein, however differences, especially at the extracellular domain are visible (Zhang et al. 2017; Jazayeri et al. 2017; Liang et al. 2018; Zhao et al. 2020), evidencing that there is still a lot unknown about GLP-1R activation and structure. Isolated ECD as full-length GLP-1R show Class B long N-terminal domain connected to the seven-helices transmembrane domain (Hallbrink et al. 2001; Al-Sabah and Donnelly 2003a; Lopez de Maturana and Donnelly 2002; Runge et al. 2008; Underwood et al. 2010).

GLP-1R NTD shares the Class B fold and three disulphide bonds linking residues 46-71, 62-104 and 85-126, and six conserved residues (Asp67, Trp72, Pro86, Arg102, Gly108, Trp110) critical for the stability of the structure, ligand binding affinity and selectivity (reviewed in Donnelly 2012). Full-length GLP-1R structures reveal different NTD conformations relative to TMD (Jazayeri et al. 2017; Liang et al. 2018; Zhang et al. 2017; Zhao et al. 2020). In contrast to others Class B1 receptors (PTH1R, PAC1R and CRF1R) which doesn't require the ECD for activation, the NTD is necessary for GLP-1R activation and signalling (Zhao et al. 2016). Therefore, the different engagement between NTD-ligand and ligand-TMD appear to be the responsible for the different conformation of the ECD relative to the TMD.

In accordance with the two-steps model, the NTD quickly binds to GLP-1 C-terminal, then, the N-terminal end of the ligand interact with the TMD and activate

the receptor. Despite the NTD being mostly involved in ligand binding, GLP-1 affinity also depends on the TMD. Experimental data show that mutation on residues in TMD decrease GLP-1 affinity by altering the binding site of GLP-1 N-terminal (Al-Sabah and Donnelly 2003b; Lopez de Maturana and Donnelly 2002). In contrast, for GLP-1R peptide agonists the NTD have a more important role in binding, and unlike to GLP-1, affinity is independent of mutation in TMD (Al-Sabah and Donnelly 2003b; Lopez de Maturana and Donnelly 2002). Moreover, recent non-peptide agonists PF-06882961 and TT-OAD2 shows the ability of non-peptide ligands to activate GLP-1R by interacting with the NTD without engaging deep into the TMD core (Zhao et al. 2020; Griffith et al. 2020). These recent non-peptide agonists highlight the importance of the NTD-ligand interaction and opens the possibilities for development of more Class B agonists (Zhao et al. 2020; Griffith et al. 2020).

The position of the NTD relative to the TMD is allowed by the stalk, an extension of TM1 into the ECD linking the NTD with the transmembrane domain. The existence of the stalk in GLP-1R was at first considered as an artifact during the crystallisation process, however cryo-EM structures show the existence of this structure. Although visible in cryo-EM structures of activated GLP-1R-G protein complex the stalk is poorly resolved hinting towards high flexibility (Lianget al. 2018). The role of the stalk is unclear since experimental data show that the stalk participates in peptide binding (Lei et al. 2018). In the recent inactive GLP-1R without orthosteric ligand, the stalk and ECL1 form a short α-helix and when comparing structures reorientation of the stalk is evident, supporting its high flexibility and role in binding and position of the ECD relative to the TMD, needed for activation and signalling (Siu et al. 2013; Lianget al. 2018; Wu et al. 2020).

Besides the stalk, ECLs are also involved in peptide recognition and binding, positioning of the ECD relative to the TMD and signalling. The ECL2 has special importance in Gs-dependant signalling (Dods and Donnelly 2015; Zhang et al. 2020a; Griffith et al. 2020; Zhao et al. 2020) and structural stability forming one of three conserved disulphide with TM3 (Dods and Donnelly 2015; Zhang et al. 2020a; Griffith et al. 2020; Zhao et al. 2020). In addition to ECL2, ECL1 and ECL3 are also involved in biased agonism (GLP-1 agonists, oxyntomodulin and exendin-4) (Woottenet al. 2016).

At the TMD the most visible characteristic of the active state is the outward movement of the intracellular end of TM6 and TM7 creating a cavity for G protein

binding (Liang et al. 2018; Liang et al. 2017; Siu et al. 2013; Zhang, Qiao, et al. 2017; Zhang et al. 2018), however, activation involves further conformational changes. At the extracellular end of the TMD interaction between residues of the N-terminal end of GLP-1 with residues forming the binding pocket are crucial for GLP-1R activation, especially His7*, Ala8* and Glu9* (Montrose-Rafizadeh et al. 1997; Zhang et al. 2020a; Zhang et al. 2017). GLP-1 His7* interacts with TM3 and TM5 (Gln234$^{3.37b}$, Val237$^{3.40b}$, Trp306$^{5.36b}$, Arg310$^{5.40b}$, Ile313$^{5.43b}$) triggering changes at the side chains, while Ala8* hydrophobic interactions with TM7 (Glu387$^{7.42b}$, Leu388$^{7.43b}$) (Zhang et al. 2020a) may contribute to TM7 destabilisation facilitating TM7 movement over TM6 using G395$^{7.50b}$ as a pivot, similar to class A (Fredriksson et al. 2003; Zhang et al. 2017). GLP-1 Glu9* interaction with Arg190$^{2.60b}$ and Leu388$^{7.43b}$ is critical for signalling and stabilisation of the active conformation (Lei et al. 2018; Wootten, et al. 2016; Zhang et al. 2020a). At the extracellular side it can be seen the interaction between Asp15* with Asp372^{ECL3}, Arg380$^{7.34b}$ and Leu384$^{7.38b}$ (Zhang et al. 2020a) which may to contribute to destabilise TM7. In addition, within activation, TM1 movement over TM7 breaks Ser155$^{1.50b}$ and Leu396$^{7.51b}$ hydrogen bond destabilising TM7 kink (Hollenstein et al. 2013; Jazayeri et al. 2016; Siu et al. 2013; Zhang et al. 2017).

In the centre of GLP-1R core conserved residues His$^{2.50b}$ and Glu$^{3.50b}$ form an ionic lock which in the inactive state is strengthened by Thr$^{6.42b}$ and Tyr$^{7.57b}$ polar interaction, forming Class B HETx polar network (Hollenstein et al. 2014; Conner et al. 2005). In the middle region of TM6 is located the conserved PxxG motif (Pro$^{6.47b}$-Leu-Leu-Gly$^{6.50b}$). This motif, due to proline and glycine lowest helix propensity, so it can be easily distorted during activation, facilitating the outward movement of TM6 resulting in a kink in the middle region of TM6 and the exposure of carbonyl oxygen from Pro$^{6.47b}$-Leu-Leu-Gly$^{6.50b}$ backbone allowing the formation of polar interactions with residues in TM3, TM5 and TM7 to stabilise the active conformation (Bailey and Hay 2007; Conner et al. 2005; Zhang et al. 2017).

At the intracellular side of the receptor, residues Arg$^{6.37b}$ and Glu$^{7.63b}$ form the intracellular ionic lock maintaining the receptor in an inactive conformation. Binding of the ligand destabilises HETx polar network and TM6 leading to the disruption of the intracellular lock and releasing residues for interaction with the G protein (Zhang et al. 2017). Additional conformational changes in the receptor facilitate the binding of the G protein (Hilger et al. 2019). These changes involve the upward shift of TM3

by at least a helical turn and the movement of the intracellular end of TM5 towards TM6. This creates a hydrophobic cavity formed by TM3 and TM6 for the binding of Y391 of α5 helix (Zhang et al. 2017). Further hydrophobic interactions between TM6 and TM8 with TM2, and hydrophilic interactions between residues at the intracellular end of TM5 and the G protein facilitate the binding of the G protein (Hilger et al. 2019).

1.5.2.2. Signalling

In pancreatic β-cells, activation of GLP-1R increases insulin secretion through activation of $G\alpha_s$. Active $G\alpha_s$ upregulates adenylate cyclase activity and increases cAMP production and further activation of PKA and other downstream messengers leading to insulin release.

Although GLP-1R mainly activates $G\alpha_s$ subunit, it has been observed to also couple to other G protein subunits (Hallbrink et al. 2001; Garant, Yang, and Bernier) and even to recruit β-arrestins and signal in G protein-independent pathways (reviewed in Holz and Chepurny 2005).

Studies using GLP-1R coupled to G protein chimera suggest that agonist affinity is determined by G protein preference, and G preference appear to be determined by the C-terminal end of the peptide (Weston et al. 2014). Surprisingly the recent structure of a GLP-1R activated by a non-peptide agonist biased towards G protein (PDB 6XOX) revealed the compound bound to a binding pocket located in the ECD where it interacts with $Trp33^{NTD}$ (Kawai et al. 2020). Suggesting that in the presence of the G protein subunit subtle changes triggered by the NTD of the receptor can stabilise GLP-1R-G protein complex.

1.6 PTH receptors and ligands

1.6.1 Ligands

Parathyroid hormone (PTH), parathyroid hormone-related protein (PTHrP) and tuberoinfundibular peptide of 39 amino acids (TIP-39) are endogenous peptide ligands involved in calcium and inorganic phosphate metabolism (Segre et al. 1979; Bergwitz et al. 1996; Zhao et al. 2019). Although being immunologically different, these peptides share a similar N-terminal with the first 34 amino acids being critical for activity (Segre 1996), however, differences in the sequence of the ligands and receptors are responsible for affinity and efficacy.

Table 1.13. Comparison PTH ligands

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
PTH (1-34)	S	V	S	E	I	Q	L	M	H	N	L	G	K	H	L	N	S	M	E	R	V	E	W	L	R	K	K	L	Q	D	V	H	N	F			
PTHrP (1-36)	A	V	S	E	H	Q	L	L	H	D	K	G	K	S	I	Q	D	L	R	R	R	F	F	L	H	H	L	I	A	E	I	H	T	A	E	I		
TIP39	S	L	A	L	A	D	D	A	A	F	R	E	R	A	R	L	L	A	A	L	E	R	R	H	W	L	N	S	Y	M	H	K	L	L	V	L	D	A	P	39

1.6.1.1 Parathyroid hormone (PTH)

The parathyroid hormone (PTH) is an 84-residue hormone that regulates calcium homeostasis. It is secreted by the parathyroid glands to regulate extracellular calcium concentrations by acting in bones, kidneys and intestines (Mayer 1979). Calcium regulates physiological processes including hormones, neurotransmitter and gastrointestinal secretion, blood coagulation, muscle contraction and skeletal development (Cheung 1980; Carafoli. 1979). Low levels of extracellular calcium cause the secretion of PTH whose main functions are to increase the rate of mineral bone dissolution, calcium filtration at renal glomerulus, calcium absorption from the small intestine and stimulation of the synthesis of vitamin D (Cheung 1980; Carafoli. 1979; Segre et al. 1979). Contradictory, if PTH is administrated intermittently, it promotes bone formation which has led to the use of N-terminal residues hPTH(1-

34), also known as teriparatide, being developed for the treatment of osteoporosis (Miller et al. 2016).

Residues at position 1-14 are critical for signalling with the 4 first residues being critical for activation, as the potent antagonist LA-PTH(5-36) show (Zhao et al. 2019; Segre 1996).

1.6.1.2Parathyroid hormone-related protein (PTHrP)

PTHrP is a 141 residues ligand with PTH-like effects, but unlike PTH, it is expressed in different tissues (skin, kidney, mammary gland, uterus, placenta) where it acts locally. PTHrP is well known for its role in chondrocyte differentiation. However, tumour produced PTHrP is consider a major mediator of humoral hypercalcaemia and cachexia. In conjunction with PTH and vitamin D, PTHrP increase skin immune function (Muehleisen et al. 2012).

Recently PTHrP analogue (abaloaratide) was approved for the treatment of osteoporosis (Kendler et al. 2018) and it has been considered as a treatment for cancer-related cachexia (Hesse et al. 2019; Suvannasankha et al. 2018).

1.6.1.3Tuberoinfundibular peptide 39 (TIP39)

Tuberoinfundibular peptide of 39 residues (TIP39) is a neuropeptide associated with pituitary function modulating stress response, thermoregulation, prolactin release and nociception and in the skin, it regulates keratinocyte calcium homeostasis and differentiation (Muehleisen et al. 2012; Sato et al. 2016b; Usdin, Gruber, and Bonner 1995). TIP39 activates the adenylyl cyclase pathway in PTH_2 receptor and elevates intracellular calcium but acts as an antagonist in PTH_1 receptor (Weaver et al. 2017; Dobolyi et al. 2012; Hoare, Gardella, and Usdin 2001).

1.6.2 Parathyroid hormone receptor

Parathyroid hormone receptor 1 (PTH_1 receptor) and parathyroid hormone receptor 2 (PTH_2 receptor) are Class B GPCRs (Segre et al. 1979; Usdin, Gruber, and Bonner 1995; Abou-Samra et al. 1992). Even though PTH_1 and PTH_2 receptors share 50% identity and are activated by PTH hormone, PTH_1 receptor is activated by TIP39 but not by PTHrP, in the same way PTH_2 receptor is activated by PTH but not by PTHrP suggesting the role of different residues in activation.

1.6.2.1 Parathyroid hormone receptor 1

Activation of PTH_1 receptor in kidneys and bone by the endogenous ligands PTH and PTHrP activates cAMP and PLC/PKC pathways controlling calcium homeostasis and a target for the treatment of osteoporosis (Lanske et al. 1996; Rankin, Grill, and Martin 1997; Abou-Samra et al. 1992; Juppner et al. 1991; Bringhurst et al. 1993; Cheloha et al. 2015).

Recent full-length crystal PTH_1 receptor structures in complex with peptide ligands confirmed that PTH_1 receptor share Class B NTD fold (α-β-β-α) but also showed i) the presence of an interaction network between the ligand and the NTD (Ehrenmann et al. 2018) where a salt bridge between $Asp137^{NTD}$ and Arg20* is essential for affinity and ligand binding (Weaver, Wigglesworth, and Donnelly 2014); ii) different orientation of the NTD relative to the TMD depending on the peptide bound (Zhao et al. 2016; Zhao et al. 2019); iii) the greater influence helicity of the peptide ligand haver over the orientation of the NTD relative to the TMD, rather than receptor-ligand interactions (Ehrenmann et al. 2018); and how despite GPCRs having a similar packing of the TMD, iv) the position adopted by the NTD after the peptide binding modify transmembrane helices and binding pocket contributing to additionally specificity to ligand binding (Zhao et al. 2016; Luck, Carter, and Gardella 1999). This confirms the role of the NTD in the first step of the two-steps activation model, highlight the flexibility of the NTD as it adopts different conformations and orientations relative to the TMD depending on the ligand in order to accommodate the bound ligand and show the effect of the ligand over TMD conformation and signalling outcome. In addition, similar behaviour between the NTD-peptide-TMD

is seen in GCGR receptor (Qiao et al. 2020) and GLP-1R, supporting common activation mechanisms.

Unlike glucagon family receptors, no stalk is seen in the PTH_1 receptor (Ehrenmann et al. 2018) but instead PTH_1 receptor is characterised by a longer and flexible ECL1 (Zhao et al. 2016; Zhao et al. 2019). As in GLP-1R and GCGR, ECL1 contributes to the stabilisation of the peptide during binding (Yang et al. 2015a; Liang et al. 2018; Zhang et al. 2018) but unlike GCGR (Qiao et al. 2020), ECL1 is not strictly required for binding and activation (Lee et al. 1994; Lee et al. 1995; Zhao et al. 2016), and instead it may substitute the stalk role as ECL1 flexibility contributes to i) orientate the ECD relative to the TMD according to the peptide agonist (Ehrenmann et al. 2018) and ii) stabilise the position of the peptide during binding (Yang et al. 2015a; Liang et al. 2018; Zhang et al. 2018).

In addition to interacting with ECL1, the ligand also interacts with ELC2 and ECL3 (Ehrenmann et al. 2018); since these loops connect TM3-TM4 and TM6-TM7 extracellular ends, helix involved in conformational changes, ECLs-ligand interaction may trigger conformational changes from the extracellular region to the intracellular region in addition to binding.

The TMD is shaped as an open V-shaped cavity (Culhane et al. 2018; Ehrenmann et al. 2018; Zhao et al. 2019) with residues Arg233$^{2.60b}$, Asp295$^{3.43b}$, His420$^{6.52b}$ and Q451$^{7.49b}$, belonging to the central polar network and critical for activation, delimiting the orthosteric binding pocket (Ehrenmann et al. 2018; Liang et al. 2017; Zhang et al. 2018). As the NTD accommodate the peptide during the second step of the two-steps model activation, the ligand N-terminus modify the binding pocket and contribute to additionally specificity to the ligand binding (Zhao et al. 2016; Luck, Carter, and Gardella 1999). Comparison of active PTH_1 receptor with GLP-1R show a higher occupancy of the ligand in the TMD due to a tighter binding resulting in a higher and sustained potency (Zhao et al. 2019) and enhanced affinity when coupled to G proteins (Shimizu et al. 2016; Zhao et al. 2019).

The bound ligand adopts a position parallel to TM2 and mainly interacts with residues Phe184$^{1.36b}$, Glu180$^{1.32b}$, Arg181$^{1.33b}$, Tyr195$^{1.47b}$ at TM1, Arg233$^{2.60b}$, Leu244$^{2.68b}$ and Tyr245$^{2.7b2}$ at TM2, Phe288$^{3.36b}$ in TM3, Glu444$^{7.42b}$ and Met445$^{7.43b}$ in TM7 (Ehrenmann et al. 2018). These residues have the function of improving binding by stabilising the N-terminal end of the ligand and translate the signal and conformational changes to the central polar network (Ehrenmann et al. 2018).

Residues Arg233$^{2.60b}$ and Gln451$^{7.49b}$ are critical for ligand binding (Gardella et al. 1996), stabilisation of the central polar network and further conformational changes around TM6 kink and activation (Ehrenmann et al. 2018). These residues (Arg$^{2.60b}$ and Gln$^{7.49b}$) appear to be critical for function in Class B as similar interactions and function are seen between the ligand - Arg$^{2.60b}$, Gln$^{7.49b}$ in GLP-1R and GCGR (Gardella et al. 1996).

After binding of the ligand, conformational changes and the presence of water molecules mediate interhelix interactions Ser229$^{2.56b}$-Asn295$^{3.43b}$ and Tyr296$^{3.44b}$-Ser370$^{5.46b}$-Asn374$^{5.50b}$ that stabilise TM2, TM3, TM4 and TM5 interface during activation (Ehrenmann et al. 2018).

Rearrangement of the central polar network allows the shifting of residues in helix 7 after binding, rearranging residues neighbouring to TM6 and stabilising the formation of TM6 kink; i) Met445$^{7.43b}$, critical for affinity and potency (Lee et al. 1995; Gardella et al. 1994) facilitates TM6 destabilisation; ii) downward movement of Gln451$^{7.49b}$ breaks its interaction with the ligand and Arg233$^{2.60b}$ to form hydrogen bonds with the backbone oxygen of Pro$^{6.47}$ and Gly$^{6.50}$ from the conserved P$^{6.47}$xx G$^{6.50}$ motif stabilising TM6 kink; iii) H$^{6.52}$ shifts between TM6 and TM7 to interact with Gln451$^{7.49b}$ contributing to the TM6 kink. In PTH_1 receptor two helical turns away from Gln451$^{7.49b}$ is residue Ile458$^{7.56b}$, suggested to stabilise the intracellular part of TM6, as mutation to a smaller residue such Ala458$^{7.56b}$ destabilises TM6 producing continuous receptor activation associated with Jansen's metaphyseal chondrodysplasia (Schipani et al. 1999).

1.6.2.2 Parathyroid hormone receptor 2

Currently there is not a full-length PTH_2 receptor but experimental data and PTH_1 receptor structures have been used for the understanding of PTH_2 receptor as both human receptors share 51% amino acid sequence (Usdin, Gruber, and Bonner 1995).

PTH_2 receptor is found in the central nervous system and skin where is activated by TIP39, it can be *in vitro* activated by PTH but not by PTHrP (Mann, Wigglesworth, and Donnelly 2008; Weaver et al. 2017; Usdin, Gruber, and Bonner 1995; Abou-Samra et al. 1992; Hoare, Gardella, and Usdin 2001; Sato et al. 2016b).

Similarities between the ligands and receptors suggest common binding and activation mechanisms.

Different residues at the C-terminal end of PTH or PTHrP and the NTD of PTH_1 or PTH_2 receptors explain differences in affinity. The C-terminal end of PTH, Trp23*, binds PTH_2 receptor NTD, $Val41^{NTD}$. However, PTHrP short side chain Pro23* is unable to interact with the also short side chain $Val41^{NTD}$ from PTH_2 receptor (Mann, Wigglesworth, and Donnelly 2008). On the other hand $V41L^{NTD}$ mutation in PTH_2 receptor allows PTHrP Pro23* to bind the PTH_2 receptor, but not PTHrP Trp23* as the longer side chains clash avoid the interaction (Mann, Wigglesworth, and Donnelly 2008).

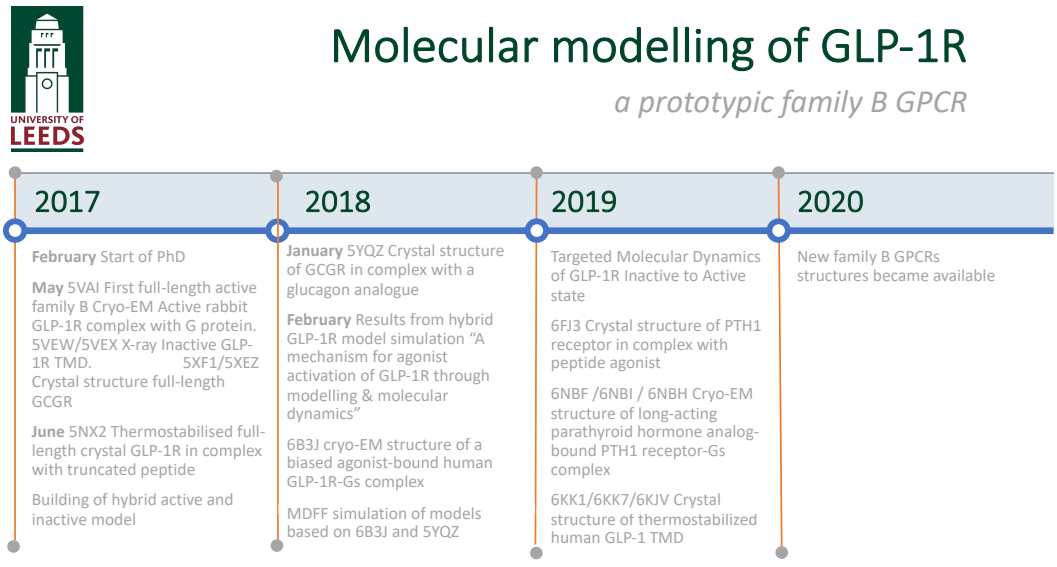
Other residues contribute to ligand binding and affinity; in the TMD residue $Tyr318^{5.39b}$ in TM5 confers high affinity binding to TIP39 where the hydroxyl group is important for ligand recognition and interaction with the N-terminal residues in the second step (Weaver et al. 2017). Removal of the first four residues, specially Leu4* from TIP39 reduces potency (Weaver et al. 2017); showing the role of the N-terminal for potency and activation and the specificity between ligands and receptor.

1.7 Aims and objectives

The over-arching aim of this project was to gain an understanding GLP-1R's molecular mechanism of action and the key interactions leading to agonist-mediated activation. At the start of this project (February 2017) there were no experimentally determined GLP-1R structures available, and hence the initial approach was to use experimentally determined structures of other family B GPCRs for homology modelling of GLP-1R. However, soon after, GLP-1R structures in active and TMD of inactive state became available (Jazayeri et al. 2017; Zhang et al. 2017; Song et al. 2017). These high-quality structures, and the ones that followed, provided valuable information about GLP-1R interactions with the ligand as well with the G protein. Nevertheless, none of these structures showed GLP-1R in complex with GLP-1, the endogenous peptide. Therefore, taking advantage of the experimental available structures, GLP-1R -GLP-1 complex models were built, in the active and inactive state with the aim of understanding GLP-1R activation and key interactions leading

to activation. In order to do this, molecular dynamics simulations (MDs) were used. MDs allows the study of experimentally obtained structures at atomic-resolution in time-steps that would be impossible to observe experimentally; since MDs rely on experimental data to guide and validate results. Through the project different MD methods were used. The main aim of this project was to build a reliable GLP-1R-GLP-1 complex in the active and inactive state and identify the key interactions characterising each state. After defining interactions in the active and inactive state, the second aim was to identify the changes that lead to activation of GLP-1R. Since Class B share structural characteristics and possibly a similar activation mechanism, another Class B GPCR, the PTH1 receptor model was built and used to compare characteristics of the active and inactive states in Class B as well activation mechanisms.

1.8 Timeline



Chapter 2: Methods

2.1 Homology Modelling

Homology modelling is a tool for the prediction of the structure from the amino acid sequence of a target protein. Homology modelling relies on the availability of the structure of one or more proteins with similar sequence (templates). Its reliability depends strongly on the similarity between the target and templates.

During the different stages of this project, models of human class B GPCRs (GLP-1R and PTH_1 receptor) in active and inactive states and their endogenous ligands were modelled relying on available template structures at the time. Most of the receptor structures were obtained from non-human organisms and agonists ligand. However, target models were human, therefore, human sequences were used as target sequences. The building of a homology model consists in different steps that can be summarised as follow:

i) Identification and selection of templates
ii) Alignment
iii) Model building
iv) Model optimization

2.1.1 Identification and selection of the templates and sequences

Templates were selected based on their availability in the PDB bank (Zhang et al. 2017; Jazayeri et al. 2017; Song et al. 2017; Liang et al. 2018; Zhang et al. 2018; Zhao et al. 2019; Ehrenmann et al. 2018), structural homology, sequence identity, root mean square deviation (RMSD), conformational state, resolution, and experimental method (Table 2.1). Human amino acid sequences were used as targets (Table 2.1).

Table 2.1 Modelling of Class B GPCRs

Model	Template PDB code	Target sequence UniProt code
Human GLP-1R hybrid active	5VAI, 5NX2	P43220
Human GLP-1R hybrid inactive	5NX2, 5VEW	P43220
Human GLP-1R active	6B3J	P43220
Human GLP-1R inactive	5YQZ	P43220
Human PTH1 receptor active state	6NBF	Q03431
Human PTH1 receptor inactive state	6FJ3	Q03431

Zhang et al. 2017; Jazayeri et al. 2017; Song et al. 2017; Liang et al. 2018; Zhang et al. 2018; Zhao et al. 2019; Ehrenmann et al. 2018

2.1.2 Alignment and Model Building

Previous alignment and model building, all residues not belonging to the receptor were removed and model building manipulations were carried out using the tools embedded within PyMOL (Schrödinger 2015).

Since the template structure coordinates and target amino acid sequence are known, target-template alignment mode 'User Template Mode' from SWISS-MODEL web server (Waterhouse, Rempfer, et al. 2018) was used for building the model and missing sidechains or residues. The FASTA target sequence and PDB file containing the coordinates of the receptor were used as input files. The high-quality structures used as templates and known target amino acid sequences resulted in valuable models.

Ligands from the original structures were *in silico* mutated to the target amino acid sequence using PyMOL (Schrödinger 2015) and then added to the model. This was done by aligning the structures to their homologous in the original structure; the model of the receptor was aligned to the receptor in the template, then the mutated ligand and in the active state models the α5-helix of the G_s protein was aligned and added. Each complex, formed by models of the receptor-ligand (inactive state) or the receptor-ligand-G protein (active state) was saved into a PDB file.

2.1.3 Model optimisation

The PDB file of complex was submitted to CHARMM-GUI (Jo et al. 2008) to create the files needed for model optimisation. Using NAMD (Bernardi et al. 2016) each complex underwent energy minimisation through steepest descent method (50 steps), followed by Newton Raphson methods (50 steps) to remove steric clashes or poor bond/torsion geometrics.

2.2 Assembly of the system

2.2.1 Orientation of the protein into the lipid bilayer

Prior assembly of the system, hydrophobic thickness and orientation of the receptor with respect to the membrane was calculated using OPM database (Lomize et al. 2012) (Figure 2.1).

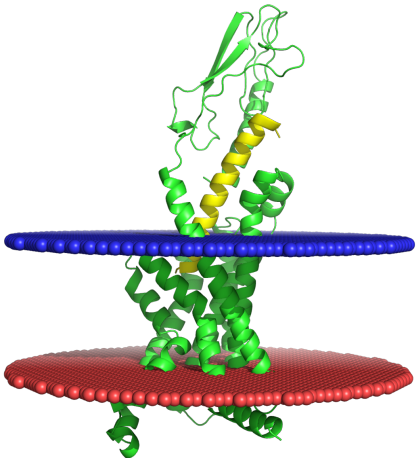


Figure 2.1. Orientation of the receptor-ligand (green and yellow respectively) in the membrane based on the hydrocarbon core of the lipid bilayer (outer – blue, inner – red) given by OPM webserver OPM database (Lomize et al. 2012).

2.2.2. Building of the system

Each system involving receptor-ligand-G protein or receptor-ligand, lipid membrane and aqueous environment was built using the CHARMM-GUI server (Jo

et al. 2008). The 'Membrane Builder' mode generated the files used for hybrid GLP-1R all-atoms MD simulation and targeted MD simulation of GLP-1R, PTH_1 receptor and β_2AR (Jo, Kim, and Im 2007; Jo et al. 2009; Lee et al. 2016; Lee et al. 2019; Wu, Cheng, et al. 2014) (Chapters 3, 5 and 6). 'MAP Utilizer MDFF' (Qi et al. 2017) mode generated the input files for GLP-1R, PTH_1 receptor and β_2AR MDFF simulation (Chapters 4 and 6).

2.2.2.1 Membrane composition

Modes were embedded into a 90 x 90 $Å^2$ lipid membrane (approximately a total of 256 lipid molecules or 128 for each leaflet). The systems were built using replacement method, as recommended for transmembrane proteins and allowing the random distribution of lipid molecules for heterogeneous membranes; no protein surface or lipid ring penetration was found (Jo, Kim, and Im 2007).

In the initial simulations, a homogenous 90 x 90 $Å^2$ lipid 1-palmitoyl-2-oleoyl-sn-glycero-3-phophochoine (POPC) bilayer membrane was used in hybrid molecules (Chapter 3). Homogeneous POPC bilayer has been previously used for the simulation of GLP-1R model (Wootten et al. 2016).

In the following simulations (Chapters 4, 5 and 6) a heterogeneous 90 x 90 $Å^2$ lipid membrane composed of cholesterol 25%, POPC 26%, POPE 20%, POPS 5%, POPI 5%, PSM 19%) was used (Table 2.2). The purpose of this membrane was to resemble pancreatic β-cells membrane. Pancreatic β-cells membranes are dynamic structures constantly modifying its composition and lipid distribution according to changes in glucose concentration and insulin secretion (Hallberg 1984a; Hagren and Tengholm 2006; Thore, Dyachok, and Tengholm 2004; Thore et al. 2005; Wuttke, Sågetorp, and Tengholm 2010; Montague and Parkin 1980; Burke and Ellenberg 2002; Berne 1975). Freshly isolated β-cells at basal state report 20% PE (Díaz et al. 1988), but this value quickly changes with variations in glucose and insulin secretion as insulin-secretory granule (ISG) fuse with the plasma membrane (MacDonald et al. 2008; Hoang Do and Thorn 2015). Pancreatic membrane usually contains 5% POPI asymmetrically distributed and has shown to be a positive regulator of cytoplasmic concentration of Ca^{2+} in pancreatic β-cells (Lin et al. 2005; Xie et al. 2016). In physiological conditions PI concentration changes during receptor-triggered

activation of phospholipase C (PLC) and glucose changes (Rana, Kowluru, and MacDonald 1986). Sphingomyelin percentage in pancreatic β-cells accounts for 19.2% (Meldolesi, Jamieson, and Palade 1971) in contrast with 7.5% reported in plasma membrane model (Koldso et al. 2014). Sphingomyelin role in β-cells is unclear, however, cholesterol stabilised sphingomyelin patches found in pancreatic β-cells show a direct correlation between insulin secretion and the amount of sphingomyelin patches (Kavishwar and Moore 2013). In another study, sphingomyelin synthase-1 (SMS-1) knockout mice show severe deficiencies in insulin secretion (Yano et al. 2011). These studies suggest a special role of sphingomyelin in the normal function of pancreatic β-cells.

Table 2.2. Composition of the lipid bilayer based on the pancreatic β-cell membrane composition used in the simulations presented in Chapters 4 to 6.

Lipid	%	Outer bilayer	Inner bilayer
Cholesterol	25	25	25
Phosphatidylcholine (POPC)	26	40	12
Phosphatidylethanolamine (POPE)	20	8	32
Phosphatidylserine (POPS)	5	2	8
Ganglioside G_{M3}	/	/	/
Phosphatidylinositol (POPI)	5	/	10
Sphingomyelin	19	28	10

The percentage of lipids was used to build the lipid bilayer of the systems described in Chapters 4 to 6. Koldson et al. 2014; Díaz et al. 1988; Lin et al. 2005; Xie et al. 2016; Meldolesi, Jamieson, and Palade 1971; Kavishwar and Moore 2013.

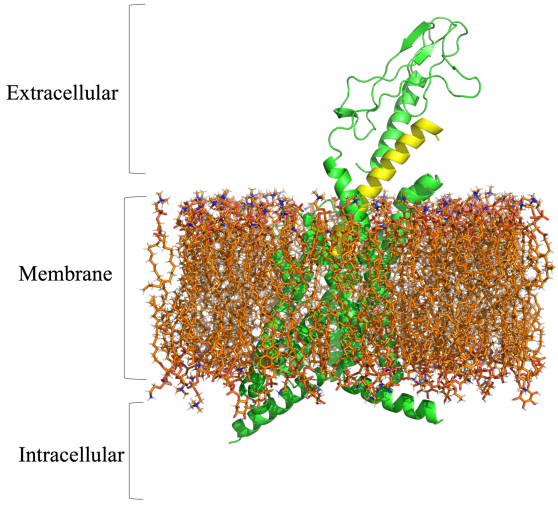


Figure 2.2. Cartoon representation of receptor-ligand complex (green and yellow respectively) embedded into a heterogeneous lipid bilayer (orange stick representation) as predicted by OPM server (Lomize et al. 2012) (Figure 2.1).

2.2.2.2 Solvation

The system was solvated with TIP3P water molecules, with an approximately water thickness 22.5 Å above and below the system. 0.15 M NaCl were added to neutralise the net charge of the system by placing them using Monte Carlo method (Figure 2.3).

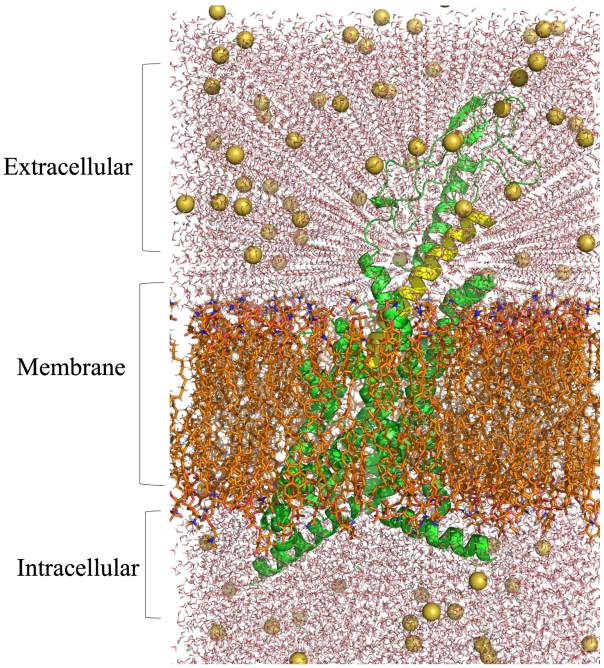


Figure 2.3. Cartoon representation of receptor-ligand complex (green and yellow respectively) embedded into a heterogeneous lipid membrane (orange stick representation) solvated with TIP3P water (red) and ions (yellow).

2.3 Molecular Dynamics Simulation

Molecular dynamics simulations compute atomic trajectories by solving the equations of motion. Since the potential energy function is a function of 3N atomic position, it calculates the force experienced by any atom given the position of other atoms. Newton's second law: $F = ma$, where F is the force on an atom, m is the mass of the atom and a is the acceleration of the atom, describes how those forces affect the movement of the atoms. So, the system of particles is governed by the equation of motion:

$$m_i \left(\frac{d^2 r_i}{dt^2} \right) = f_i$$

where m_i is the mass of i particle, f_i is the force acting on i, and r_i represents the position coordinates of i. Therefore, the forces fi acting on atoms are usually derived from a potential energy $U(r^N)$ where $r^N = (r_1, r2, \ldots r_N)$ represents the complete set of 3N atomic coordinates:

$$f_i = -\left(\frac{\partial U(r_1, r_2, \ldots r_N)}{\partial r_i}\right)$$

and the potential energy (U) is obtained from:

$$U = \sum_{\text{bonds}} \frac{1}{2}k_b(r - r_0)^2 + \sum_{\text{angles}} \frac{1}{2}k_a(\theta - \theta_0)^2 + \sum_{\text{torsions}} \frac{V_n}{2}[1 + \cos(n\phi - \delta)] + \sum_{\text{improper}} V_{imp} + \sum_{\text{LJ}} 4\epsilon_{ij}\left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}}\right) + \sum_{\text{elec}} \frac{q_i q_j}{r_{ij}}$$

Bonded interactions, local contributions or intramolecular contributions to the total energy

Repulsive, Van der Waals and the Coulombic contributions to the total energy

where r_0 is obtained from X-ray diffraction experiments, k is the spring constant that may be estimated from infrared or Raman spectra, ϕ is the torsional angle, δ is the phase, n defines the number of minima or maxima between 0, V_n determines the height of the potential barrier, ij interaction between two atoms, $i, j,$ 12-6 correspond to the Lennard-Jones (LJ) potential, ϵ_{ij} is the LJ potential corresponding depth of the well for the interaction between two atoms, i, j, σ represents the value at which the potential becomes zero. Because an enormous computational effort would be needed to do this for every atom, standardized empirically derived force fields (FF) based on tested properties and experimental data (MacKerell 2000) are used to maintain and reproduce molecular geometries of bond lengths, bond angles, bond torsions, non-bonding or improper and electrostatic interactions. FF subdivide potential function into the sum of: Bonded interactions, also known as local contributions, which include covalent bond-stretching (Σ_{bonds}), angle-bending (Σ_{angles}), torsion potential ($\Sigma_{torsion}$) and improper torsion potential ($\Sigma_{improper}$); and non-bonded interactions which include Lennard-Jones (Σ_{LJ}) (short-range Pauli repulsion and long-range Van der Waals attraction) potential and Coulomb electrostatics (Σ_{elec}) (Guvench and MacKerell 2008; MacKerell 2000). To reduce the cost, local interactions (bonded, van der Waals) are calculated at each timestep and longer range interactions are calculated less often.

At each time step of the simulation the force f_i needs to be computed to update the position r_i. and velocity of atoms; and at the new positions, the atomic forces are re-calculated. For simplicity, integration algorithms such the Verlet algorithm, and its variations are used. The Verlet algorithm integrates Newton's equation of motion; in the Verlet algorithm velocities are not explicitly solved but calculated from first order central difference and directly relate the force to the position (Verlet L. 1968; Swope et al. 1982). The simulations used the velocity Verlet algorithm which has an improved accuracy compared to standard Verlet; it starts with the position (r) of atoms at the previous timestep (t) and velocity expansions; positions (r), velocities (v) and accelerations (a):

Calculates the new positions at time t+ Δt

$$r(t + \Delta t) = r(t) + \Delta t v(t) + \frac{1}{2}(\Delta t)^2 a(t)$$

Calculates velocities at mid-step:

$$v\left(t + \frac{1}{2}\Delta t\right) = v(t) + \frac{1}{2}\Delta t a(t)$$

Calculates the positions at the next step:

$$r(t + \Delta t) = r(t) + v\left(t + \frac{1}{2}\Delta t\right)/\Delta t$$

Force evaluation carried out from the potential and providing:

$$\Delta t a(t + \Delta t)$$

for the next step

Updates the velocities:

$$v(t + \Delta t) = v\left(t + \frac{1}{2}\Delta t\right) + \frac{1}{2}\Delta t a(t + \Delta t)$$

To reduce the computational cost the Verlet-I algorith divides the potential energy as long-range interactions, which are computed less frequently than short-range interactions, which are computed more frequently:

$$U = U^{\text{long-range}} + U^{\text{short-range}}$$

In this way, the velocity Verlet integration method is used to advance the positions and velocities of the atoms in time and to reduce computational cost.

In the molecular dynamics simulations described in the following chapters, including equilibration and production, were performed using NAMD software (Bernardi et al. 2016), using CHARMM36 force-field (Huang and MacKerell 2013), and the Impulse-based Verlet-I, also known as r-RESPA method, performing multiple timestep integration at a 2fs for short-range nonbonded forces and 4 fs for long-range electrostatics.

2.3.1 Equilibration

The assembled system underwent six steps of equilibration (375 ps) prior to MDs production (Figure 2). Equilibration steps 1 and 2 were performed at constant volume and temperature (NVT) and steps 3 to 6 at constant pressure, area, and temperature dynamics (NPAT). Through equilibration, constraints applied to protein backbone and sidechains, water, lipids and ions molecules were slowly released.

2.3.2 All-atoms Classic Molecular Dynamics

All-atoms molecular dynamic method was used for the simulation of hybrid GLP-1R (Chapter 3). Simulation of the system involving protein, water and lipids were performed using CHARMM36 force-field (Huang and MacKerell 2013) and periodic boundary conditions, at 303.15 K temperature and 1 atm pressure using Langevin dynamics with a damping coefficient of 1-ps to control the temperature and a Nose-Hoover Langevin piston to control the pressure. Particle Mesh Ewald method was used to calculate long-range electrostatic interactions. Bond lengths involving hydrogen atoms were fixed using the SHAKE algorithm (Barth et al. 1995) using a 2 fs time step.

2.3.3 Molecular dynamics flexible fitting (MDFF)

Molecular dynamics flexible fitting (MDFF) is an extension of MDs. MDFF restrain the dynamics by adding an external harmonic potential dependent on the cryo-EM electronic density map gradient (U_{EM}), MD potential energy (U_{MD}) and the potential of secondary structures (U_{ss}) preserving secondary structures while preventing distortion and overfitting (Trabuco et al. 2008; Trabuco et al. 2009). MDFF uses the information contained in the electronic density map, which includes regions where the density is low and the structure unresolved.

$$U_{total} = U_{MD} + U_{EM} + U_{SS}$$

High-density areas from the density map correspond to energy minima and data not corresponding to the molecule, like the solvent contribution, is removed (Trabuco et al. 2008). Therefore, the potential from the density gradient map is defined by

$$U_{EM}(R) = \sum_j w_j V_{EM}(r_j) \qquad \text{where} \quad V_{EM}(r) = \begin{cases} \xi \left[1 - \dfrac{\phi(r) - \phi_{thr}}{\phi_{\max} - \phi_{thr}}\right] \end{cases}$$

as w_j corresponds to per-atom weight typically set to atomic mass, ξ is an arbitrary scaling factor ($\xi > 0$) which is the same for all the atoms, $\phi(r)$ is the EM density at position r, ϕ_{max} is the maximum value of the EM density map and ϕ_{thr} is a density threshold selected in accordance with the density histogram which removes data not corresponding to the molecule, like the solvent contribution (Trabuco et al. 2008).

2.3.3.1. Fitting models into electron density map

PTH1 receptor and GLP-1R complexes were simulated using MDFF method. Prior simulation, the models were fitted into their respective density maps (Liang et al. 2018; Zhang et al. 2018; Zhao et al. 2019). The fitting into density maps was done using 'Colores' program from Situs package (Wriggers 2012; Wriggers and Birmanns 2001).

The templates used to build the inactive model of GLP-1R and PTH1 receptor were based on the crystal structures reported in 5YQZ (Zhang, Qiao, et al. 2017) and 6FJ3 (Ehrenmann et al. 2018) respectively. Since CHARMM-GUI requires ccp4 format, the X-ray diffraction maps were first converted from DSN6 to ccp4 using CCP4i Software (McNicholas et al. 2011; Potterton et al. 2003). Correlation between models and density map show no overfitting.

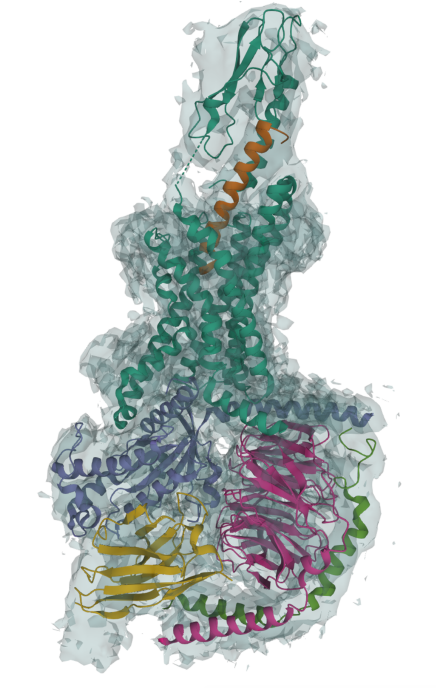


Figure 2.4. Cartoon representation of receptor-ligand-G protein complex embedded into the density map (grey shadow).

2.3.3.2 MDFF Production

Once the models were fitted into the density maps, input files for MDFF production were generated using MDFF Utilizer in CHARMM-GUI server (Jo et al. 2008; Jo et al. 2009; Lee et al. 2016; Lee et al. 2019; Qi et al. 2017; Wu, Cheng, et al. 2014). A heterogeneous lipid membrane with similar lipid composition as the pancreatic cell was used (Cholesterol 25%, POPC 26%, POPE 20%, POPS 5%, POPI 5%, PSM 19%). Each model was independently embedded into a 90 x 90 $Å^2$ heterogeneous lipid membrane before adding neutralizing ions (0.15 M NaCl) and a water/solvent box (TIP3P).

MDFF simulations were performed using NAMD software (Bernardi et al. 2016) as the restraints-free simulations described above, at the same temperature and pressure, however, restraints were applied to Cα atoms during the first 100 ns of the simulation to maintain the gross position of the protein while allowing loops and sidechains to reach favourable positions and interactions while being fit into the map. The scaling factor determines the strength of the potential derived from the electron density map over the protein (Trabuco et al. 2008). The default scaling factor $\xi = 0.3$ kcal/mol (Trabuco et al. 2008) was used since high-quality density maps were used

and higher values resulted in stronger forces acting on the system. A constant force of 200 kcal mol $^{-1}$ Å^{-2} on dihedral angles was used during simulations (Trabuco et al. 2008). After the initial 100 ns, restraints derived from the cryo-EM map were steadily removed from the ligand, transmembrane domain and G protein. Restraints on the N-terminal domain of the receptor were kept avoiding the N-terminal bending over the membrane. The resulting segments were analysed.

2.3.4 Targeted molecular dynamics

Despite advances in computational power, spontaneous conformational changes are still challenging to achieve and are not expected to happen at the timescale used due to the size of the system. Therefore, targeted molecular dynamics simulations method was used to study the activation of human GLP-1R and PTH_1 receptor (Chapters 5 and 6). In targeted MD steering forces guide a subset of atoms towards the final 'target' structure. The force acting in each atom is given by the gradient of the potential:

$$U_{TargetedMD} = 0.5 \frac{k}{N} [RMS(t) - RMS^*(t)]^2$$

k = the spring constant in kcal/mol/Å^2
N = the number of targeted atoms
RMS(t) = RMSD of the current structure
RMS(t)* = RMSD of the targeted structure

At each timestep, the gradient of potential ($U_{TargetedMD}$) is obtained from the difference from the RMSD value from the current structure *RMS(t)*, and the target structure *RMS*(t)*, by aligning the target structure to the current structure. The spring constant *k*, is scaled down by the number *N* of targeted atoms (Bernardi et al. 2016).

2.3.4.1 Target and initial coordinates

The last frame from MDFF simulation of active and inactive cryo-EM models were independently saved to a PDB file. Since targeted MD method requires the same number of atoms, the α5 helix from G_s protein was added to the inactive model by

placing it in the space corresponding to the intracellular side of the receptor with no contact with the receptor. The inactive structure was used as the starting model and the active structure was used as the target coordinates.

2.3.4 Targeted molecular dynamics production

Harmonic restraints were applied only to Cα atoms of the target structure to guide the receptor from the inactive state into the active; the rest of individual atoms were free to move in accordance to the force field and environmental parameters (Schlitter, Engels, and Kruger 1994). Since the targeted MD depends on the coordinates and time, the production time was set for a total of 100 ns. At each time step the distance and expected distance to the target structure were computed showing the conformational changes.

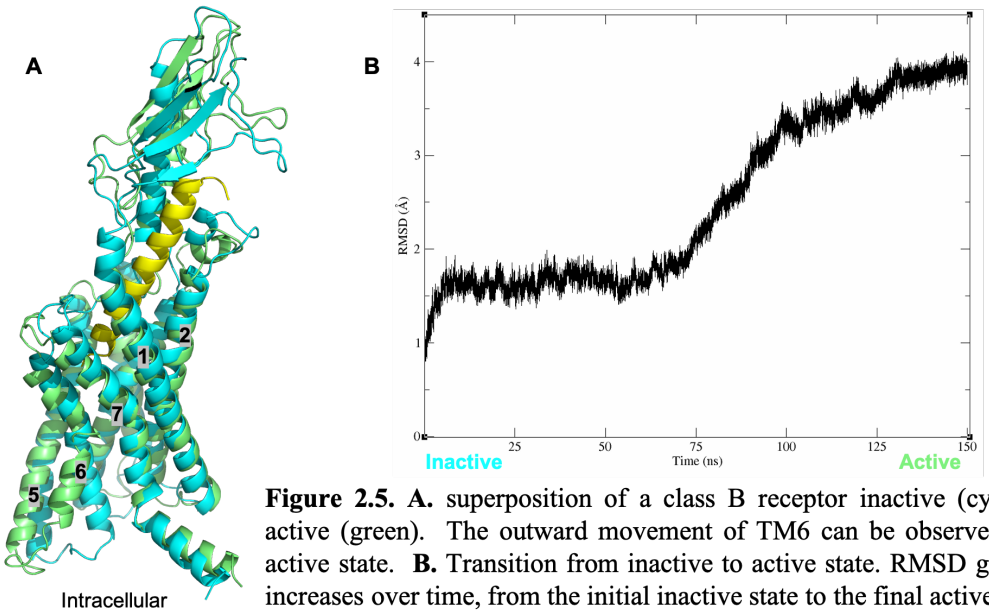


Figure 2.5. A. superposition of a class B receptor inactive (cyan) and active (green). The outward movement of TM6 can be observed in the active state. **B.** Transition from inactive to active state. RMSD gradually increases over time, from the initial inactive state to the final active state.

2.3 Analysis

The output trajectories were visualised and analysed and manipulated using tools embedded in VMD (Humphrey, Dalke, and Schulten 1996), WORDOM (Caflisch et al. 2007), PyMOL (Schrödinger 2015) and Bio3D (Grant et al. 2006).

2.3.1 Trajectory manipulation and analysis

Trajectories were manipulated using WORDOM software command line (Caflisch et al. 2007). This include conversion or extraction of trajectory psf files into pdb files; RMSD and distances measure and monitorisation of the receptor complex was done using WORDOM, provided information about of conformational changes along the trajectory (Caflisch et al. 2007).

2.3.2 Hydrogen bonds

Hydrogen bonds were obtained using VMD plugin HBonds version 1.2 (Luo 2007). A hydrogen bond is formed between an atom with a hydrogen (H) bonded to it, donor (D), and another atom or acceptor (A); and defined as the distance between D-A within 3.5 Å and angle D-H-A less than 25°.

2.3.3 Salt bridges

Salt bridges were searched using VMD Salt Bridges plugin (Villa 2006). The oxygen-nitrogen distance cut-off was defined as 3.2 Å or less. Resulting salt bridges were saved into txt files.

2.3.4 Solvent Accessible Surface Area

Solvent accessible surface area (SASA) was calculated using accessible surface area and accessibility calculation for proteins web server "Accessible surface area and accessibility calculation for protein" version 1.2 from the Centre for Informational Biology, Ochanomizu University (Center for Informational Biology 2012). Solvent accessible residues were the residues that showed a difference of at least 20% between the SASA value ($Å^2$) in the active and inactive state. These residues were followed and studied in detail.

2.3.5 Rotameric angles

Rotameric angles were determined for solvent accessible residues. Rotameric dihedral angles were obtained using torsion.pdb Bio3D package from R (Grant et al. 2006). In order to get rotameric angles, first, water and lipid bilayer were removed from the trajectory, then the trajectory was converted into PDB, files using WORDOM and alanine and Glycine residues were removed from the file. Then, using torsion.pdb Bio3D package from R, each frame was analysed and rotameric angles were saved. Since the frames were analysed in chronological order, it was possible to identify segments with changes in rotameric angles.

2.4 Software

Visualisation, alignment and modelling of the protein complex was done using PyMOL (Schrödinger 2015).

Structure homology-modelling was done using SWISS-MODEL web server (Bienert et al. 2017; Waterhouse, Rempfer, et al. 2018) an automated protein structure homology-modelling server.

OPM database provides spatial arrangements of membrane proteins with respect to the hydrocarbon core of the lipid bilayer (Lomize et al. 2012). OPM was used to orient the receptors into the membrane.

Generation of the input files used for the simulations (equilibration and production), including the building of the lipid bilayer and solvation of the complex, were generated using CHARMM-GUI web server (Lee et al. 2016; Wu, Cheng, et al. 2014).

NAMD – Scalable molecular dynamics software (Phillips et al. 2005; Bernardi et al. 2016) was used to simulate GPCRs complex.

VMD was used for trajectories visualisation and analysis (Humphrey, Dalke, and Schulten 1996). HBonds Version 1.2 plugin (Luo 2007) was used for the detection of hydrogen bonds formed throughout the trajectory. Salt bridges plugin, version 1.1 was used to find the salt bridges throughout the trajectory (Villa 2006).

Manipulation of the trajectory was done using the command line utility WORDOM software (Caflisch et al. 2007).

Colores module from Situs software that enable rigid-body docking of the atomic structure into the density map prior flexible fitting of the atomic structure into the density map (Wriggers 2012; Wriggers and Birmanns 2001).

Conversion of electronic maps from the format DSN6 into ccp4 was done using CCP4i Software (McNicholas et al. 2011; Potterton et al. 2003).

The scripts used rely on the torsion.pdb function from R Bio3D package (Grant et al. 2006) to determine rotameric angles from residues of interest in the trajectory.

Chapter 3: Hybrid GLP-1R models

3.1 Introduction

At the time this PhD project started (February 2017) there were no full-length Class B GPCR structures available - the only experimentally determined structures available for Class B GPCRs were the isolated NTD and isolated TMD structures (Chapter 1). This included the structure of the GLP-1R isolated NTD in complex with various ligands (PDB 3IOL, 3C5T and 3C59; Runge et al. 2008; Underwood et al. 2010). While there was no GLP-1R TMD structure, the closely related glucagon receptor structure had been solved (PDB 4L6R; Siu et al. 2013) but the relative orientation of the two domains was unknown. The situation changed significantly in May 2017, when an active state GLP-1R-Gs complex (PDB code 5VAI; Zhang et al. 2017) became the first full-length Class B structure available; this was a cryo-EM structure of rabbit GLP-1R in complex with a human GLP-1 ligand and a Gs heterotrimeric G-protein stabilised by a nanobody (Nb35). The 150kDa complex had 3.9 Å resolution (Zhang et al. 2017) and represented the first GLP-1R coupled to its natural ligand and its primary G protein. At the same time, two crystallography studies revealing the GLP-1R structures were published (PDB code 5NX2 and 5VEW/5VEX). The X-ray crystal structure of 5NX2 represented a thermos-stabilised human GLP-1R in complex with a non-natural C-truncated peptide agonist (Jazayeri et al. 2017) This structure had 11 stabilising mutations and displayed wild type-like agonist affinity but with reduced potency, suggesting a partially active conformation. It is important to note that the agonist was a C-terminally truncated peptide built with non-natural variants based on GLP-1 N-terminus. The 5VEW/5VEX structures represented the inactive human GLP-1R isolated TMD, without any NTD, to 2.7 Å and 3.0 Å resolution respectively. To stabilise the inactive state ten mutations, including a disulphide bond between helices five and six, and two different negative allosteric modulators were used (Song et al. 2017). These structures provided a toolbox from which full-length models of both the active and inactive states could be reconstructed and simulated in order to understand agonist binding and activation.

3.1.1. Rationale for building and simulating hybrid GLP-1R models

The only *inactive* state structures for the TMD of GLP-1R available at the time this work was carried out were the apo 5VEW/5VEX coordinates, while the only apo structure of the NTD of GLP-1R was part of 5NX2. Therefore, to obtain an approximation of a complete inactive full-length GLP-1R, a hybrid model was constructed by aligning the TMDs of 5VEX and 5NX2 and then melding the NTD of 5NX2 to the TMD of 5VEX as described in the methods.

Despite the 5VAI and 5NX2 structures both representing agonist-bound *active* states, with the characteristic helix six outward movement, there were marked differences that needed to be recognised and resolved in order to model the "best" active agonist-bound state and therefore attempt to understand agonist binding and activation mechanisms. These differences start at the extracellular domain where the position of the NTD relative to the TMD differs between 5VAI and 5NX2. However, given the flexibility between domains in Class B receptors (Yang et al. 2015a), this is perhaps not surprising. The NTD of Class B GPCRs is connected to the TMD by a linker region named as the "stalk". The stalk was first observed in the structure of the human glucagon receptor (PDB 4L6R) as a helical extension of TM1 (Siu et al. 2013). While 5NX2 structure displays a non-helical ten-residue linker sequence, the B-factor from the electron density map shows high degree flexibility of the area (Jazayeri et al. 2017). Similarly, the cryo-EM density map of 5VAI shows low-resolution in this region perhaps hinting at its flexibility (Zhang et al. 2017). The dynamism of the stalk region may facilitate the movement between domains as it guides peptide ligands into the orthosteric binding pocket agreeing with the two-domain-binding mechanism (Inooka et al. 2001). Furthermore, while the GLP-1 ligand straddles the two domains in 5VAI, perhaps limiting their relative movement, the shorter 11-residue ligand in 5NX2 is bound exclusively to the TMD enabling more conformational flexibility between the domains. Hence, given the unnaturally short ligand in 5NX2 and its thermo-stabilised sequence, relative to 5VAI representing a wild type GLP-1R sequence coupled to GLP-1 and G_s, I elected to utilise the NTD conformation from 5VAI in the hybrid model.

A second, and more surprising, difference between 5NX2 and 5VAI involves the conformation of TM6. While both structures display the expected outward movement of TM6 relative to that observed in the inactive 5VEW structure, which

creates the required binding site for α5 of G_s, the observed conformation and position of TM6 differs significantly in each structure. The TM6 in 5VAI is highly distorted forming a kink in the central region and closely packed against the rest of the TMD bundle in the central region (Figure 3.1A), in close agreement to that observed in the structure of the calcitonin receptor TMD in complex with G_s (Figure 3.1B; PDB code 5UZ7; Liang et al. 2017). However, TM6 in 5NX2 is much more regular and linear but is displaced away from the rest of the TM bundle such that there appears to be an unusual "channel" joining the extracellular peptide binding cavity to the intracellular space (Figure 3.1C). This appeared likely to be unstable and, indeed, in the molecular dynamic simulations using 5NX2 without its ligand, the TM bundle fractured between TM5 and TM6 and allowed bilayer lipids to penetrate the receptor core (data not shown). Hence, given this 5NX2 instability, coupled with the good correlation between 5VAI and 5UZ7 (Liang et al. 2017), and also that 5VAI represents a wild type GLP-1R sequence coupled to GLP-1 and G_s, I elected to utilise the TM6 conformation from 5VAI in the hybrid model.

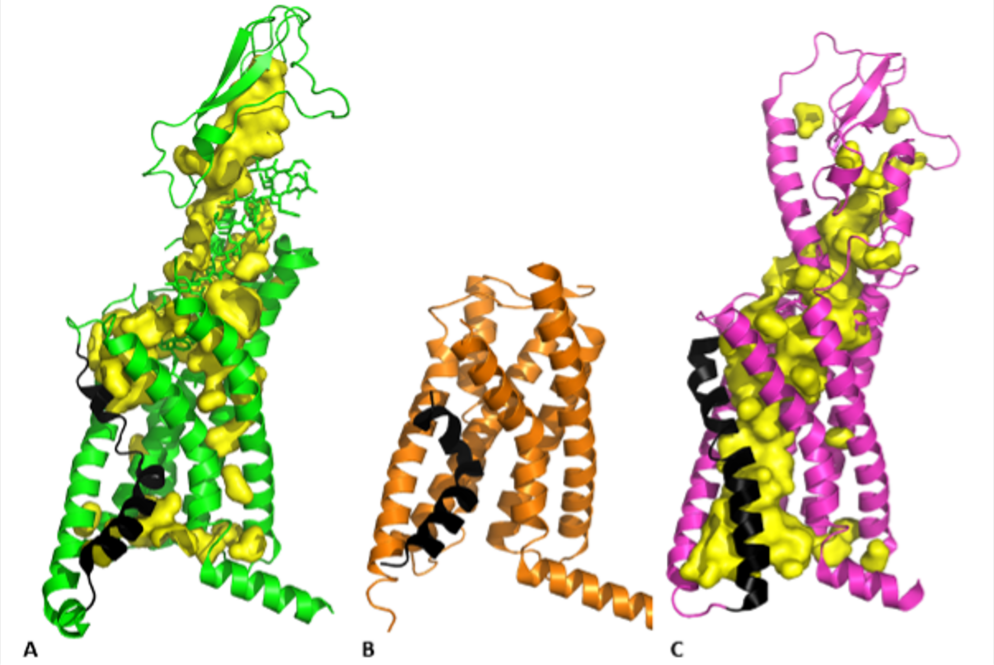


Figure 3.1. Ribbon representation of 5VAI (A), 5UZ7 (B) and 5NX2 (C). Ligands are shown as sticks, TM6 is shown in black, and the cavity space is shown for A and C as yellow. TM6 is highly distorted in 5VAI and 5UZ7 and bends inwards at the centre to pack closely against the rest of the helical bundle. On the other hand, TM6 in 5NX2 is more regular and leaves an unusual cavity within the helical bundle that joins the cytoplasmic side to the extracellular binding pocket.

Nevertheless, there are some aspects of the 5VAI structure that appear to be potentially less reliable than 5NX2. In particular, the Cα positions of Arg380$^{7.34b}$ and Lys288$^{7.37b}$ at the extracellular end of TM7 are displaced by one helical turn between the two structures (Figure 3.2). While in 5NX2 they are facing towards the ligand and Arg380$^{7.34b}$ is potentially interacting with Asp15* as expected from site-direct mutagenesis data and previous modelling studies (reviewed in de Graaf et al. 2016), and Chapter 1), the refined B-factor at the top of TM6, including Arg380$^{7.34b}$, suggests a high degree of flexibility (Jazayeri et al. 2017). In the 5VAI structure, Arg380$^{7.34b}$ and Lys288$^{7.37b}$ point away from the TM bundle and would hence interact with the lipid head-group region of the bilayer. As a consequence, the contacting regions on TM1, part of TM2 and extracellular end of TM7 from 5NX2 in the hybrid model were used in order to maintain helix-helix contacts (Figure 3.3). While such hybrid modelling represents the best approach (in the absence of reliable high-resolution structures) for creating a starting model for molecular dynamics simulations, there are clearly likely to be a number of approximations and associated error in the model.

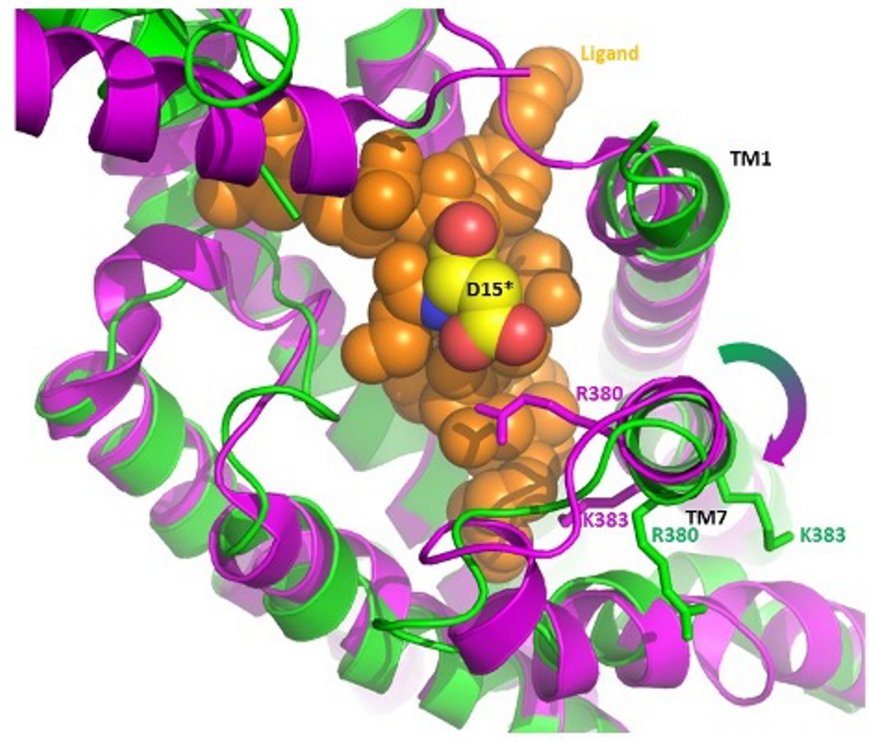


Figure 3.2. Comparison of 5NX2 (magenta with agonist as space-fill) and 5VAI (green). While Arg380$^{7.34b}$ and Lys383$^{7.37b}$ face the ligand in 5NX2, in 5VAI they are displaced by 100° and face the lipid head-group region of the bilayer.

The PDB structures used as template had some missing residues or sidechains, including the stalk region. Since the template structure coordinates and target amino acid sequence are known, target-template alignment mode 'User Template Mode' from SWISS-MODEL web server (Waterhouse, Rempfer, et al. 2018), were SWISS-MODEL homology server extracts the initial structural information from the template and then the sequence is aligned to conserved residues. The missing residues are then modelled using SWISS-MODEL rotamer library. Finally to avoid unfavourable interactions, the structure undergoes energy minimisation.

Furthermore, the newly created receptor-ligand interactions and the melding points between the different fragments used to create the hybrid models may have resulted in less-than-optimal interactions and conformations, although long relaxation times during the equilibration stages were used, which should have fixed local inaccuracies. Nevertheless, it is important to interpret the output carefully in light of the mutagenesis and molecular pharmacological data in order to identify meaningful outcomes.

3.1.2. Aims and Objectives of Chapter 3

In order to fully understand how agonists bind and activate the human GLP-1R, I set out to build a high quality, complete, active-state, human GLP-1R model, docked with GLP-1, using a combination of both the 5VAI and 5NX2 structures as templates. For the active model, most of 5VAI structure was used as it represented an active GLP-1R in complex with the G protein and ligand, providing information about the interaction between ligand and receptor as well residues interactions. However, the position of residues in helix 1 and extracellular end of helix 7 region of 5VAI receptor appear in disagreement with previous mutagenesis data and modelling studies (reviewed in de Graaf et al. 2016), therefore the homologous residues in 5NX2 were used to build the hybrid active model. This hybrid model was subjected to molecular dynamics simulations to refine the model and ascertain the dynamic details of the peptide-receptor interaction.

A second objective was to generate a complete inactive wild-type human GLP-1R model, by combining the inactive TMD of 5VEW with the empty NTD of

5NX2, and then to examine the stability of this model using molecular dynamics in order to compare the inactive conformation to that of the agonist-bound GLP-1R.

3.2 Methods

All molecular model building manipulations were carried out using the tools embedded within PyMOL (Schrödinger 2015) unless otherwise stated.

3.2.1. Building active GLP-1R

(i) 5VAI and 5NX2 were structurally aligned by superimposing the residues of their TMDs using the align function in PyMOL. (ii) All residues in 5NX2 were deleted, apart from Glu139$^{1.34b}$-Lys197$^{2.67b}$ and Glu373^{ECL3}-Phe393$^{7.48b}$. Then this region was aligned to their homologue residues in 5VAI (iii) The G-protein, nanobody, ligand and regions Glu139$^{1.34b}$-Lys197$^{2.67b}$ and Glu373^{ECL3}-Phe393$^{7.48b}$ were deleted from the 5VAI structure. (iv) The remaining atoms from both molecules were merged and was then used as a template to build a model of human GLP-1R using the homology modelling web server SWISS-MODEL (Waterhouse, Bertoni, et al. 2018; Bienert et al. 2017). Residues Ser129^{ECD} to Glu138$^{1.33b}$, connecting the NTD and TMD, were built by SWISS-MODEL, as were any missing or altered side chains in the template. All other residues were built from the template and matched the starting conformations. The output PDB file from SWISS-MODEL represented the apo hybrid model of active human GLP-1R from Thr29^{ECD} to Arg421^{CTT}. (v) GLP-1 and α5 of G_s were added back into the model by structurally aligning the SWISS-MODEL output with 5VAI, extracting the co-ordinates for the GLP-1 and α5 (374-394) segments from 5VAI, and merging them with the active GLP-1R hybrid model.

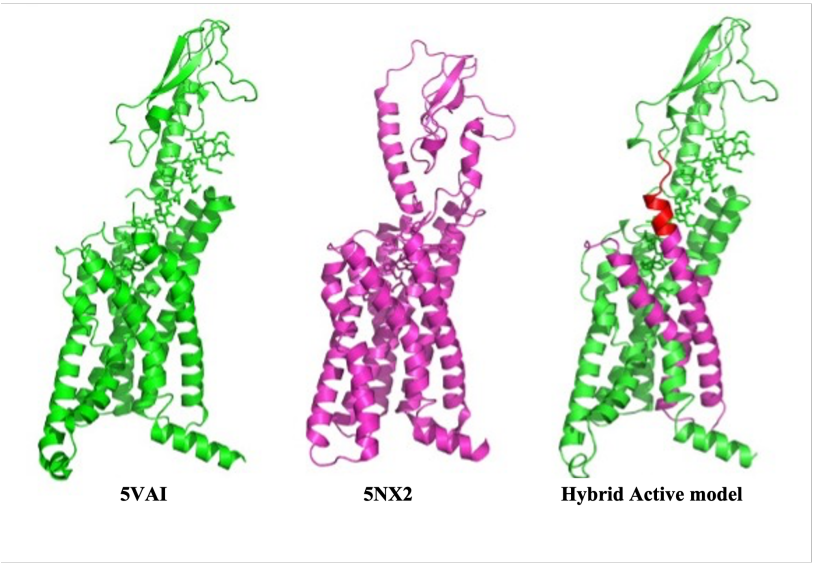


Figure 3.3. The hybrid active state model of human GLP-1R was built using 5VAI (green) and 5NX2 (magnenta). Missing regions were built by SWISS-MODEL (red).

3.2.2. Building inactive GLP-1R

The NTD from 5NX2 represents the only apo structure of this domain and was hence selected as the NTD to fuse with the TMD from 5VEW. (i) One monomer from the crystallographic dimer of 5VEW was deleted, as was the fusion partner. (ii) The remaining residues from 5VEW were structurally aligned with 5NX2 by superimposing the residues of their TMDs using the align function in PyMOL (Figure 3.4). (iii) All residues in the TMD of 5NX2 were deleted, leaving only the NTD residues from Thr29^{ECD} to Glu133$^{1.28B}$ and also Met204^{ECL1} to Asp215^{ECL1} (the latter is the part of ECL1 which contacts the NTD on 5NX2 but is missing in 5VEW). (iv) This was used as a template to build a model of inactive human GLP-1R using SWISS-MODEL. All the missing regions in the template were built by SWISS-MODEL, as were all missing or mutated side chains. All other residues built from the template matched the starting conformations. The output represented the apo hybrid model of inactive human GLP-1R from Thr29^{ECD} to Arg421^{CTT}.

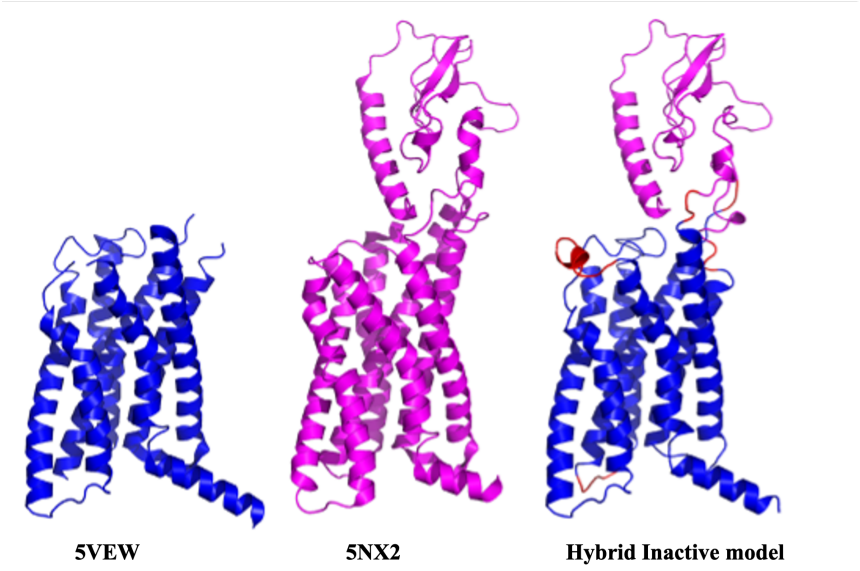


Figure 3.4. Inactive GLP-1R built from 5VEW (blue) using the NTD from 5NX2 (magenta). Missing regions were built by SWISS-MODEL (red).

3.2.3. Molecular Dynamics

3.2.3.1. System building

The starting models were first orientated with respect to the hydrocarbon core of the lipid bilayer utilising the OPM database and server (Lomize et al. 2012) and used to build a protein/membrane complex and generate the inputs necessary for molecular dynamics simulation using CHARMM-GUI (Jo, Kim, and Im 2007; Brooks et al. 2009). The four disulphide bonds of GLP-1R were first defined. A simple POPC bilayer has been previously used to simulate a model of GLP-1R (Wootten et al. 2016). For this simulation, the models were embedded into a similar 90 x 90 $Å^2$ simple lipid bilayer containing POPC lipids before a water/solvent box was added (TIP3P and 0.15 M NaCl). Realistic lipid membranes are challenging for *in silico* modelling because not all lipids force field parameters have been tested against experimental data. The resulting lipid bilayer build by CHARMM-GUI and used in the simulations had a good correlation between the OPM lipid boundary predictions, as phospholipid head groups were just outside the predicted OPM

boundaries (Figure 3.5). Simulations were all set at 303.15 K temperature and 1 atm pressure, using Langevin dynamics with a damping coefficient of 1/ps to control the temperature and a Nose-Hoover Langevin piston to control the pressure.

3.2.3.2. System simulation

Active and inactive models were subjected each to equilibration (70 ns) and production (385 ns for inactive state and 460 ns for active state) using NAMD (Phillips et al. 2005) with NPT ensemble and CHARMM36 force field (Huang and MacKerell 2013). A time-step of 2 fs was used, and Particle Mesh Ewald to account for long-range electrostatics.

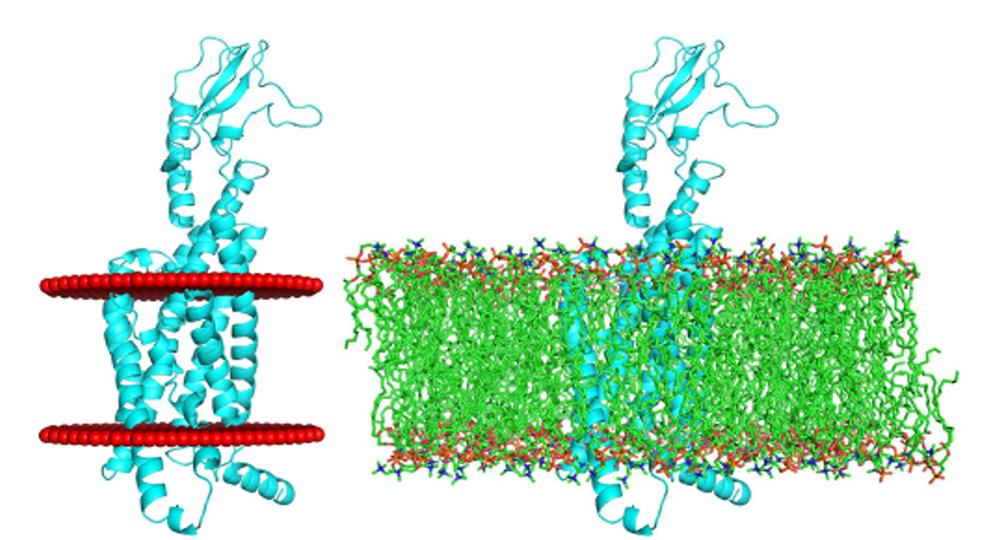


Figure 3.5. The active state model of human GLP-1R shown in ribbon form (cyan) with the hydrophobic bilayer boundaries as predicted by OPM server (Lomize et al. 2012) as red space filled dummy atoms (left) and the POPC bilayer built by CHARMM-GUI (Jo, Kim, and Im 2007; Brooks et al. 2009). In stick form (right). The carbon atoms (green) depict the hydrophobic region of the bilayer, while the phosphate, oxygen and nitrogen atoms (orange, red, blue) form the phospholipid head groups just outside the OPM boundary. Since the model was pre-orientated with respect to the membrane normal (Z axis), there was very good correlation between the OPM lipid boundary prediction and the POPC bilayer built by CHARMM-GUI.

3.2.3.3. Trajectory analysis

Output trajectories were analysed using WORDOM and VMD (Humphrey, Dalke, and Schulten 1996; Caflisch et al. 2007). The simulation trajectory for the active hybrid model was analysed in detail for between 230-370 ns and the simulation of the inactive hybrid model trajectory was analysed between 200-380ns as during these segments each structure reached stability. The analysis was done using VMD in order to estimate the fraction of time in which all residue-residue hydrogen bonds were formed (defined as D-A distance of 3.5 Å or less, with an angle between D-H-A of 25° or less).

3.3. Results and Discussion

3.3.1. Trajectory analysis

Due to the higher resolution of the single template structure used to build the TMD of the inactive model, relative to the two lower resolution structures used to build the active hybrid TMD, the former reached stability earlier in the MD simulations with a positional root mean square deviation (RMSD) relative to the TMD of starting model plateauing at 60 ns and remaining stable throughout the remainder of the 385 ns simulation (Figure 3.6C). The active hybrid model took longer to reach stability but nevertheless the RMSD plateaued after 230 ns (Figure 3.6C).

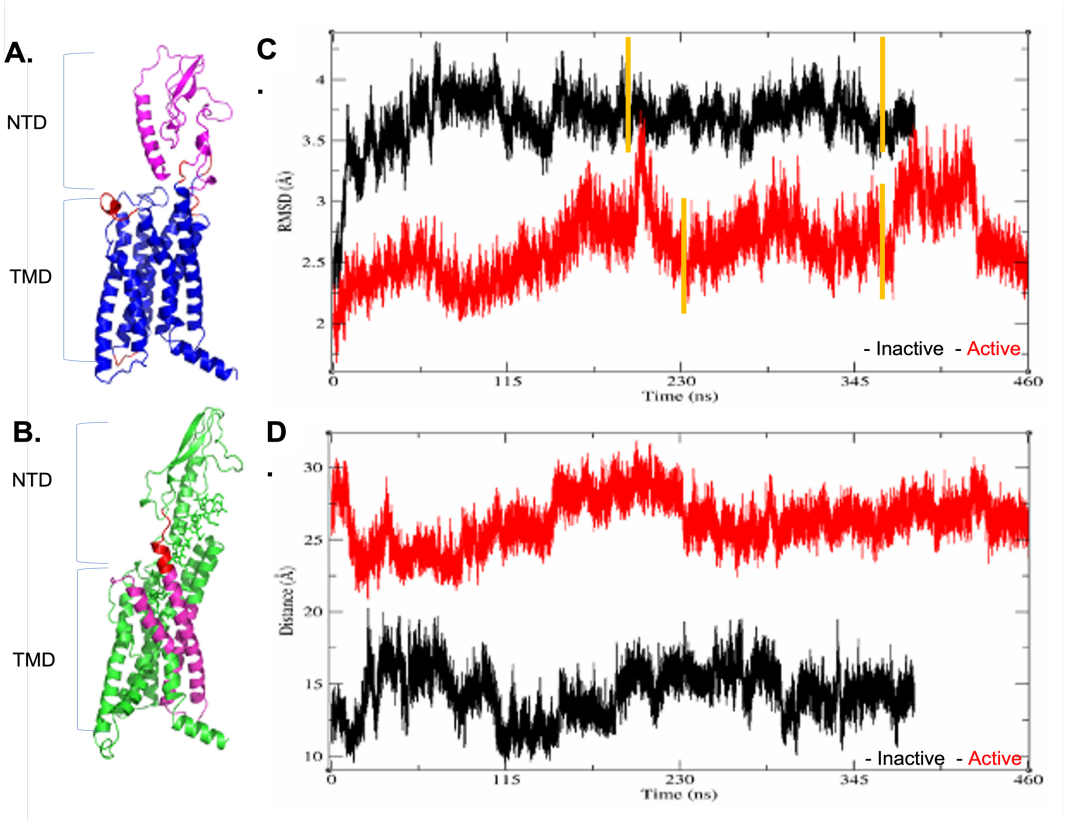


Figure 3.6. Ribbon representation of the hybrid inactive (**A**) and hybrid active (**B**) starting models with GLP-1 shown in stick form in the active state. Regions derived from 5VEW (blue), 5NX2 (magenta), 5VAI (green), and built by SWISS-MODEL (red). **C.** Root mean square deviation (RMSD) of the TMD region relative to the starting model for the simulations of the hybrid GLP-1R models - inactive (black, 385 ns) and active (red, 460 ns). Yellow line limits the segment of the trajectory that was analysed. **D.** Distance between the Cα atoms of Arg-176 (TM2) and Arg-348 (TM6) – active state (red) and inactive state (black).

The principal difference between the inactive and active conformations of GPCRs is the outwards movement of the cytoplasmic half of TM6 away from the transmembrane bundle in order to create the binding site for α5 of the G protein (Rasmussen et al. 2007). Such movement can be observed when comparing the inactive 5VEX and 5VEW structures with those of the active conformations 5VAI and 5NX2 (Jazayeri et al. 2017; Song et al. 2017; Zhang et al. 2017). With the outward movement of the intracellular end of TM6 to create a cavity for binding of the G protein, the distance between TM2 and TM6 increases. Therefore, while the distance between the Cα atoms of residues Arg176$^{2.46b}$ in TM2 and Arg348$^{6.37b}$ in TM6 is 23.7 Å in 5NX2 and 25.3 Å in 5VAI, it is only 12.1 Å in 5VEW. The starting distance in the active state hybrid model was 23.7 Å while it was 12.1 Å in the inactive

state of the hybrid model; the distances were maintained throughout the simulations of both our inactive and active models (Figure 3.6D).

Although not observed in the simulations, the outward movement of the cytoplasmic end of TM6 upon activation appears to be accompanied by a large perturbation of the helix at Gly361$^{6.50b}$ and a movement of the segment of TM6 incorporating His363$^{6.52b}$ and Glu364$^{6.53b}$ towards the cytoplasmic side (Cα movements of 5.5 Å both from the simulated inactive GLP-1R to the simulated active GLP-1R at 340 ns each – Figure 3.7). In addition, TM7 tilts towards TM6 at the extracellular end, with the hinge being at Gly395$^{7.50b}$, with an additional clockwise twist of the helix to move Arg380$^{7.34b}$ and Arg383$^{7.37b}$ from the exterior of the protein into the core of the receptor where they interact directly with the ligand (Figure 3.8).

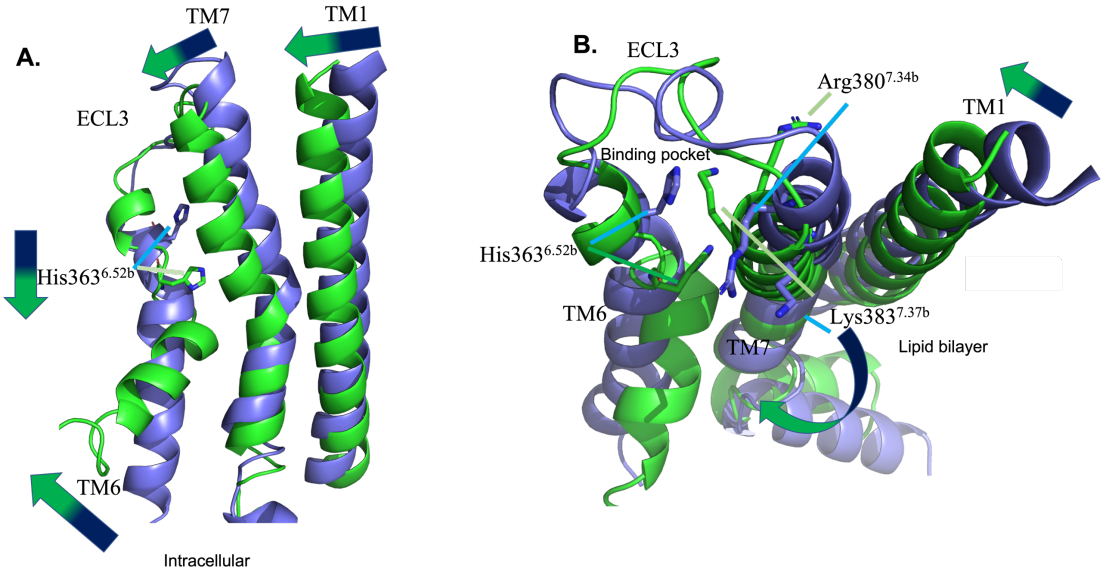


Figure 3.7. Alignment of the hybrid GLP-1R comparing the position of TM1, TM6 and TM7 in the active state (green) and the inactive state (blue) at 340 ns of simulation each. **A.** Lateral view, arrows describe the direction of the movement; in this view the shift of TM1 over TM7 and TM7 over TM6 can be observed, as well the outward movement of the intracellular segment of TM6 in the active state (green) relative to the inactive state (blue) and the movement of His363$^{6.52b}$ (stick representation) towards the intracellular side. **B.** Extracellular view displaying the rotation of the extracellular end of TM7; rotation of residues Arg380$^{7.34b}$ and Lys383$^{7.37b}$ (stick representation of residues) from the lipid bilayer towards the binding pocket can be observed in the active state (green) relative to the inactive state (blue).

ECL3, which links the top of TM6 and TM7, undergoes a significant conformational rearrangement from the initial model to the final simulated structure as a result of TM6 and TM7 helical movements. The space created by the movement of TM7 towards TM6 is filled by the tilting of the extracellular side of TM1 around

Gly361$^{6.50b}$. The final consequence is the outward movement of TM6 at the cytoplasmic end where the G protein docks. In the simulations, residues 382-394 region of the α5 helix of Gs remained stable and helical through the 460 ns active state simulation (Figure 3.8).

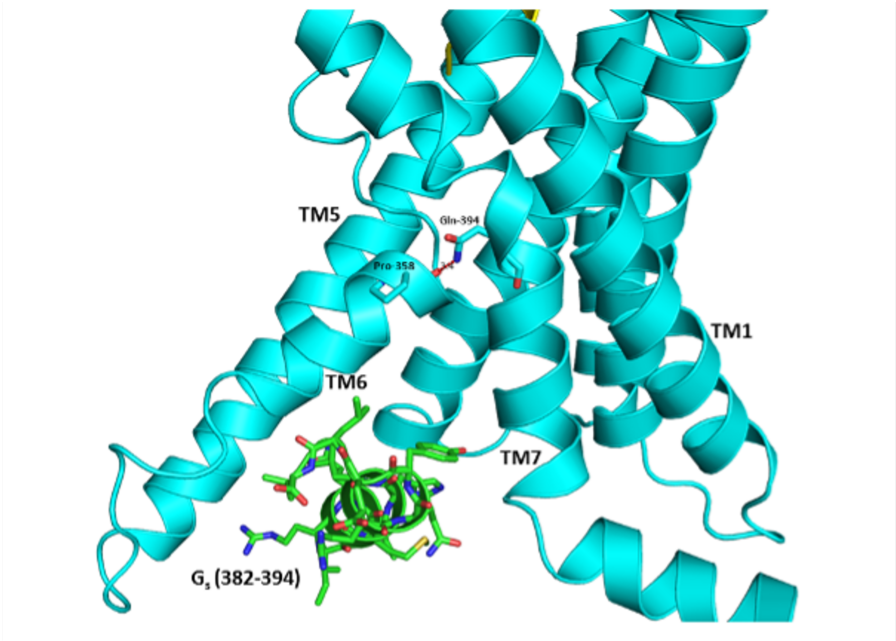


Figure 3.8. Snapshot of the active (cyan) model at 460 ns. Stick representation of residues display oxygen (red), nitrogen (blue), and hydrogen bonds (red dash). At the middle region of the receptor interaction between Gln394$^{7.49b}$ (helix 7) and Pro358$^{6.47b}$ (TM6) possible help stabilising TM6 kink formation. In the intracellular region the α5 helix (residues 382-394) of Gs remained stable and helical through the simulation of the active state.

In the inactive hybrid model, Arg190$^{2.60b}$ was able to remain hydrogen bonded with Gln394$^{7.49b}$ for 40% of the trajectory (Figure 3.9A, blue). However, as it can be seen, this interaction was abolished by ligand binding (Figure 3.9A, green) since Arg190$^{2.60b}$ (yellow) instead interacted with Glu9* of the ligand for 84% of the trajectory. In addition, Glu9* also interacts with the side chain hydroxyl groups of Tyr148$^{1.43b}$ and Tyr152$^{1.47}$ (78% and 76% respectively), both residues that have been identified as being involved in GLP-1 recognition (reviewed in de Graaf et al. 2016). The movement of Arg190$^{2.60b}$ to interact with Glu9* results in Gln394$^{7.49b}$ being free to move (Figure 3.9A) and interact with the exposed main chain carboxyl oxygen of Pro358$^{6.47b}$ of TM6 (Figure 3.9B), providing a helix cap for the cytoplasmic half of

this helix following its disruption at Gly361$^{6.50b}$ (Figure 3.9B). This represents a way in which agonist binding could stabilise one type of active state of GLP-1R. GLP-1 potency is sensitive to Arg190$^{2.60b}$ substitution, as is a synthetic variant of oxyntomodulin in which the native Gln at the third position was replaced by Glu (Wootten, Reynolds, Koole, et al. 2016). Given that native oxyntomodulin (and synthetic Q9-GLP-1) can activate the cAMP pathway but are not sensitive to Arg190$^{2.60b}$ substitution (Wootten, Reynolds, Koole, et al. 2016), it is likely that they stabilise a different active state which does not rely upon direct Arg190$^{2.60b}$ binding or the breakage of the Arg190$^{2.60b}$ / Gln394$^{7.49b}$ interaction.

A second clear change in hydrogen bond formation resulting from ligand binding can be observed with Arg380$^{7.34b}$. In the inactive model, this residue forms a very stable hydrogen bond with Glu373^{ECL3} throughout the trajectory (96%). Upon ligand binding, Arg380$^{7.34b}$ undergoes a significant movement to interact with Asp15* of the ligand (77%) (in agreement with mutagenesis data (Dods and Donnelly 2015), while the interaction with Glu373^{ECL3}is also maintained (82%), albeit requiring a substantial conformational rearrangement of ECL3 (Figure 3.9D).

Hence the movement of Arg380$^{7.34b}$ towards Asp15* of the ligand encourages the rearrangement of ECL3, which is ultimately an essential requirement for the movement of TM6 and the creation of a Gs binding site. The position of Arg380$^{7.34b}$ is conserved as a positively charged residue (Arg or Lys) in the receptors for glucagon, GLP-2 and GIP, all of which also have a negatively charged residue (Asp or Glu) in the ligand at the position equivalent Asp15* in GLP-1. Indeed, the oppositely charged residues have been interchanged from ligand to receptor in an elegant study by Moon et al. (Moon et al. 2015) showing a reciprocal rescue which strongly implicates the two residues in an interaction. Moon et al. replaced Arg380$^{7.34b}$ with Asp, resulting in almost a 2000-fold reduction in potency. However, Arg9*-GLP-1, which had almost 100-fold lower potency at wild-type GLP-1R, was shown to have 120-fold improved potency at the Arg380$^{7.34b}$ –Asp mutant receptor.

Figure 3. 9. Comparison of interactions between active and inactive hybrid GLP-1R

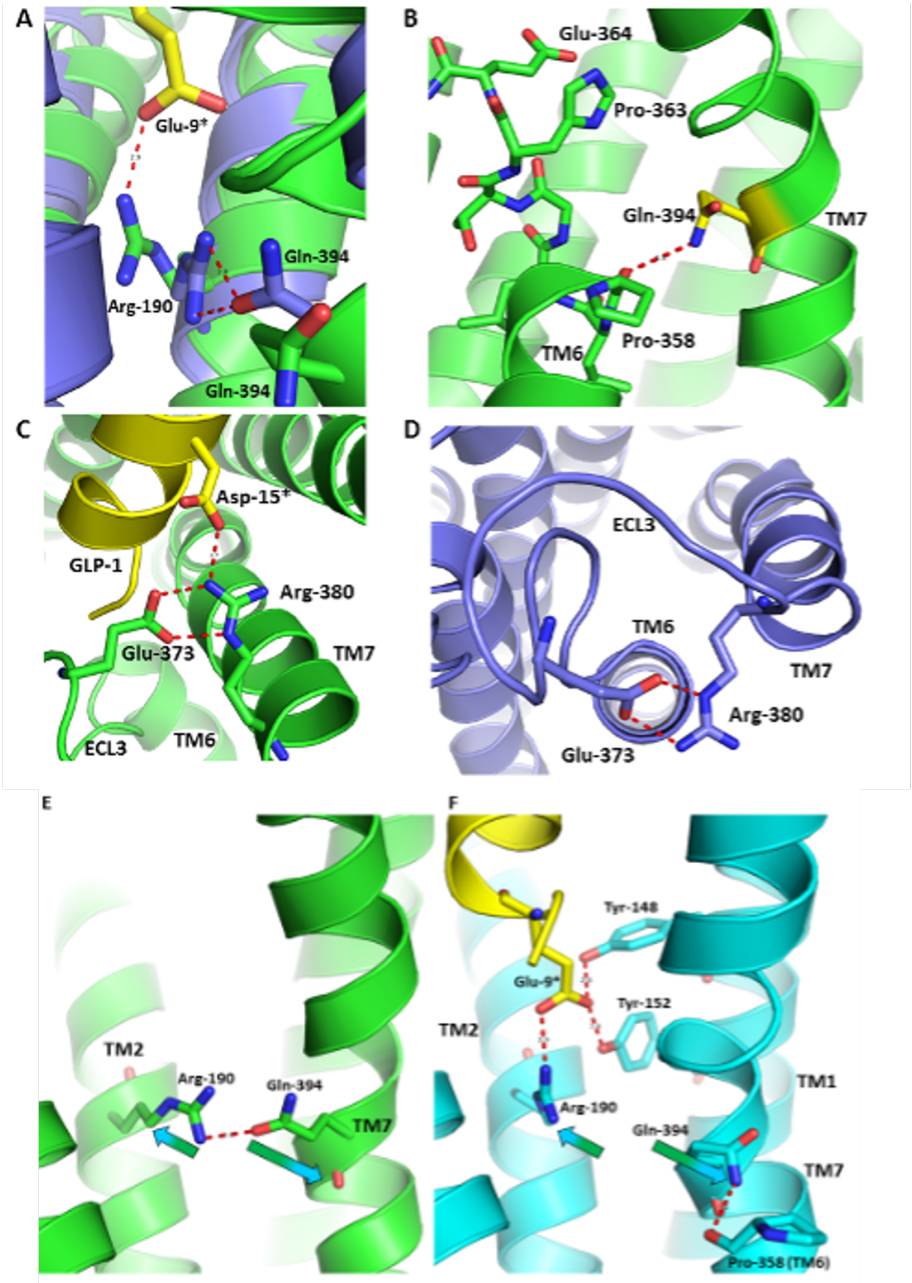


Figure 3.9. Snapshots of the inactive (blue) and active (green) models taken at 285 ns and 325 ns respectively with stick representations residues showing oxygen (red), nitrogen (blue), and hydrogen bonds (red dashes). **A.** Overlay of both models showing Arg190$^{2.60b}$ and Gln394$^{7.49b}$ interaction in the inactive state, whereas in the active state Arg190$^{2.60b}$ instead interacts with Glu9* of GLP-1 (yellow), freeing Gln394$^{7.49b}$ to move and interact elsewhere. **B.** Gln394$^{7.49b}$ (yellow) in the active state provides an N-cap for the cytoplasmic half of TM6 at the carbonyl oxygen of Pro358$^{6.47}$. **C.** Arg380$^{7.34b}$ in the active receptor points towards the centre of the helical bundle and interacts with both Glu373^{ECL3} and Asp15* of GLP-1. **D.** Arg380$^{7.34b}$ in the inactive receptor points away from the helical bundle and interacts with Glu373^{ECL3}. The ligand-induced movement of Arg380$^{7.34b}$ may be responsible for rearranging ECL3. **E – F.** Arg190$^{2.60b}$ to Gln394$^{7.49b}$ interaction seen in the inactive state is broken as Arg190$^{2.60b}$ binds Glu9* of GLP-1, allowing Gln394$^{7.49b}$ to form a hydrogen bond with Pro358$^{6.47b}$, possibly stabilising TM6 active conformation.

A third arginine residue (Arg299^{ECL2}) also has substantially different interactions in the active and inactive model trajectories. Despite starting off pointing downwards towards the receptor core as in the 5VAI template, Arg299^{ECL2} rapidly moved out of the binding pocket and then interacted with both Glu294^{ECL2} (57%) and Glu21* (40%) of the ligand (3.10A) as previously predicted from earlier mutagenesis and modelling studies (Dods and Donnelly 2015). Indeed, Arg299^{ECL2} was placed in this external position in the 5VEX X-ray structure, adding further evidence towards this conformation in the active state.

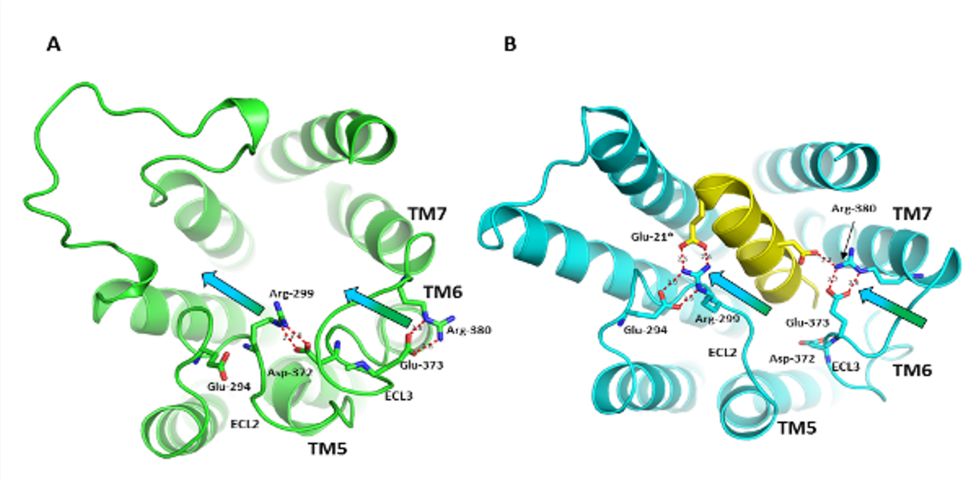


Figure 3.10. Snapshots of hybrid models, inactive receptor (green) and active receptor (cyan) with GLP-1 (yellow), taken at 285 ns and 325 ns respectively. Arg380$^{7.34b}$ and Glu373^{ECL3} interact and move towards the ligand, with Arg380$^{7.34b}$ interacting with Asp15*. Arg299^{ECL2} also moves towards the ligand and interacts with Glu21* (and Glu294^{ECL2}). The Arg299^{ECL2} to Asp372^{ECL3} interaction seen in the inactive conformation is broken, as well as numerous other ECL2-ECL3 interactions.

In the inactive model, Arg299^{ECL2} forms a very stable interaction with both Glu34^{ECD} (97%) on the NTD, and with Asp372^{ECL3} (97%) of ECL2. Indeed, a characteristic of the active state is the close interaction between ECL2 and ECL3 – for example, Glu294^{ECL2}, Asn300^{ECL2} and Asn304^{ECL2} interact with Asp372^{ECL3} and Arg376^{ECL3} (Figure 3.10 B). This inter-loop interaction is absent in the active state, which would allow the opening of the binding pocket towards the ligand.

3.3.2. Ligand binding in active conformation

Glu9* is clearly a critical residue, interacting with Arg190$^{2.60b}$ (84%), Tyr148$^{1.43b}$ (78%) and Tyr152$^{1.49b}$ (76%), and to a lesser extent with Thr391$^{7.44B}$ (16%) during the segment of the simulation analysed. Likewise, Arg380$^{7.43B}$ is important since it interacts with Asp15* (77%). Additional interactions are Thr13* with Lys297^{ECL2}, and Glu21* with Arg299^{ECL2} – both residues have been shown to be involved in agonist recognition (reviewed in de Graaf, Donnelly et al. 2016). However, the key residue for agonist-induced receptor activation in GLP-1R is the N-terminal His7*. In the early stages of the simulation, the positively charged N-terminal moiety interacted with Glu9*, forming a four-way interaction with Arg190$^{2.60b}$ and Glu364$^{6.53b}$. However, this intra-ligand salt bridge eventually broke and consequently the distance between Arg190$^{2.60b}$ and Glu364$^{6.53b}$ increased. While Arg190$^{2.60b}$ continued to interact with Glu9*, the N-terminal amine of positive charged His7* interacted with both Glu364$^{6.53b}$ (78%) and Glu387$^{7.42b}$ (66%) (Figure 3.11). It is interesting to note that, while there have been a number of mutations of these residues (reviewed in de Graaf et al. 2016), it was only the double substitution by Yang et al. (Yang et al. 2016) which abolished ligand binding, suggesting that either of these Glu residues can partially compensate for the other in the less deleterious single mutations. The sharing of the interaction with the positively charged amino-terminus of GLP-1 by these two acidic side chains forms the centre of a network of salt bridges and hydrogen bonds involving several residues known to be critical for full agonist recognition (Asn300^{ECL2}, Trp306$^{5.36b}$, Arg310$^{5.40b}$, Asp372^{ECL3}, and Lys383$^{7.37b}$).

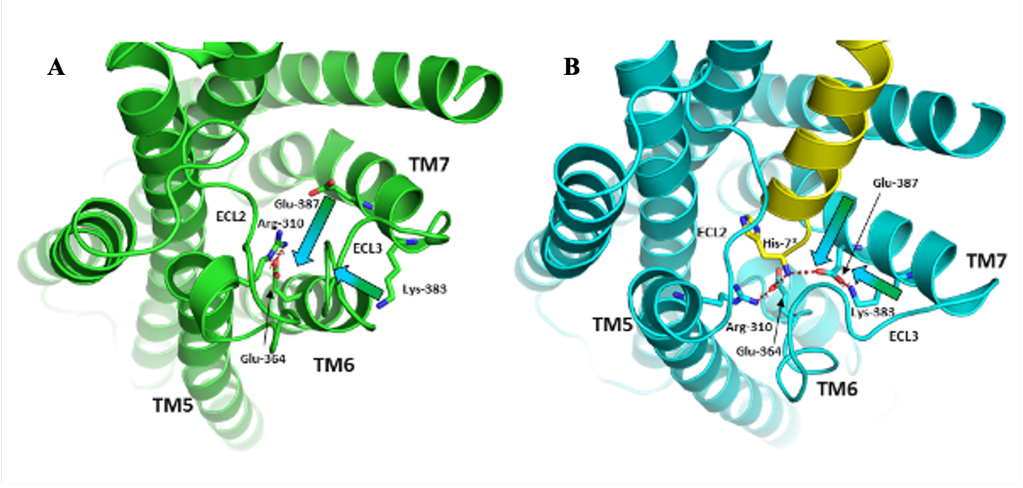


Figure 3.11. Snapshots of hybrid models, inactive receptor (green) and active receptor (cyan) with GLP-1 (yellow), taken at 285 ns and 325 ns respectively. **A.** Interaction of Lys383$^{7.37b}$ and Glu387$^{7.42b}$, interacting directly with the N-terminal nitrogen of His7*. **B.** Lys383$^{7.37b}$ and Glu387$^{7.42b}$ are distant in the inactive conformation. The Arg310$^{5.40b}$ to Glu364$^{6.53b}$ interaction observed in the inactive state is reformed in the active state but Glu364$^{6.53b}$ moves towards the ligand and also interacts directly with the N-terminal nitrogen of His7*.

It is interesting to note that Arg310$^{5.40b}$ and Lys383$^{7.37b}$ both directly stabilise Glu364$^{6.53b}$ and Glu387$^{7.42b}$, respectively, but display no direct interactions with GLP-1. This is in keeping with mutagenesis and pharmacological data which showed their mutation caused minimal disruption of GLP-1 binding affinity compared with much more significant impairment of GLP-1 mediated receptor activation (Dods and Donnelly 2015).

3.4. Conclusions

This chapter has demonstrated and rationalised the importance of interactions between (i) Asp15* and Arg380$^{7.34b}$, rotating TM7 and re-configuring ECL3 by pulling Glu373^{ECL3} towards the ligand (Figure 3.10 A and B); (ii) Glu21* and Arg299^{ECL2}, freeing Asp372^{ECL3} and contributing further to the ECL3 reconfiguration (Figure 3.10 A and B); (iii) His7* with Glu364$^{6.53b}$ (stabilised by Arg310$^{5.40b}$) and with Glu387$^{7.42b}$ (stabilised by Lys383$^{7.37b}$) – the movement of Glu364$^{6.53b}$ is possibly a critical characteristic of the movement and distortion of TM6 during the formation

of the active state (Figure 3.11); and (iv) Glu9* and Arg190$^{2.60b}$, freeing Gln394$^{7.49b}$ to stabilise the distortion in TM6 (Figure 3.8, 3.9). By acting together, these various ligand-receptor interactions result in the stabilisation of an active receptor state in which TM6 moves to enable Gs to bind (Figure 3.8). Since N-terminally truncated GLP-1 can activate GLP-1R at high ligand concentrations in recombinant systems (Donnelly 2012), it is likely that interactions (i) and (ii) are sufficient to generate the active state through the Asp-15*/Glu21*-mediated disruption of ECL3. However, clearly the additional interactions via His7* make this transition substantially more efficient.

In summary, this work has demonstrated that fragments of several GLP-1R structures can be used to create stable and meaningful hybrid receptor models which can be simulated and analysed to answer important questions linking structure to function, and to propose agonist-mediated receptor activation mechanisms. The work described in later chapters was based on later improved structures which were published after the work described here was complete. The work described in Chapter 3 was published in early 2018 (Gómez-Santiago, Paci, and Donnelly 2018).

In the next months, and since then, new structures of GLP-1R became available representing an opportunity to update and improve the work presented in this chapter.

Chapter 4: Simulation of Cryo-EM based models of GLP-1R

4.1 Introduction

The number of solved GLP-1R has increased since 2017, when the first cryo-EM structure of an activated full-length GLP-1R (PDB 5VAI) became available alongside the crystal structures of a thermostabilised active GLP-1R in complex with a truncated peptide agonist (PDB 5NX2) and the TMD of the human GLP-1R in the inactive state (PDB 5VEW/5VEX) (Zhang et al. 2017; Jazayeri et al. 2017; Song et al. 2017). The work described in Chapter 3 was carried out in 2017 and submitted and published in early 2018 (Gómez-Santiago, Paci, and Donnelly 2018). Shortly after submission, a new cryo-EM structure became available of the human active GLP-1R in complex with $G\alpha_s$ and a Gs-biased agonist peptide, exendin-P5, at a global 3.3 Å (PDB code 6B3J; Liang et al. 2018). This new 6B3J structure represented a full-length human GLP-1R in the active state, with a better resolution from the previous rabbit GLP-1R PDB code 5VAI. There were differences between the two structures, especially on the extracellular end of TM1, TM6 and TM7 and ECL3, which interestingly was largely in agreement with the hybrid model used in Chapter 3. In the TMD the backbone conformation of TM1 was conserved but differed in the overall position of the sidechains. The intracellular structures and the engagement with $G\alpha_s$ -α5 helix was also different, possibly reflecting the biased nature of the agonist. The limited density at the intracellular end of TM5 and ICL3 suggested a more flexible region in the agonist-bound GLP-1R than in previous GLP-1-bound receptor. Similar to 5VAI, 6B3J was missing residues belonging to the stalk region (S129-S136) and ICL3 (N338-T343). As this receptor was complexed with a biased peptide, exendin-P5, activation mechanisms and interactions are expected to be subtly different from the endogenous GLP-1 ligand. Nevertheless, at this time in the project, this new *active state* 6B3J structure represented the best available data for the study of agonist binding and activation of family B GPCRs requiring the work in Chapter 3 to be updated and improved.

What was also needed was an *inactive* state structure of GLP-1R. However, at this time in the project, the experimentally determined structures of GLP-1R in the inactive state only showed the isolated TMD, leaving NTD conformation and position relative to the TMD as well interaction with the ligand uncertain. Nevertheless, a full-length peptide-bound glucagon structure in the inactive state became available at 3.0 Å resolution (PDB code 5YQZ; Zhang et al. 2018). Although glucagon has physiologically opposite role to GLP-1R in glucose homeostasis (Creutzfeldt et al. 1996), both receptors have a high sequence similarity and are structurally homologue; therefore, the glucagon structure could be used as the initial template for GLP-1R homology modelling. The earlier crystal structure of full-length inactive GCGR structure (PDB code 5XF1; Zhang, Qiao, et al. 2017) in complex with an antibody, mAb1 at 3.19 Å resolution had been available earlier but the ECD orientation relative to the TMD was not compatible with the two-steps binding. In 5XF1 structure the antibody is found interacting with the αA helix and loops L2, L4 and L5 of the GCGR NTD, behaving as an antagonist as it blocks the access of the ligand into the binding pocket and restricting the conformational flexibility between the ECD and the TMD. The new inactive state 5YQZ structure was in complex with a glucagon analogue, the partial agonist NNC1702, providing insight of the ligand-receptor interaction and the relative orientation of the ECD to the TMD prior to activation, providing an opportunity for the study of receptor-ligand interaction in the inactive state (Figure 4.1). The inactive 5YQZ structure showed a similar binding site, for example the ligand (Asp9*) -receptor TM7 ($Arg^{7.34b}$) interaction which was equivalent to that seen in GLP-1-GLP-1R ($Asp15^{*}$-$Arg^{7.34b}$), supporting conserved binding mechanisms and the suitability of GCGR being a template for GLP-1R.

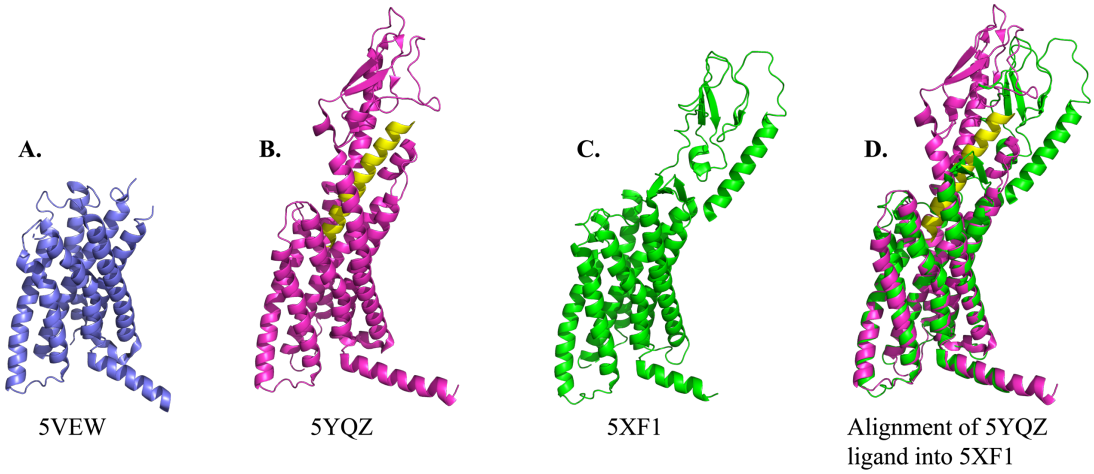


Figure 4.1. A. Inactive TMD GLP-1R PDB VEW (blue), **B.** inactive full-length GCGR (pink) in complex with a GCG analogue (yellow) PDB 5YQZ, **C.** full-length GCGR PDB code 5XF1 (green). The orientation between the ECD and TMD in 5XF1 structure is incompatible with class B two-steps activation model and a change in orientation could be require to allow the binding of the ligand. **D.** Alignment of 5YQZ (pink) and 5XF1 (green) show clashes between ECL1, TM2 and NTD of 5XF1 with the ligand (yellow) (Song et al. 2017; Zhang et al. 2018, Zhang et al. 2017).

4.1.1. Aims and objectives

In order to better understand the peptide-receptor interactions characteristic of generation GLP-1R activity, an *active state* GLP-1R models complexed with GLP-1 and the Gs protein N-terminal α-helix was built using the 6B3J structure. To generate a peptide-bound full-length *inactive state* structure, the homologous glucagon receptor 5YQZ was used to build an inactive-state GLP-1R model. Two additional improvements to the methodology used in Chapter 3 were also included. Firstly, to increase the quality of the simulation a heterogeneous membrane with a similar composition to the pancreatic β-cell was utilised (Díaz et al. 1988; MacDonald et al. 2008; Hoang Do and Thorn 2015; Lin et al. 2005; Xie et al. 2016; Rana, Kowluru, and MacDonald 1986; Meldolesi, Jamieson, and Palade 1971; Kavishwar and Moore 2013). Secondly, Molecular dynamics Flexible Fitting (MDFF) was used to include experimental data from the cryo-EM into the simulations (Zhang et al. 2018; Liang et al. 2018; Trabuco et al. 2008; Jo et al. 2008; Jo et al. 2009; Lee et al. 2016; Lee et al. 2019; Qi et al. 2017; Wu, Cheng, et al. 2014).

4.2. Methods

Model building manipulations, alignments and *in silico* mutations were done using tools from PyMOL software (Schrödinger 2015) unless otherwise stated.

4.2.1. Building inactive GLP-1R

The inactive model was based on the crystal structure of glucagon receptor, PDB 5YQZ (Zhang et al. 2018). (i) All residues not belonging to the receptor were removed and the receptor. (ii) SWISS-MODEL homology modelling server (Waterhouse, Rempfer, et al. 2018) was used to build the inactive model using 5YQZ as a template and GLP-1R human sequence (UniProt P43220). Missing residues from ICL2 region (Glu260 - Thr257) and missing side chains were built by SWISS-MODEL relying in SWISS-MODEL structural database and backbone-dependent rotamer library (iii) The glucagon analogue found in 5YQZ structure was *in silico* mutated to GLP-1 ligand; (iv) because this peptide is two residues shorter than GLP-1, residues His7* and Arg37* were added (Figure 4.2) and residues were renumbered according GLP-1 numbering. (v) GLP-1R and GLP-1 ligands were aligned onto 5YQZ coordinates, GLP-1R model was aligned to GCGR, while GLP-1 ligand was aligned to the glucagon analogue (vi) GLP-1R-GLP-1 complex saved into a single file (vii) the model was subjected energy minimisation to optimise minor steric clashes and poor bond/torsion geometries using NAMD (Bernardi et al. 2016).

4.2.2. Building active GLP-1R

(i) 6B3J structure was stripped from residues not belonging to the receptor (ii) Residues belonging to the stalk and TM1 ($Glu128^{ECD}$ - $Ala158^{1.53b}$) from inactive GLP-1R model were aligned to their homologous residues in 6B3J ($Pro137^{1.32b}$-$Ala158^{1.53b}$) in 6B3J to build the stalk, then (iii) residues $Glu128^{ECD}$ - $Glu138^{1.33b}$ were merged to 6B3J in order to build the missing stalk and the new hybrid receptor was saved (iv) Using the template and the human GLP-1R sequence, missing residues from ICL3 and mutated or missing side chains were built by SWISS-MODEL server

(Waterhouse, Rempfer, et al. 2018) (v) the ligand was built using exendin-P5; residues E1, P32 and S33 from ExP5 were removed and remaining residues were *in silico* mutated and renumber according to GLP-1 and sequence (vi) the new ligand and receptor models were aligned to their homologous in 6B3J (vii) $G_s\alpha5$ segment (residues $M371^{ECL3}$ -$Leu394^{7.49b}$) was added to the model based on 6B3J coordinates (viii) The model was subjected energy minimisation through steepest descent method (50 steps), followed by Newton Raphson methods (50 steps) to optimise minor steric clashes and poor bond/torsion geometries using NAMD (Bernardi et al. 2016).

A. Glucagon analogue: SQGTFTSEYSKYLDSRRAQDFVKWLLNT

GLP-1: HAEFTFTSDVSSYLEGQAAKEFIAWLVKGR

B. Exendin-P5: ELVDNAVGGDLSKQMEEEAVRLFIEWLKNGGPS

GLP-1: HAEFTFTSDVSSYLEGQAAKEFIAWLVKGR

Figure 4.2. Ligand sequence comparison **A.** Glucagon analogue was mutated to GLP-1 to create the ligand used in the inactive structure. In colour red residues mutated to GLP-1 sequence and colour blue residues added. **B.** Exendin-P5 ligand was mutated to GLP-1 sequence to create the ligand for the active model. In colour red residues mutated to GLP-1 sequence and colour grey residues removed.

4.2.3. Molecular dynamics flexible fitting

In early stages of the simulation of the active and inactive state the NTD "fell over" and lay parallel to the lipid membrane (data not shown). The cause was unknown and was suspected to be an artefact. Therefore, in order to improve the simulation and reproduce GLP-1 interaction with GLP-1R, while preserving the dynamism of proteins and taking advantage of cryo-EM large data sets, the experimental data from density maps was added to the simulations. Molecular dynamics Flexible Fitting (MDFF) was used (Trabuco et al. 2009; Qi et al. 2017; Trabuco et al. 2008). MDFF has been successfully used to model ribosomes and their substrates, Mot-TBP complex, and viruses (reviewed in McGreevy et al. 2016; Wollmann et al. 2011; Bharat et al. 2014; Lorenz and Holmes 2010; Gogala et al.

2014). MDFF incorporates forces proportional to the density gradient map to the MD force field (Qi et al. 2017; Trabuco et al. 2008; Trabuco et al. 2009). Despite providing atomic details, three-dimensional proteins built from density maps display atomic assemblies where atoms are fitted into a single position based on the density map. However, proteins are in constant movement and the position of atoms are constantly changing. Since high density area represent energy minima, the map potential drives the atoms into the map density while MD potential preserves physical parameters of the model (Qi et al. 2017; Trabuco et al. 2008; Trabuco et al. 2009). In this way, sidechains and flexible areas were able to move and find a favourable position based on the MD potential while forces from the density map maintain the gross structure of the receptor.

4.2.3.1. System building

Models were orientated with respect to the membrane using OPM database (Lomize et al. 2012). Models were fitted into density maps using 'Colores' program from Situs package (Wriggers 2012; Wriggers and Birmanns 2001). The active model was fitted into 6B3J density map (Liang et al. 2018) and the inactive model into 5YQZ density map (Zhang et al. 2018). The CCP4i software (McNicholas et al. 2011; Potterton et al. 2003) was used to convert the density map of 5YQZ format DSN6 into ccp4 format. Files containing oriented models in the membrane and models fitted into density maps were submitted into CHARMM-GUI server (Jo et al. 2008) where MDFF Utilizer (Qi et al. 2017) was used to generate the input files for simulation. GLP-1R disulphide bonds were first defined and each model was embedded into a 90 x 90 $Å^2$ heterogeneous lipid membrane (Cholesterol 25%, POPC 26%, POPE 20%, POPS 5%, POPI 5%, PSM 19%) before adding neutralizing ions (0.15 M NaCl) and a water/solvent box (TIP3P). Simulations were set at 303.15 K temperature and 1 atm pressure, using Langevin dynamics with a damping coefficient of 1-ps to control the temperature and a Nose-Hoover Langevin piston to control the pressure.

4.2.4. System simulation

Each system underwent equilibration for 70 ns and a total of 350 ns production using all-atoms molecular dynamics flexible fitting (MDFF) method (Trabuco et al. 2008; Trabuco et al. 2009). Simulations were done using CHARMM36 force-field (Huang and MacKerell 2013) and NAMD software (Bernardi et al. 2016) at a 2 fs timestep, using Particle Mesh Ewald to account for long-range electrostatics.

In both trajectories, harmonic restraints obtained from the density map potential were applied to all Cα atoms. Such restraints guided Cα atoms towards high-density regions and allow backbone and sidechains to be fitted into the map. During the first 100 ns, restraints on Cα atom were steadily removed from TMD and the ligand. The following 200 ns restraints were only applied to Cα atoms of the NTD but TMD and ligand were simulated without restraints.

4.2.5. Trajectory analysis

The resulting trajectories were analysed using WORDOM (Caflisch et al. 2007) and VMD (Humphrey, Dalke, and Schulten 1996). From the total output trajectory for the active and inactive models the segment between 150 to 350 ns was in detail analysed.

4.3. Results

4.3.1. Trajectory analysis

Both models were built using high quality structures as templates. Despite mutation of residues and change in the ligand, the simulation of both models quickly reached stability. In the segment analysed, 150-350 ns, the average RMSD of the active model was 1.02 Å and of the inactive model 1.48 Å (Figure 4.3). After 60 ns of simulation restraints from Cα atoms of the TMD in the active and inactive models

were removed. Restraints on the ligand were removed after 100 ns in the active and inactive model. Visual inspection of the structure did not show major changes in the gross structure. Since the inactive model was built using the homologous glucagon receptor, it had an expected higher change in the RMSD compared with the active model built directly from a GLP-1R structure. Despite changes in the RMSD, restrained and unrestrained trajectories fell into acceptable values for simulations (Theoretical and Computational Biophysics Group. 2016 ; Humphrey, Dalke, and Schulten 1996) and show the stability of both models.

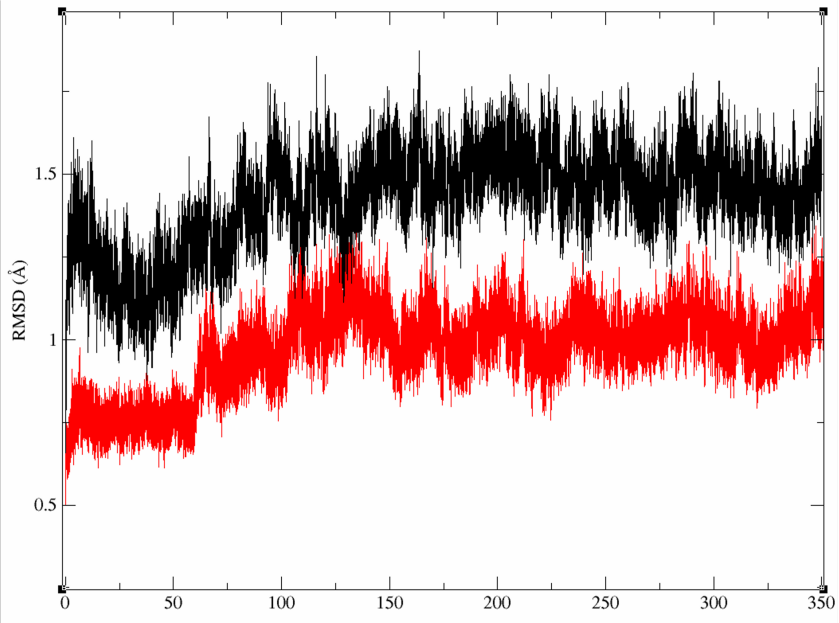


Figure 4.3. RMSD of the transmembrane domain of the active (red) and inactive (black) model during the full trajectory (350 ns). Cα atoms in inactive and active complexes were restraint to the map and steadily removed during the first 100 ns of the simulation. No restraints were applied in ligand and TMD after 100 ns. Plateauing of the RMSD can be observed after 150 ns of the active and inactive models.

GPCRs are in constant movement in the search for a favourable conformation, with some conformations being more favourable than others (Bockenhauer et al. 2011) and this can be observed during the simulation.

Despite changes in RMSD in both models when restraints were removed, RMSD value remained stable with minor changes, meaning that the models achieved stability in their conformational state.

The main difference between the inactive and active conformations in the characteristic outward movement of TM6 which creates a cavity for G-protein binding (Rasmussen, Choi, et al. 2011a). The active model was simulated with the α5-helix of the G protein, which remain in the cavity interacting with the receptor through the simulation; in the inactive model TM6 remains in the closed conformation. Throughout the simulations, and even when restraints were removed from the TMD, the distance between Cα atoms of residues Arg176$^{2.46b}$ (TM2) and Arg348$^{6.37b}$ (TM6) was maintained; on average 23.75 Å in the active model and 11.84 Å in the inactive (Figure 4.4).

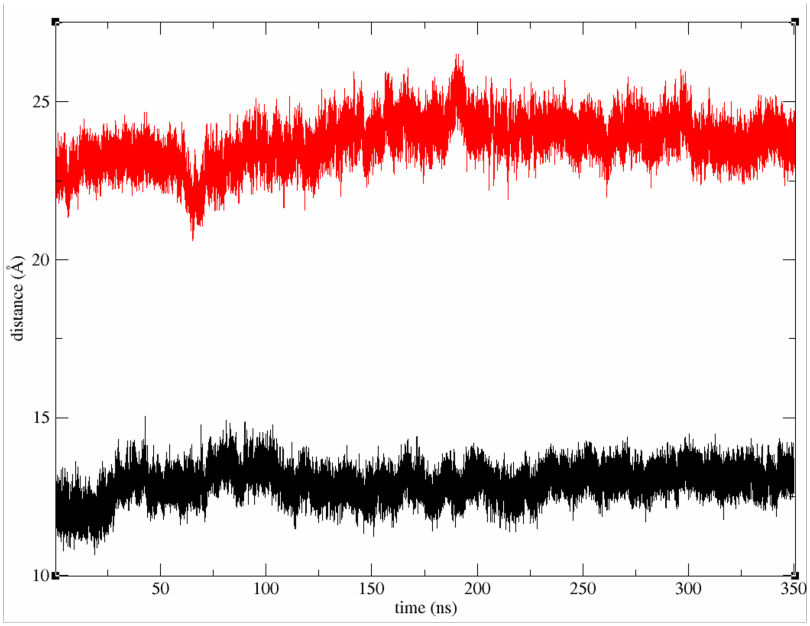


Figure 4.4. Distance between Cα atoms of Arg175$^{2.45b}$ (TM2) and Arg348$^{6.37b}$ (TM6) in inactive (black) and active (red) state throughout the trajectory. The distance is greater in the active state (red) than in the inactive state (black) as consequence of TM6 outward movement. The distance in each model was maintained during the trajectory show the stability of the models in each state.

Even though activation is not seen in the simulation, differences in the inactive and active models are consistent with reported structures as well with previous simulation (Chapter 3). In the active state TM1 moves towards TM7, the extracellular half of TM7 bends towards TM6 and TM6 intracellular end moves outwards. It can also be seen TM6 extracellular half bending towards the membrane and the outward movement of the intracellular end of TM3 and TM5.

4.3.2. Inside the inactive receptor

In the simulation of the inactive model, three networks are seen: i) an extracellular polar network, ii) a hydrophobic transmembrane network and iii) an intracellular network or ionic lock. It is important to identify the characteristics and interactions in the inactive state to determine the order of events leading to activation.

At the extracellular end of the receptor two polar clusters are found delimiting the binding cavity. The first polar cluster is formed by Asn240$^{3.43b}$ - Arg190$^{2.60b}$- Tyr152$^{1.47b}$; and the second polar cluster is formed by Tyr241$^{3.44b}$- Glu-364$^{6.53}$- Thr391$^{7.44b}$. These two polar clusters may be connected via Thr391$^{7.44b}$ linking TM1, TM2, TM3, TM6 and TM7 (Dods and Donnelly 2015); (Table 4.1). It is interesting to note that the interaction of the main chains in the middle region of TM3 during the inactive state is more flexible or less helical (Table 4.1), conferring flexibility to the binding pocket, while still maintaining the receptor in the inactive state (Wootten, Reynolds, Koole, et al. 2016; Wootten, Simms, et al. 2013). In contrast, in the active state, the middle region of TM3 increase the hydrogen bonds between the main chain becoming more stable (hydrogen bonds between Tyr241$^{3.44b}$ -Val237$^{3.40b}$ 93% active, 50% inactive; Tyr241$^{3.44b}$ -Leu245$^{3.48b}$ 78% active, 28% inactive). In the active simulations and as previously described, Thr391$^{7.44b}$ links both polar clusters forming hydrogen bonds with Arg190$^{2.60b}$ (48%) and Glu364$^{6.53b}$ (88%) (Dods and Donnelly 2015); (Table 4.1). Binding of the ligand N-terminal rearranges extracellular residues (Dods and Donnelly 2015).

Table 4.1. Hydrogen bond occupancy of residues at the extracellular end in the active and inactive state

Hydrogen bonds occupancy (%)	Active	Inactive
Arg190$^{2.60b}$– Asn240$^{3.43b}$	95	94
Arg190$^{2.60}$ – Tyr152$^{1.47b}$	48	23
Arg190$^{2.60}$ main –Ser186$^{2.56b}$	87	49
Arg190$^{2.60}$ main – Val194$^{2.64b}$ main	82	98
Tyr241$^{3.44}$– Glu364$^{6.53b}$	84	86
Tyr241$^{3.44}$ main – Val237$^{3.40b}$ main	93	50
Tyr241$^{3.44}$main – Leu245$^{3.48b}$ main	78	28
Thr391$^{7.44b}$ - Arg190$^{2.60b}$	48	*
Thr391$^{7.44b}$ - Glu364$^{6.53b}$	88	*
Thr391$^{7.44b}$ main – Glu387$^{7.42b}$ main	99	81
Glu387$^{7.42b}$ main – Lys383$^{7.37b}$ main	99	74
Hydrogen bonds between sidechains, unless marked as main, in that case the hydrogen bond is formed with the main chain. * Residues within bonding distance (3.5 Å)		

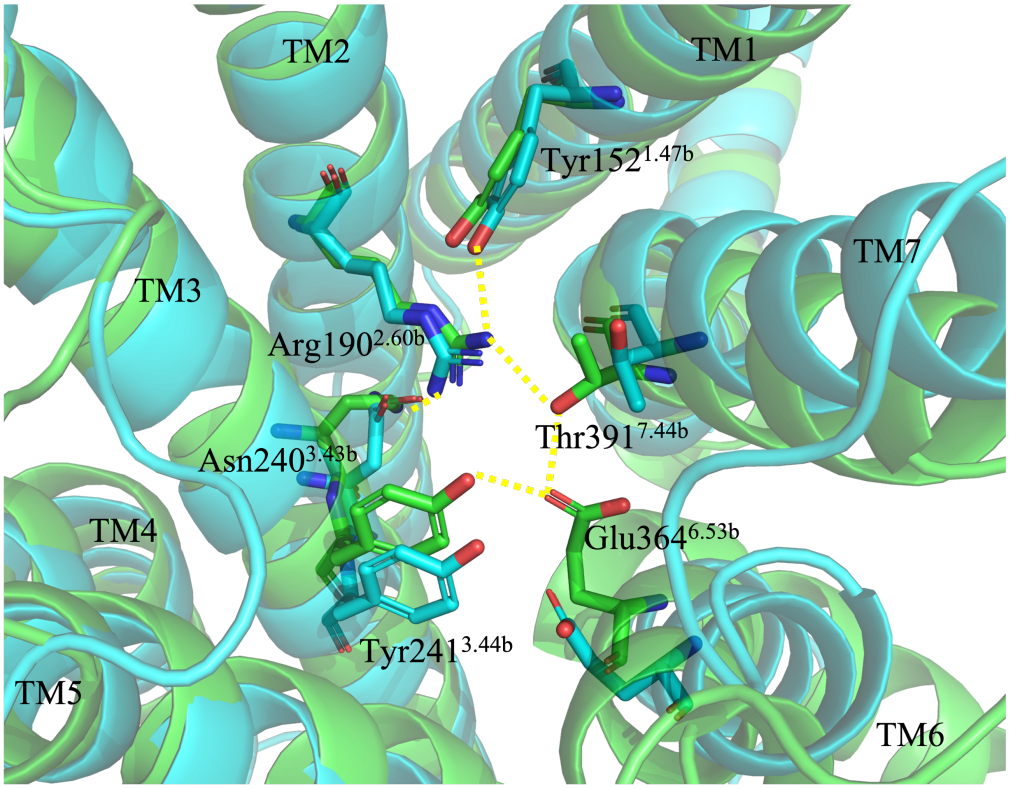


Figure 4.5. Extracellular view comparing the position of residues at the extracellular side in active (green) and inactive (cyan) state. Stick representation of residues display oxygen (red), nitrogen (blue), hydrogen bonds yellow dashed line.

In the core of the receptor residues from Leu183$^{2.53b}$, Leu244$^{3.46b}$, Leu356$^{6.45b}$, Leu360$^{6.549b}$ and Val398$^{7.53b}$ form a hydrophobic network. During activation, nearby polar residues Asn320$^{5.50b}$ and Asn240$^{3.43b}$ will affect the rotamers of these residues and trigger further changes. However, in the inactive conformation, the side chains of these hydrophobic residues are pointing towards the core of the receptor, closing the core of the receptor and "locking" TM6 in the inactive conformation (Figure 4.6).

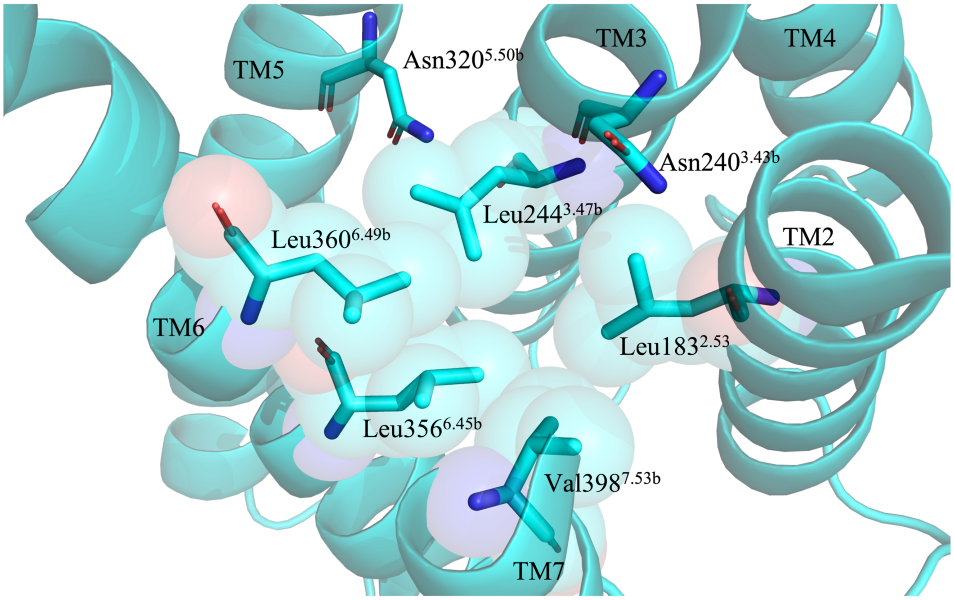


Figure 4.6. Hydrophobic network of the inactive (cyan) state. Stick representation of residues display oxygen (red), nitrogen (blue). Helix 1 and the extracellular portion of helix 7 were hidden.

Table 4.2. Hydrogen bond occupancy of residues belonging to the hydrophobic transmembrane network in active and inactive state

Hydrogen bonds occupancy (%)	Active	Inactive
Leu183$^{2.53b}$ main – Ser186$^{2.56b}$	53	13
Leu244$^{3.47b}$ main – Asn320$^{5.50b}$ side	54	7
Leu356$^{6.45b}$ main – Ser352$^{6.41b}$ main	48	85
Leu360$^{6.49b}$ main – Leu356$^{6.45b}$main	53	99
Leu360$^{6.49b}$ main – Glu364$^{6.53b}$ main	0	75
Val398$^{7.53b}$ main – Gln394$^{7.49b}$ main	97	97
Hydrogen bonds between sidechains, unless marked as main, in that case the hydrogen bond is for the main chain.		

In concordance with previous data, at the intracellular side of the receptor residues His180$^{2.50b}$, Glu247$^{3.50b}$, Thr353$^{6.42b}$ and Tyr402$^{7.57b}$ form a polar network, in the so-called ionic lock (Figure 4.7 and Table 4.3), keeping the receptor in a closed conformation characteristic of the inactive state (reviewed in Wootten, Simms, et al. 2013). However, the function of these residues is not limited to keep the receptor in the inactive state but also are involved in activation and interaction with the G protein as mutagenesis show that removal of the charges when mutated to alanine reduces GLP-1 efficacy (Wootten, Simms, et al. 2013).

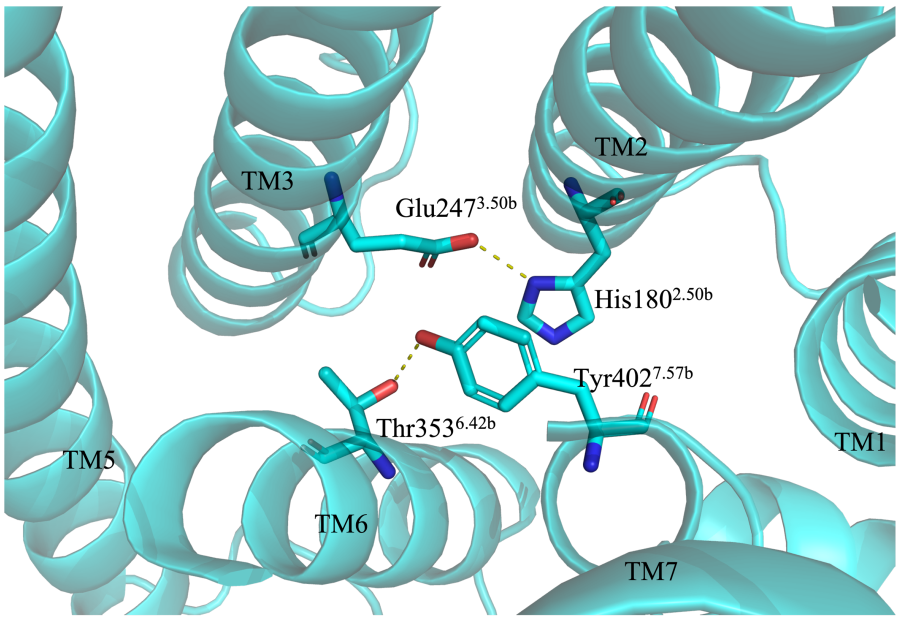


Figure 4.7. Extracellular view of the ionic lock or intracellular polar network in the inactive (cyan) state. Stick representation of residues display oxygen (red), nitrogen (blue) and hydrogen bonds in yellow dashed lines.

Table 4.3. Hydrogen bond occupancy of residues belonging to the intracellular polar network in the active and inactive state.

Hydrogen bonds occupancy (%)	Active	Inactive
His180$^{2.50b}$ - Glu247$^{3.50b}$	56	47
Glu247$^{3.50b}$ – Tyr402$^{7.57b}$	84	0
Thr353$^{6.42b}$ –Tyr402$^{7.57b}$	0	72
Thr353$^{6.42b}$ – Glu247$^{3.50b}$	0	2
Hydrogen bonds formed between sidechains		

4.3.3. Ligand binding in the active conformation

Binding of the ligand disrupts interactions at the extracellular side of the receptor, leading to changes at the extracellular end of TM1 and TM7. Although binding of the ligand is a process not seen during these simulations, comparison of resulting structures, active and inactive, shows different interactions between TM1 and TM2. It appears that the hydrophobic and bulky Phe12*, shifts the extracellular end of TM1 towards TM7, breaking Tyr148$^{1.43b}$ – Lys197$^{2.67b}$ interaction (hydrogen bond occupancy 53% inactive, 1% active) to allow them to interact with the ligand.

TM1 clockwise movement towards TM7 appear to be facilitate by residues at the middle region of TM1 and TM7, Gly151$^{1.46b}$ - Ser155$^{1.50b}$ -Ser392$^{7.47b}$, acting as a pivot (Fig 4.8 and 4.9.B), while at the extracellular end TM1 and TM7, aromatic residues Phe143$^{1.38b}$ and Phe385$^{7.40b}$ stabilise the transmembrane end of the receptor in the active state as spatial restraints avoid the movement of TM7 back to the inactive coordinates (Figure 4.8).

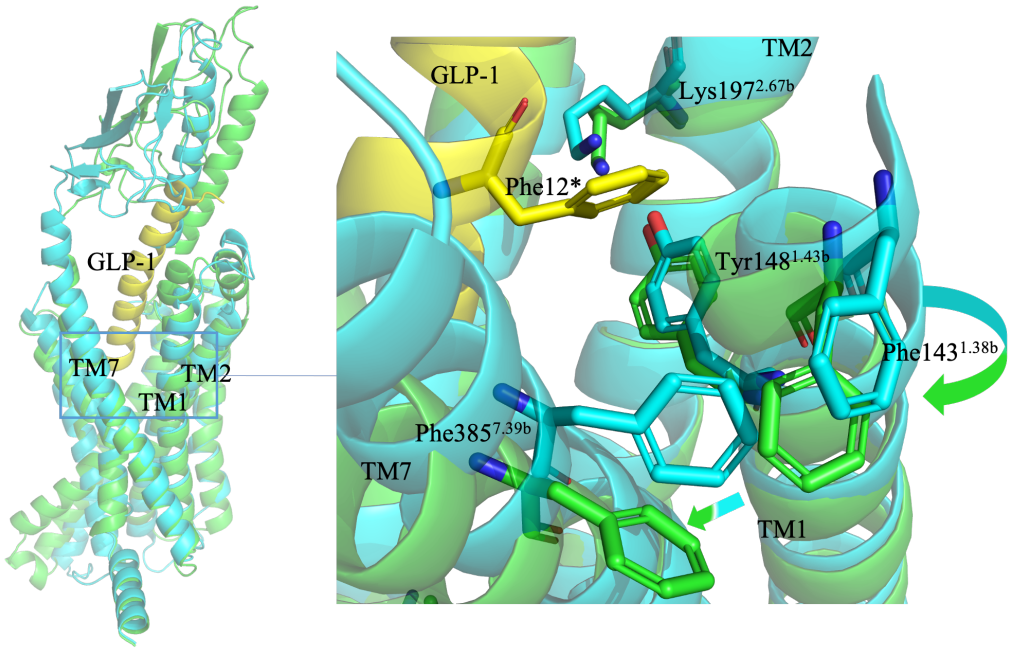


Figure 4.8. Extracellular view comparing TM2-TM1-TM7 in active (green), inactive (cyan) and GLP-1 ligand (yellow) in the active coordinates comparing the position of residues at the extracellular side state. Stick representation of residues display oxygen (red), nitrogen (blue), arrows show the direction of the movement. As result of the ligand movement into the binding pocket, the bulky side-chain of Phe12* disrupts Lys197$^{2.67b}$ – Tyr148$^{1.43b}$ which frees TM1 to allow it subsequent movement towards TM7.

The position of the ligand in the inactive model places Ala8* interacting with Glu387$^{7.42b}$ (hydrogen bond occupancy of 98% during inactive state, 49% active state). Although Ala8*- Glu387$^{7.42b}$ interaction is seen in the active state, this interaction decreases as TM7 extracellular end bends towards TM6. The shift of TM7 rearranges residues Glu387$^{7.42b}$, Phe390$^{7.43b}$, Thr391$^{7.44b}$, Gln394$^{7.49b}$ in TM7 (Figure 4.9 C-D and Table 4.4). As TM7 bends towards TM6 and in the presence of the ligand, the rotamer change in the Thr391$^{7.44b}$ side chain allows the formation of a

hydrogen bond with Glu364$^{6.53b}$ (hydrogen bond occupancy 88% active, 0% inactive) forcing Gln394$^{7.49b}$ movement towards the intracellular side where it forms a hydrogen bond with Pro358$^{6.47b}$ main chain stabilising the kink formation in TM6 (hydrogen bond occupancy 84% active, 0% inactive; Table 4.4), while being in binding distance from His363$^{6.53b}$. The possible π-stacking effect of Phe390$^{7.43b}$ over His363$^{6.52b}$ may shift TM6 and with it His363$^{6.52b}$ into the core within binding distance of Gln394$^{7.49b}$ adding stability to TM6 kink formation (Figure 4.9 C-D). Comparison of the inactive and active models show a small outward movement of the intracellular end of TM5 along with TM6. Most residues at the extracellular end of TM5 are involved in cAMP stimulation (Mathi et al. 1997; Takhar et al. 1996).

In the middle region of TM5 Asn320$^{5.50b}$ forms hydrogen bonds with Leu244$^{3.47b}$ and the main chain of Leu360$^{6.49b}$ that is exposed with TM6 outward movement and kink formation (figure 4.9, Table 4.2 and 4.4). Also, in TM5, Phe324$^{5.54b}$ aromatic ring moves pointing towards the core of the receptor, possible aiding the kink formation and determining the extension of TM6 outward movement affecting the size of the intracellular cavity for G protein.

Table 4.4. Hydrogen bond occupancy of residues surrounding TM6 kink in active and inactive state

Hydrogen bonds occupancy (%)	Active	Inactive
Ser155$^{1.50b}$ - Gly151$^{1.45b}$ main	95	65
Ser155$^{1.50b}$ - Ser392$^{7.47b}$ main	12	8
Asn320$^{5.50b}$ – Leu244$^{3.47b}$ main	54	7
Asn320$^{5.50b}$ – Leu360$^{6.49b}$main	35	0
Asn320$^{5.50b}$ main – Phe324$^{5.54b}$ main	90	76
Glu387$^{7.42b}$ – Ala8* (main)	49	98
Thr391$^{7.44b}$ - Glu364$^{6.53b}$	88	0
Thr391$^{7.44b}$- Arg190$^{2.60b}$	48	0
Thr391$^{7.44b}$ main - Gln394$^{7.49b}$ main	29	99
Gln394$^{7.49}$- Pro358$^{6.47b}$ main	84	0
Hydrogen bonds between sidechains, unless marked as main, in that case the hydrogen bond is formed with the main chain.		

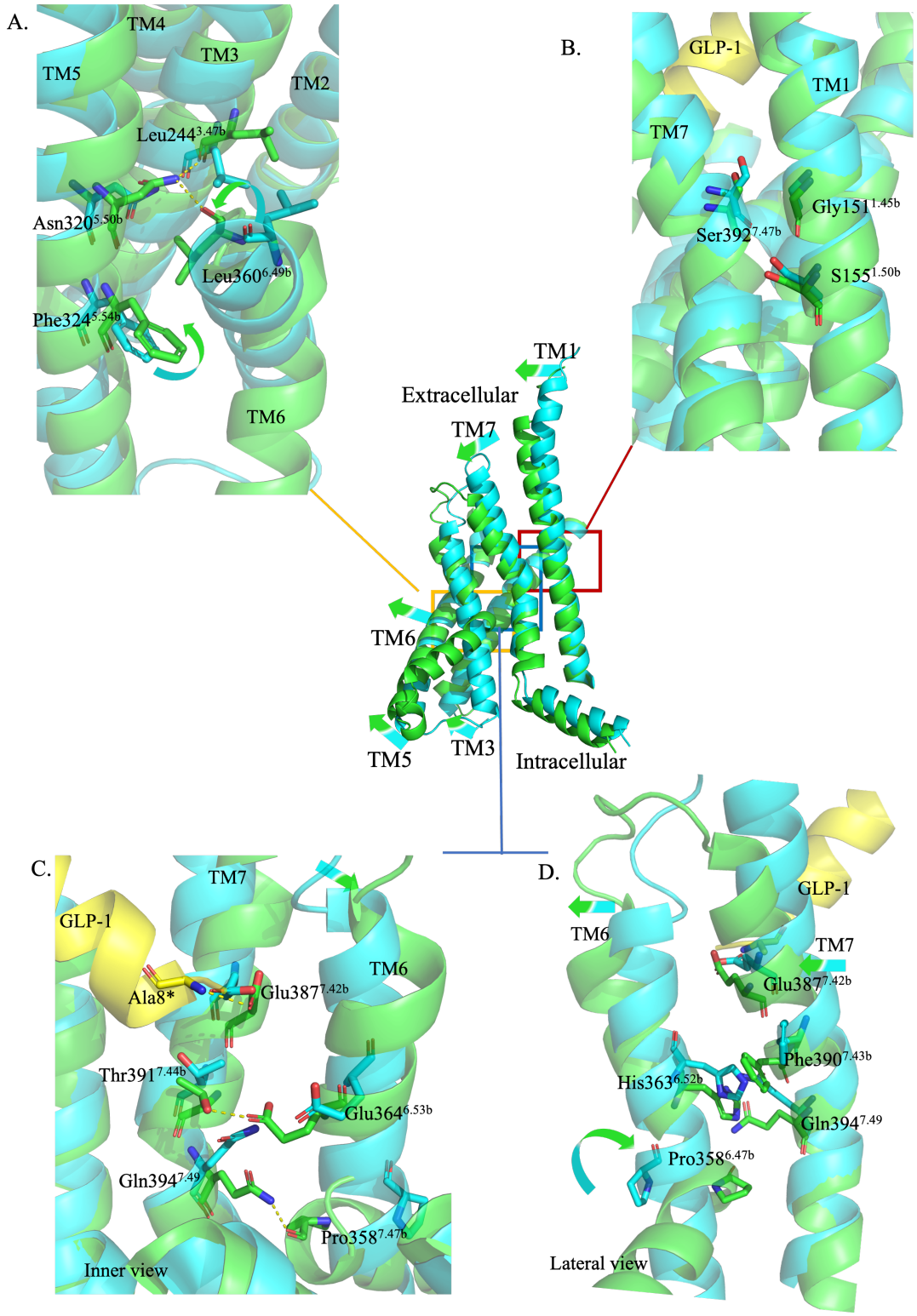


Figure 4.9. Comparison of active (green), inactive (cyan) and GLP-1 ligand (yellow) in the active coordinates. Stick representation of residues display oxygen (red), nitrogen (blue), hydrogen bonds (yellow dash) and arrows show the direction of the movement.

Binding of the ligand triggers conformational changes through the receptor. Although activation mechanisms are not seen in our simulation, comparison of the models in the active and inactive state show structural differences.

At the extracellular side, binding of the ligand rearranges the extracellular polar network, allowing the rotation of TM7 and its movement towards TM6. Comparison of the ligand in the inactive model with the active position show His7, and Ala8* as the first residues to interact with the polar network. In the inactive state His7* forms hydrogen bonds with Glu387$^{7.42b}$ (38% occupancy) and Ala8* (98% occupancy) shifting the extracellular end of TM7 towards TM6 to accommodate His7* and Ala8* via Glu387$^{7.42b}$ (Figure 4.10), mutagenesis data show Glu387A mutation reduces GLP-1 affinity (Dods and Donnelly 2015) but not G387D, suggesting the sidechains' role to 'catch' the ligand for TM7 rearrangement. However, comparison with the active state, show the decrease in hydrogen bonds in the active state between His7* - Glu387$^{7.42b}$ (38% occupancy), and Ala8* - Glu387$^{7.42b}$ (49% occupancy) suggest the movement of the ligand into the binding pocket.

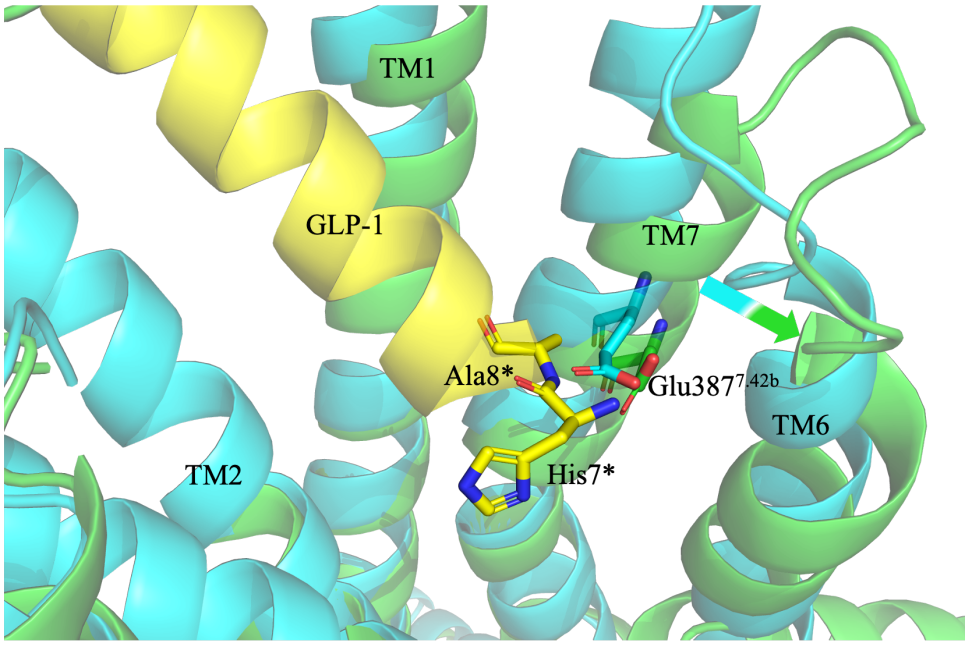


Figure 4.10. Extracellular view showing the overlay of active (green), inactive (cyan) and GLP-1 ligand in active coordinate (yellow). The effect of the ligand over the extracellular end of TM7 shift Glu387$^{7.42b}$ and in consequence TM7 over TM6. Stick representation of residues display oxygen (red), nitrogen (blue), arrows show the direction of the movement.

As the ligand moves forwards into the binding pocket, Glu9* disrupts the extracellular polar network by forming hydrogen bonds with Tyr152$^{1.47b}$ and Arg190$^{2.60b}$ (Table 4.5). The presence of the ligand shifts the Arg190$^{2.60b}$ side chain to the intracellular side strengthening Arg190$^{2.60b}$-Asn240$^{3.43b}$ hydrogen bond. In contrast with the previous hybrid model (Chapter 3), there was no interaction between Arg190$^{2.60b}$-Gln394$^{7.49b}$ in the inactive state. During the active model simulation, Pro358$^{6.47b}$ forms a hydrogen bond with Gln394$^{7.49b}$ stabilising TM6 kink. Pro358$^{6.47b}$ is part of the highly conserved class B P$^{6.47}$xxG$^{6.50}$ (P$^{6.47}$LLG$^{6.50}$ in GLP-1R) motif involved in coordinating large conformational changes [reviewed in (Liang et al. 2018). The low helical propensity of Pro-58$^{6.47b}$ (Pace and Scholtz 1998) allows the exposed backbone to interact with Gln394$^{7.49b}$ (85% occupancy) and stabilise the TM6 kink. The flexibility of the region is key for TM6 outward movement. Mutagenesis data show that double mutation N240A/Q394A affects GLP-1R activation (ΔLog τ_c = 0.70) (Wootten, Reynolds, Koole, et al. 2016). A reason for this decrease in activation may be due to the difficulty of spatial restraints removal by Asn240$^{3.43b}$ easing spatial restraints removal and the lack of pivot found in Gln394$^{7.49b}$ by TM6 indirectly affecting residues that stabilise TM6 kink (Figure 4.9 and 4.11).

Table 4.5. Hydrogen bond occupancy in active and inactive state after ligand interaction (TM1-TM2)

Hydrogen bonds occupancy (%)	Active	Inactive
Glu9* – Arg190$^{2.60b}$	37	5
Glu9* - Tyr152$^{1.47b}$	70	0
Tyr152$^{1.47b}$ – Arg190$^{2.60b}$	48	23
Arg190$^{2.60b}$ – Asn240$^{3.43b}$	95	94
Arg190$^{2.60b}$ main – Ser186$^{2.56b}$ main	87	50
Asn240$^{3.43b}$ – Ser186$^{2.56b}$	70	8

Hydrogen bonds between sidechains, unless marked as main, in that case the hydrogen bond is formed with the main chain.

In this same region, a second interaction between TM3-TM2 is seen as a hydrogen bond between Asn240$^{3.43b}$- Ser186$^{2.56b}$ (Figure 4.4.). The importance of these interactions appears to be that Asn240$^{3.43b}$ acts as a fulcrum reducing distance between TM2-TM3 and removing any possible spatial restraint from the core of the

receptor by shifting Phe187$^{2.57b}$ out of the core (Figure 4.11), as these residues are opposite to the TM6 kink. Binding of the ligand Glu9* with Arg190$^{2.60b}$ in the active state shift Arg190$^{2.60b}$ towards the intracellular side. At the same time hydrogen bonds interaction between, Arg190$^{2.60b}$ – Ser186$^{2.56b}$ and Asn240$^{3.43b}$ – Ser186$^{2.56b}$ increase in the active state (Table 4.5) resulting in the packing of the middle region of TM2 and TM3 as consequence of the ligand binding. The role of Arg190$^{2.60b}$ in ligand binding is well known as mutagenesis show. If Arg190$^{2.60b}$ is mutated to GLP-1 efficacy decreases (ΔLog τ_c = 0.53 Wootten et al. 2011), suggesting the need of a charged sidechain to bind with the negative charge of Glu9* (or Glu3* in Ex4). Similarly, Y152A mutation reduce GLP-1 and (exendin4) potency and affinity (Coopman et al. 2011) suggesting the role of this residue in and the inward movement of TM1 to close the binding pocket. Mutagenesis data show N240Q had no effect on GLP-1R (Wootten et al. 2016b) as Gln polar side chain replace Asn allowing Arg190$^{2.60b}$-Asn240$^{3.43b}$ interaction, but not N240A, as it affects GLP-1R activation (ΔLog τ_c = 0.67 Wootten et al. 2016b) hampering TM3 movement towards TM2 (Figure 4.11).

Although the order of events is not seen in our simulations, comparison of active and inactive states show rotamer changes of hydrophobic residues contributing to TM6 kink formation. Movement of the Leu244$^{3.47b}$ sidechain towards TM2 creates a small void in the centre of the receptor and removes spatial restraints allowing the rotation of Leu360$^{6.49b}$, and TM6, in direction to TM5 (Figure 4.12). The low helical propensity of TM6 at PxxG motif facilitate the exposure of Leu360$^{6.49b}$ main chain which only in the active state forms a hydrogen bond with Asn320$^{5.50b}$ sidechain (35% hydrogen bond occupancy) contributing to stabilise TM6 kink at TM5-TM6 interface (Figure 4.12, 4.13).

TM6 kink is located at the middle region of TM6 involving the conserved class B P$^{6.47}$xxG$^{6.50}$ motif (reviewed in Liang et al. 2018). In the active model, the hydrogen bond between Pro358$^{6.47b}$ and Gln394$^{7.49b}$ side chain (hydrogen bond occupancy 84% only present in the active state) contributes to the stabilisation of the kink at TM6-TM7 interface (Figure 4.12).

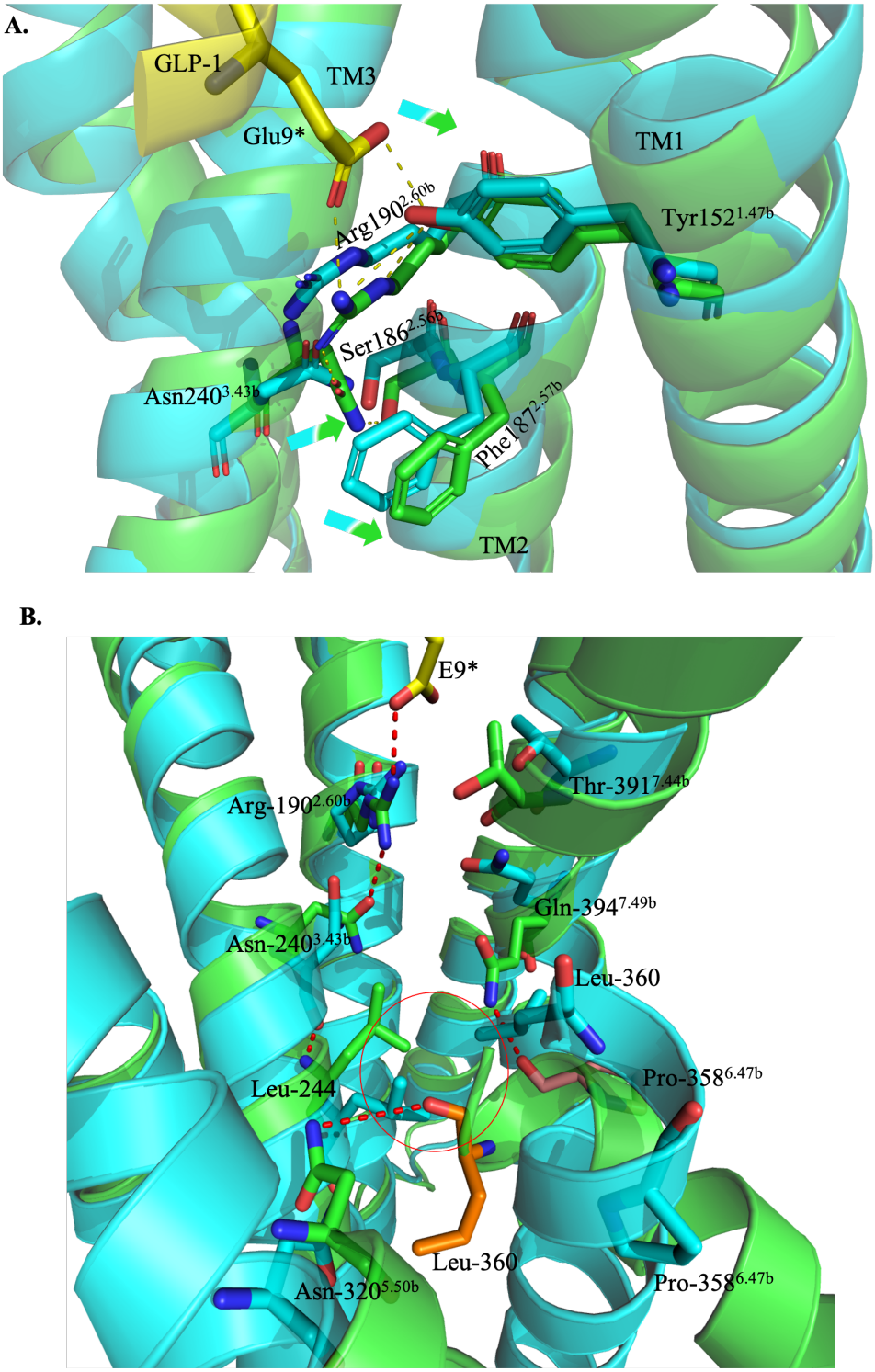


Figure 4.11. Overlay of active (green), inactive (cyan) and GLP-1 ligand in the active coordinates (yellow). Stick representation of residues display oxygen (red), nitrogen (blue), arrows show the direction of the movement. **A.** TM3-TM2-TM1 view, extracellular end of TM2 and TM5, TM6 and TM7 were hidden. This image shows Glu9* binding to Arg190$^{2.60b}$ and triggering further conformational changes in Asn240$^{3.43b}$, Ser186$^{2.56b}$, and Phe187$^{2.57b}$. **B.** The importance of TM3 rearrangement creates an inner cavity (red circle) allowing the rotation of Leu360$^{6.40}$ (orange-movement) and Pro358$^{6.47b}$ (orange: movement) in the active state, and stabilisation of TM6 kink via a hydrogen bond between Gln394$^{7.49b}$ – Pro358$^{6.47b}$.

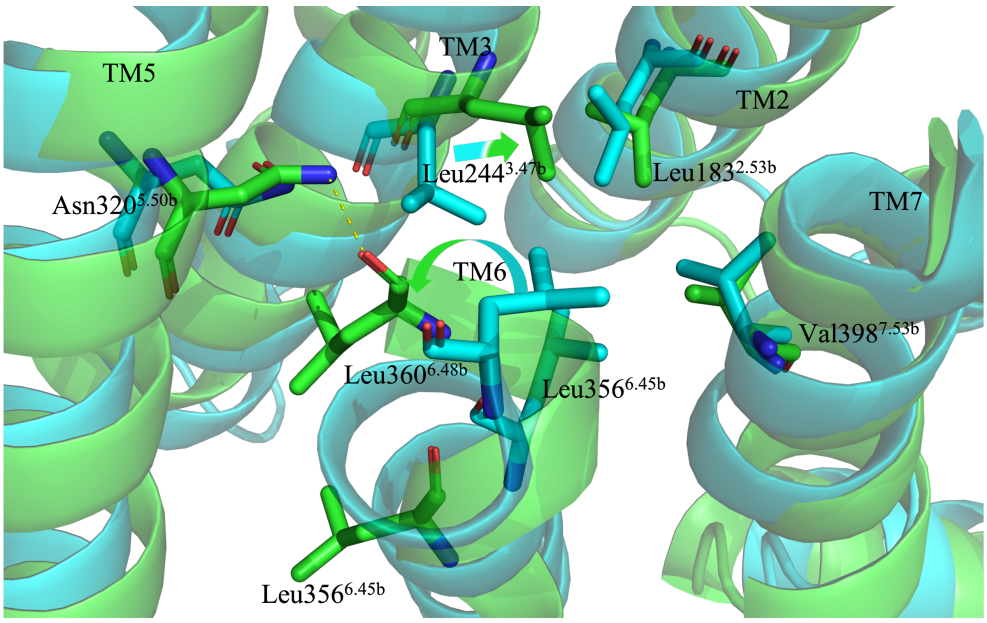


Figure 4.12. Extracellular view of hydrophobic residues in the core of the receptor overlaying active (green) and inactive (cyan) states. Stick representation of residues display oxygen (red) and nitrogen (blue). Arrows show the direction of the movement. Extracellular end of TM6 and TM7 were hidden. Shift of Leu244$^{3.47b}$ creates a small cavity which allows the rotation movement of Leu360 and kink formation.

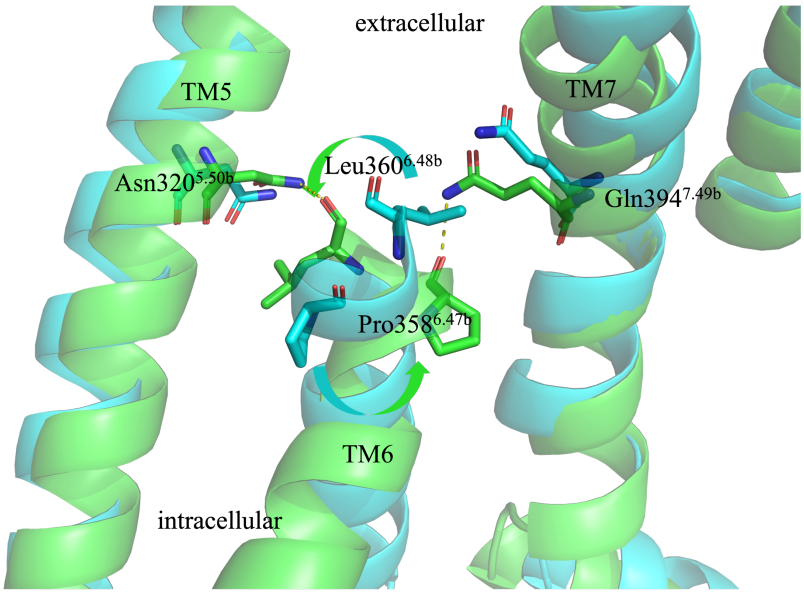


Figure 4.13. Lateral view of TM5, TM6 and TM7 overlaying active (green) and inactive (cyan) states. Stick representation of residues display oxygen (red) and nitrogen (blue). Hydrogen bonds in yellow dashed lines. Arrows show the direction of the movement. The rotation movement of Leu360$^{6.48b}$ allows kink formation. The extracellular half of TM6 and TM1, TM2, TM3 and TM4 were hidden.

4.4. Discussion

There were structural improvements from the previous hybrid models since the active model was based on a biased agonist-bound human GLP-1R -Gs complex,(PDB 6B3J) and the inactive model was based on a full-length glucagon receptor in complex with an agonist (PDB 5YQZ) giving information about the interactions between the ligand and the receptor. The density gradient improved the models by driving the structures to high density areas while maintaining the flexibility of structures reported by low density gradient, since the receptor is composed of different domains with different degrees of freedom, thus preventing overfitting and providing better results.

Comparison between the inactive and active models show that in the active state hydrophobic residues from the ligand, specially Phe12*, disrupt TM1-TM2, facilitating the movement of TM1 towards TM7. Recently a GLP-1R-GLP-1(9-36) and a PAM (LSN3160440) was able to fully activate the GLP-1R by binding of the PAM in the TM1-TM2 interface, allowing access of the ligand (Bueno et al. 2020). This shows the importance of the hydrophobic interactions between the ligand and TM1-TM2 to facilitate the entrance of the ligand into the binding pocket but also relying in residues His7* and Ala8* at the N-terminal of the ligand to stabilise the binding (Bueno et al. 2020). In GLP-1(9-36) the absence of His7* and Ala8* result in different rotamers of Phe12* which might be the reason for a lower affinity of GLP-1(9-36) (Bueno et al. 2020). The shift of TM1 over TM7 followed by TM7 movement over TM6 improves ligand binding by shrinking the binding pocket. In addition, a difference observed in this MDFF simulation was the lack of interaction between His7* and Glu364$^{6.53b}$ (as seen in Chapter 3) or with any other residues in the extracellular end of TM5 or TM6. This is surprising since residues Arg310$^{5.40b}$ and Glu364$^{6.53b}$ are involved in GLP-1 efficacy (Dods and Donnelly 2015; Wootten et al. 2016a; Wootten et al. 2016b) and a direct interaction with the ligand could be expected. Instead His7* interacts with Glu387$^{7.42b}$ as it shifts TM7 extracellular end towards TM6. This suggests a greater role of Glu387$^{7.42b}$ and TM7 in activation as in might improve the stability of the ligand. Indeed, Glu387$^{7.42b}$ is required for GLP-1 efficacy (Dods and Donnelly 2015) and mutation of residues at the extracellular end of TM7 had no effect in ExP5 but affected GLP-1R signalling (Liang et al. 2018). It is important to consider that this model was based in a GLP-1R in complex with a

ExP5, a G protein-biased agonist that is two residues longer than GLP-1. The ligand His7* interaction with Glu387$^{7.42b}$, might be at the early stages of activation, as the ligand enters the binding pocket, and will shift TM7 shrinking the pocket and guiding His7* sidechains towards TM5-TM6 extracellular.

In TM2 is the characteristic Arg190$^{2.60b}$ – Glu9*. This interaction is not only present in other class B GPCRs, but its mutation affects receptor-endogenous ligand efficacy (Donnelly 2012; Unson, Gurzenda, and Merrifield 1989; Gardella and Jüppner 2001; Hinke et al. 2001; Rivier, Rivier, and Vale 1984; Solano et al. 2001; Wootten et al. 2018). Comparison between state show conformational changes after binding in the helical segment below Arg190$^{2.60b}$ to the central hydrophobic network where rotamer changes allows the creation of a small cavity allowing TM6 rotation and kink formation.

4.5 Conclusions

Binding of the ligand shrinks the binding space and increases the packing of the helices at the extracellular side. Disruption of TM1-TM2 interface triggers further changes: i) shift of TM1 towards TM7 and TM7 towards TM6 with the objective to guide the ligand's N-terminus towards TM5-TM6. ii) Interaction between Glu9*-Arg190$^{2.60b}$, binds and accommodates the ligand and in addition to the movement of TM2 towards TM3 pack the core of TM2-TM3 creating a cavity that will allow TM6 kink formation. iii) Previous reported interaction between Arg380$^{7.34b}$ -Asp15* (reviewed in de Graaf et al. 2016) was not present.

Chapter 5:. Targeted Molecular Dynamics of GPCRs

5.1. Introduction

While the simulations described in Chapters 3 and 4 enable an understanding of ligand-receptor interactions in one conformational state, larger and more complex conformational transitions, such activation events. are not accessible via this approach since they occur over longer time scales (picoseconds to seconds). Despite the constant grown in computational power, it is still almost impossible to simulate such events using classic MD at temperatures inside physiological ranges. Targeted molecular dynamics is an alternative method that can be used to induce conformational changes to a known target by applying time-dependent constraints at physiological temperature (Schlitter, Engels, and Kruger 1994). Target molecular dynamics guides a group of atoms towards the final 'target'.

In order to understand GLP-1R activation by its endogenous ligand, high-quality GLP-1R models embedded in a realistic heterogeneous membrane were simulated using MDFF method to allow the ligand, sidechains and missing residues to accommodate and find a favourable conformation in the active and inactive state (Chapter 4). In this chapter, these resulting receptor-ligand complexes were used as the input for targeted MD simulations, in order to define the GLP-1R activation mechanism, the order of events, and the identification of important residues involved in the process. This method requires an initial structure, for which we used the final state from MDFF simulation in the inactive state of GLP-1R, while for the final or 'target' state the final state from MDFF simulation of the GLP-1R active state was used (Chapter 4). Going from the inactive to the active state allowed us to simulate activation of the receptor. From earlier use of targeted MD in the study of the unfolding of the α-helical portion of insulin and searching of pathways of conformational transitions (Schlitter, Engels, and Kruger 1994), targeted MD has been successfully used for the identification and validation of biologically relevant drug molecules (sirtuins) (Rumpf et al. 2015; Jones et al. 1997; Kiviranta et al. 2008), protein unfolding (Ferrara, Apostolakis, and Caflisch 2000), and conformational changes that participate in the closure of ligand-gated ion channels (Rovšnik et al. 2021) and identification of allosteric mechanisms of calmodulin (Liang, Pang, et al.

2017) and identification of active site in aldehyde dehydrogenases and cofactor flip (Rahuel-Clermont et al. 2019).

To validate these targeted MD simulations for GLP-1R, the β_2 adrenergic receptor (β_2AR) was used as a test case since its activation mechanisms has been extensively studied (Bang and Choi 2015; Benovic et al. 1987; Chan, Filipek, and Yuan 2016; Dror et al. 2011; Laporte et al. 1999; Liu, Horst, et al. 2012; Lohse et al. 1990; Luttrell et al. 1999; Manglik et al. 2015; Manna et al. 2016; Noda et al. 1994; Rasmussen et al. 2007; Rasmussen, Choi, et al. 2011b; Rasmussen, DeVree, et al. 2011; Samama et al. 1993b; Shi et al. 2002; Strader, Sigal, and Dixon 1989; Yao and Kobilka 2005). β_2AR is a Class A GPCR with an activation mechanism which involves conformational changes required to accommodate catecholamines and enable G protein activation and signalling. Ionic and hydrogen bond interactions between Asp113$^{3.32}$, Asn312$^{7.39}$ (superscripts in Ballesteros-Weinstein numbering, (Ballesteros and Weinstein 1995) with the ligand are important for the binding affinity, while hydrophobic residues Val114$^{3.33}$, Phe193$^{5.32}$, and Phe290$^{6.52}$ enclose the agonist into the orthosteric binding pocket (Rosenbaum et al. 2007). At the catechol end of the ligand, hydrogen bonds between Ser203$^{5.42}$ and Ser207$^{5.46}$ with the agonist phenoxy moieties results in a tightening of the binding pocket with a shift of TM5 towards TM6. As the result, Pro211$^{5.50}$ moves into the space occupied by Ile121$^{3.40}$ in the inactive structure, forcing the Ile121$^{3.40}$ rotamer to change and shift towards TM6 into the space occupied by Phe282$^{6.44}$ in the inactive state. Phe282$^{6.44}$ and TM6 move away from TM3, tilting the intracellular end of TM6 away from TM3, creating an intracellular cavity for G protein binding.

5.2. Methods

5.2.1. β_2 Adrenergic Receptor

5.2.1.1 Model Building

Molecular model building manipulations and alignments were carried out using the tools embedded within PyMOL (Schrödinger 2015) unless otherwise stated.

5.2.1.1.1. Inactive model

The structure of the turkey β_1 adrenergic receptor bound to agonist isoprenaline, PDB 2y03 (Warne et al. 2011) was used to build the inactive model of a human β_2AR. (i) Residues not belonging to the receptor were removed. (ii) The receptor was mutated to the human β_2AR (UniProt P07550) using the homology modelling web server SWISS-MODEL (Waterhouse, Bertoni, et al. 2018). (iii) Residues were renumbered according to β_2AR. (iv) Isoprenaline was *in silico* mutated to adrenaline and added to the model based on 2Y03 coordinates. (v) β_2AR-adrenaline complex was subjected to energy minimisation through steepest descent method (50 steps), followed by Newton Raphson methods (50 steps) to optimise minor steric clashes or poor geometries using NAMD (Bernardi et al. 2016).

5.2.1.1.2. Active model

The structure of β_2 adrenoceptor bound to adrenaline and an engineered nanobody (Nb80), PDB 4LDO (Ring et al. 2013), was used to build the active state. (i) Residues not belonging to the receptor were removed, while ICL3 region from 2Y03 was used to build missing residues. (ii) The receptor was used as a template to build a model of human β_2AR using the homology modelling web server SWISS-MODEL (Waterhouse, Bertoni, et al. 2018). (iv) Adrenaline and nanobody were added to the model based on 4LDO coordinates, since the nanobody stabilises a physiologically relevant state (Rasmussen, Choi, et al. 2011a). (v) β_2AR-adrenaline-Nb complex was subjected to energy minimisation through steepest descent method (50 steps), followed by Newton Raphson methods (50 steps) to optimise minor steric clashes or poor geometries using NAMD (Bernardi et al. 2016).

5.2.1.2. Molecular Dynamics Flexible Fitting

5.2.1.2.1. Building the system

Models were fitted into density maps using the 'Colores' program from Situs package (Wriggers and Birmanns 2001; Wriggers and Chacón 2001). Active and inactive models were fitted into 4LDO (Ring et al. 2013) and 2Y03 density maps (Warne et al. 2011) respectively. MDFF Utilizer using CCP4i Software (McNicholas et al. 2011; Potterton et al. 2003) was used to convert the density map from the DSN6 format into ccp4. Models were orientated with respect to the hydrocarbon core of the lipid bilayer utilising OPM database and server (Lomize et al. 2012). Files containing oriented models in the membrane and fitted into density maps were submitted into CHARMM-GUI server (Jo et al. 2008; Jo et al. 2009; Lee et al. 2016) where MDFF Utilizer (Qi et al. 2017) was used to generate the input files for the first stage of the simulation. Disulphide bonds were first assigned, and each model was independently embedded into a 90 x 90 $Å^2$ heterogeneous lipid membrane (Cholesterol 25%, POPC 26%, POPE 20%, POPS 5%, POPI 5%, PSM 19%) before adding neutralizing ions (0.15 M NaCl) and a water/solvent box (TIP3P). Simulations were set at 303.15 K temperature and 1 atm pressure, using Langevin dynamics with a damping coefficient of 1-ps to control the temperature and a Nose-Hoover Langevin piston to control the pressure using NAMD (Bernardi et al. 2016).

5.2.1.2.2. System simulation

Each system underwent equilibration for 70 ns and a total of 350 ns production using all-atoms molecular dynamics flexible fitting (MDFF) method (Trabuco et al. 2008; Trabuco et al. 2009). Simulations were done using CHARMM36 force-field (Huang and MacKerell 2013) and NAMD software (Bernardi et al. 2016) at a 2 fs timestep, using Particle Mesh Ewald to account for long-range electrostatics.

In both trajectories, harmonic restraints obtained from the density map potential were applied to all Cα atoms. Such restraints guided Cα atoms towards high-density regions and allow backbone and sidechains to be fitted into the map. During the first 100 ns, restraints on Cα atom were steadily removed from TMD and

ligand. The following 200 ns restraints were only applied to Cα atoms of the NTD but TMD and ligand were simulated without restraints.

5.2.1.3. Targeted Molecular Dynamics

5.2.1.3.1. Building the system

Structures obtained from the final frame of MDFF simulation of active and inactive β_2AR were used for targeted MDs as 'target' and 'initial' structures respectively. Since this method requires the same number of atoms, the nanobody was added to the inactive file but was manually moved towards the intracellular space away from the receptor. The inactive or 'initial' structure was orientated with respect to the lipid bilayer utilising OPM database and server (Lomize et al. 2012) and submitted into CHARMM-GUI server (Jo et al. 2008; Jo et al. 2009; Lee et al. 2016) where membrane builder option was used to generate the input for targeted MD with NAMD (Jo, Kim, and Im 2007; Jo et al. 2009; Wu, Cheng, et al. 2014; Schlitter, Engels, and Kruger 1994). Disulphide bonds were first assigned, and the model was embedded into a 90 x 90 $Å^2$ heterogeneous lipid membrane (Cholesterol 25%, POPC 26%, POPE 20%, POPS 5%, POPI 5%, PSM 19%) before adding neutralizing ions (0.15 M NaCl) and a water/solvent box (TIP3P). The active structure obtained from the final frame of MDFF was set as the 'target' structure.

5.2.1.3.2. System simulation

Targeted MD Simulations were performed using CHARMM36 force-field (Huang and MacKerell 2013). The system underwent equilibration for 70 ns and was set for a total of 150 ns production using all-atoms targeted MDs. Steering forces guiding the initial inactive structure towards the final 'target' active structure were only applied to Cα atoms using an elastic constant for targeted MD of 200 kcal/mol/$Å^2$ (Schlitter, Engels, and Kruger 1994). Simulations were set at 303.15 K temperature and 1 atm pressure, using Langevin dynamics with a damping coefficient of 1-ps to control the temperature and a Nose-Hoover Langevin piston to control the

pressure. Simulations were performed using NAMD software (Bernardi et al. 2016) at a 2 fs timestep, using Particle Mesh Ewald to account for long-range electrostatics.

5.2.2. GLP-1R from inactive to active state

5.2.2.1. System building

Structures obtained from the final frame of MDFF simulation of active and inactive MDFF of GLP-1R (Chapter 4) were used for targeted MDs as 'target' and 'initial' structures respectively. Since this method requires the same number of atoms in each file, therefore, to match the number of atoms in the active state the G protein α5-helix was added to the inactive structure file but manually moved towards the intracellular space away from the receptor. This file was orientated with respect to the lipid bilayer utilising OPM database and server (Lomize et al. 2012) and submitted into the CHARMM-GUI server (Jo et al. 2008; Jo et al. 2009; Lee et al. 2016) where membrane builder option was used to generate the input for targeted MD simulation with NAMD (Jo, Kim, and Im 2007; Jo et al. 2009; Wu, Cheng, et al. 2014; Schlitter, Engels, and Kruger 1994).

5.2.2.2. Targeted molecular dynamics simulation

Simulations were performed using CHARMM36 force-field (Huang and MacKerell 2013) and NAMD software (Bernardi et al. 2016) at a 2 fs timestep, using Particle Mesh Ewald to account for long-range electrostatics. The system underwent equilibration for 70 ns and was set for a total of 150 ns production using all-atoms targeted MDs. Steering forces guiding the initial inactive structure towards the final 'target' active structure were only applied to Cα atoms using an elastic constant for targeted MD of 200 kcal/mol/Å^2 (Schlitter, Engels, and Kruger 1994) was only applied to Cα atoms. To study activation, the inactive state was set as the starting structure and the active state set as the 'target' structure.

5.3. Trajectory analysis

The resulting full trajectory obtained from targeted MD simulation of β2AR and GLP-1R complexes were analysed using WORDOM (Caflisch et al. 2007), VMD (Humphrey, Dalke, and Schulten 1996), and Bio3D module in R to calculate rotamer states (Grant et al. 2006; Skjærven et al. 2014). To facilitate the analysis of the trajectories in a chronological order, the resulting trajectories were divided according to RMSD values, resulting segments of varied length, resulting in segment with a plateau RMSD and segments with changes in RMSD. For each segment hydrogen bonds and rotameric angles were analysed. Hydrogen bonds were obtained as previously described (Section 2.3.2) and rotamer angles were obtained using torsion.pdb Bio3D package from R (Grant et al. 2006). In order to get rotameric angles, water and lipid bilayer atoms were removed from the trajectory, then the trajectory was converted into PDB files using WORDOM (Caflisch et al. 2007), and alanine and Glycine residues were removed from the file. Then, using torsion.pdb Bio3D package from R (Grant et al. 2006), each frame was analysed and rotameric angles were saved into a txt file. Since the frames were analysed in chronological order, it was possible to identify segments with changes in rotameric angles. Special attention was paid to residues with a solvent accessible surface area with a 20 percent difference between the inactive and active state. SASA for each residue in each state was first calculated using "Accessible surface area and accessibility calculation for protein" version 1.2 web server from the Centre for Informational Biology, Ochanomizu University (Center for Informational Biology 2012). If there was a 20% difference in the SASA value of a residue in the inactive – active state, then the rotamer angles for this residue was analysed. Then these residues were separated into 'layers' according to their position in the receptor, resulting in three categories: extracellular, middle, intracellular and two subcategories: extracellular-middle and middle-intracellular.

5.3. Results

5.3.1. β2 Adrenergic Receptor

5.3.1.1. Molecular dynamics flexible fitting trajectory analysis

Models of the inactive β2AR and active β2AR in complex with adrenaline were built and simulated for a total 70 ns using MDFF method (Figure 5.1); during that time the receptor remained stable. Expected interactions between the active state receptor and the ligand were found: i) Asn312$^{7.39}$ and Asp113$^{3.32}$ are interacting with the amino group of the ligand; ii) Asn312$^{7.39}$ forms a hydrogen bond with the hydroxyl group of the ligand, iii) residues aSer203$^{5.42}$ and Ser207$^{5.46}$ in TM5 and iv) Asn293$^{6.55}$ in TM6 form hydrogen bonds with the phenoxyl-groups of the ligand; v) at the extracellular side, hydrophobic Phe193^{ECL2} and Phe289$^{6.51}$ surround the binding pocket (Deupi and Kobilka 2007; Rasmussen, DeVree, et al. 2011; Rosenbaum et al. 2007) (Figure 5.2).

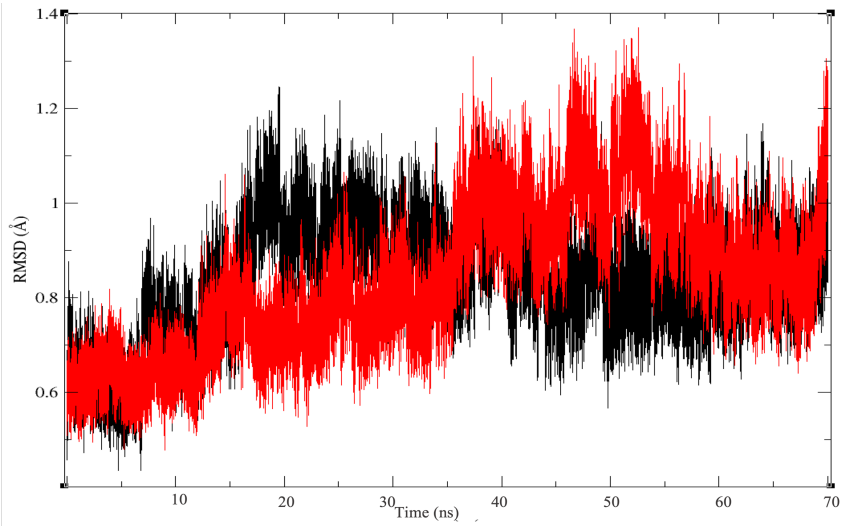


Figure 5.1. RMSD of the TMD of the active (red) and inactive (black) βAR model during MDFF trajectory (70 ns). During the first 30 ns of the simulations the Cα atoms in the inactive and active models were restrained to the electron density map.

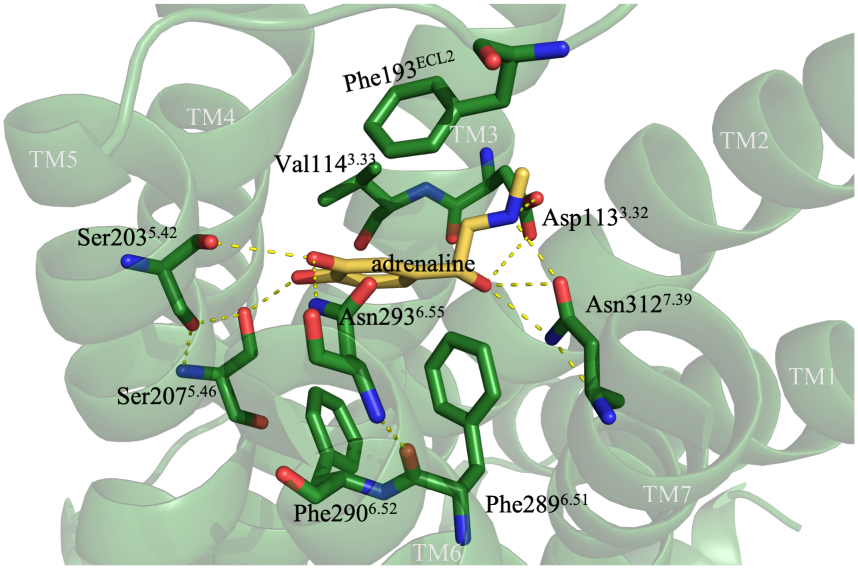


Figure 5.2. Snapshot of the last frame of MDFF of β2AR (green) in complex with adrenaline ligand (yellow) in the active state. Stick representation have oxygen red, nitrogen blue, hydrogen bonds as yellow dashed lines. In this image interactions in the extracellular end of the receptor with adrenaline ligand are seen.

5.3.1.2. Targeted molecular dynamics trajectory analysis

Targeted MD guided the receptor from a state (inactive) towards another (target active), therefore, changes in RMSD are expected. The total trajectory (150ns) was divided into seven segments, (A-G) of different length based on changes in RMSD; where there are segments with increases on RMSD reflecting major changes in the backbone more stable RMSD values reflecting minor changes in the backbone. Throughout the trajectory there were changes in the sidechains (Figure 5.3).

During targeted MD, the conformational changes typical of GPCR activation were seen. In β2AR, there is a polar network at the extracellular side, located on TM3 TM7, TM6 and TM5, that is directly involved in binding the agonist and transducing the subsequent conformational changes. Residues Asp113$^{3.32}$ and Asn312$^{7.39}$ have a key role in agonist binding (reviewed in Chan, Filipek, and Yuan 2016). Across the simulation Asp113$^{3.32}$ and Asn312$^{7.39}$ interact with the ligand, adrenaline, through ionic interactions and hydrogen bonds (average 93% and 63% respectively; Table 5.1 and 5.2). In addition, a hydrogen bond between Tyr316$^{7.43}$ and Asp113$^{3.32}$ remains constant through the simulation and contributes to the shrinking of the binding pocket

and forms the roof of an inner cavity around TM7-TM1-TM2 interface and affecting Met82$^{2.53}$ shifts (Figure 5.4 and Table 5.1).

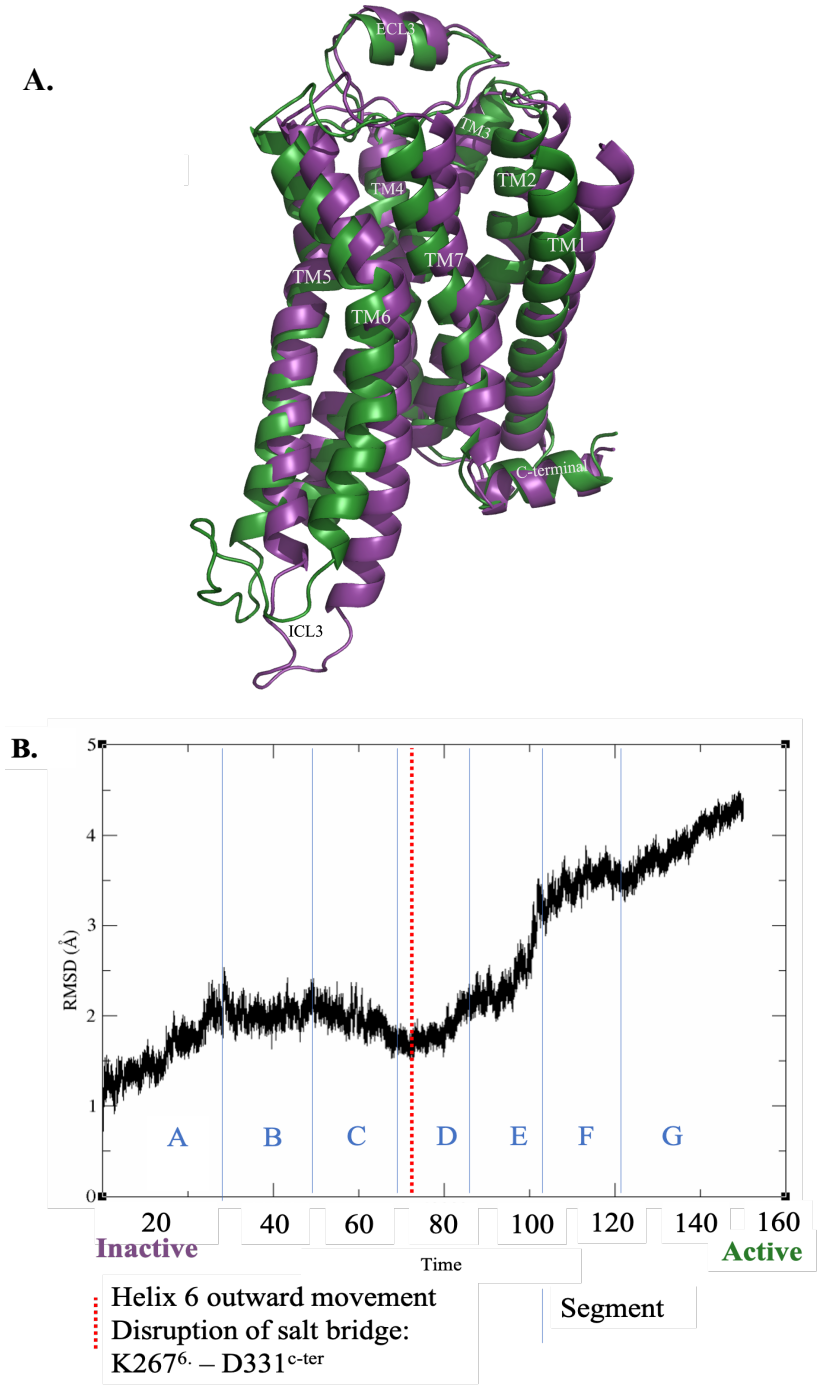


Figure 5.3. **A.** Overlaying initial inactive (purple) and resulting active (green) β2AR after 150 ns of targeted MD. **B.** The total trajectory was divided into seven segments (A-G, blue lines) according to RMSD values. Each segment was analysed. The red dashed line marks the moment in the trajectory of TM6 outward movement, disrupting K267$^{6.}$ – D331$^{c\text{-}ter}$ salt bridge.

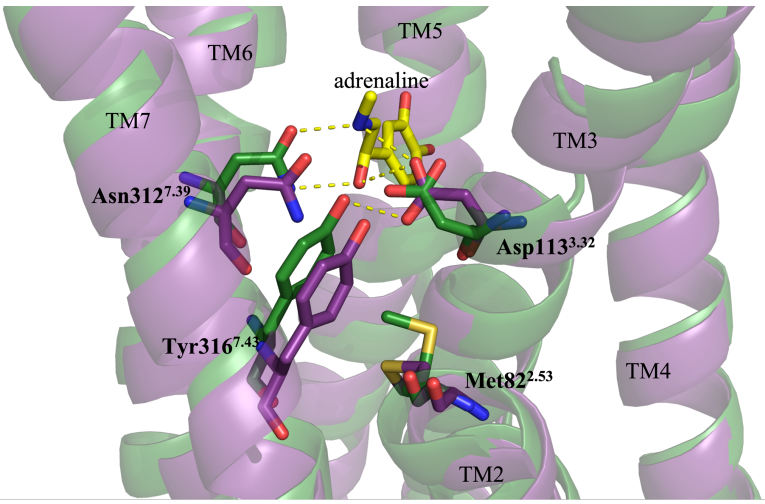


Figure 5.4. Lateral view of extracellular end overlaying initial inactive (purple) and resulting active (green) β2AR after 150 ns of targeted MD in complex with adrenaline ligand (yellow) in the final active state coordinates. TM1 and TM2 extracellular end were hidden. Stick representation have oxygen red, and nitrogen blue. Hydrogen bonds as yellow dashed lines. Arrows showing the direction of movement. Binding of the ligand, the inward movement of TM3-TM7 and the hydrogen bond between Asp113$^{3.32}$ – Tyr316$^{7.43}$ forms the 'roof' of a cavity.

The effect of the movement of TM7 over TM6 rearranges TM6 via hydrogen bonds between Tyr308$^{7.35}$-Asn293$^{6.55}$ and Asn312$^{7.39}$-adrenaline ligand that 'push' Phe289$^{6.51}$, and TM6 extracellular end, counter-clock propagating changes in Trp286$^{6.48}$ (equivalent to Glu364 in GLP-1R) contributing to TM6 rotation (Figure 5.5 and 5.6). As consequence of helices 6 and 7 movement, mainly on Asn293$^{6.55}$ interaction with the ligand, residues Ser203$^{5.42}$ and Ser207$^{5.46}$ in helix 5 interact with ligand phenoxy moieties (Chan, Filipek, and Yuan 2016; Strader et al. 1989).

Changes from the extracellular side are reflected in the middle region of the receptor as rotamer changes that affect intracellular residues. As such, in consequence of the interaction between the ligand and Asp113$^{3.32}$ (Table 5.2) there is the inward movement of TM3 reducing the binding pocket forcing a rotamer change in Val117$^{3.36}$ which due to spatial restraints force the rotamer change in Ser120$^{3.39}$. The inward movement of TM5 due to Ser207$^{5.46}$ -ligand interaction shifts Pro211$^{3.40}$ into the space previously occupied by Ile121$^{3.40}$ (Figure 5.7).

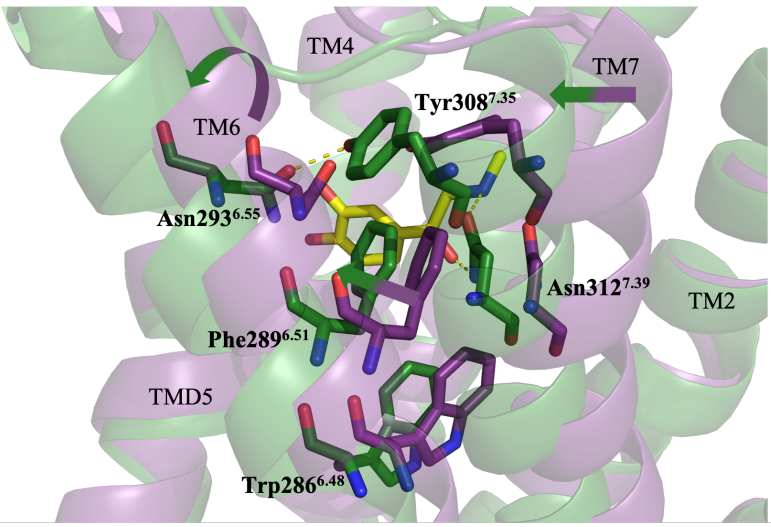


Figure 5.5. Lateral view of extracellular end overlaying initial inactive (purple) and resulting active (green) β2AR after 150 ns of targeted MD in complex with adrenaline ligand (yellow) in the final active state coordinates. TM1 and TM2 extracellular end were hidden. Stick representation have oxygen red, nitrogen blue, arrow showing the movement of TM7 towards TM6 and curved arrow shows counter-clock rotation of TM6 triggering TM6 outward movement.. Hydrogen bond (yellow dashed line) between Asn293$^{6.55}$ – Tyr308$^{7.35}$ and Asn312$^{7.39}$ with adrenaline shift Phe289$^{6.51}$.

During activation, TM7 rotates resulting in rotational changes of the side chains of residues in the helix, and rotamer changes as Asn318$^{7.45}$ pushes Ser319$^{7.46}$ inward. The role of Asn318$^{7.45}$ is to maintain the outward movement of helix 6 in the active state by avoiding its movement back to the inactive position by restraining Phe282$^{6.44}$. While Ser319$^{7.46}$ forms a hydrogen bond with Asp79$^{2.50}$ (rotamer change of neighbouring Ser120$^{3.39}$ allows the inward movement of Asp79$^{2.50}$ removal of physical restraints from sidechains and creating a small void enough for the later rotamer change in Ile121$^{3.40}$ and Phe282$^{6.44}$) reducing the space inside the core (Table 5.2, 5.3). In addition to TM7 movement, rotamer change of Ser120$^{3.39}$ disrupts Asp79$^{2.50}$ and Tyr326$^{7.53}$ hydrogen bond. The now free Tyr326$^{7.53}$ forms a hydrogen bond with the conserved Tyr219$^{5.58}$ (Probst et al. 1992; Table 5.1), critical for stabilising and maintaining the active state in the β_2-AR (Dror et al. 2011; Gabilondo, Krasel, and Lohse 1996) and enabling the foramtion of the G protein bindin site (Fleetwood et al. 2019). It is well known that Tyr219$^{5.58}$ favours TM6 outward movement and contribute to class A activation efficacy (Venkatakrishnan et al. 2016; Ragnarsson et al. 2019). Mutagenesis of either NPxxY$^{7.53}$ and Tyr$^{5.58}$ stabilises β_2-AR inactive state decreasing cAMP signalling and agonist affinity (Ragnarsson et al. 2019; Gabilondo, Krasel, and Lohse 1996), while and T326A (TM7) mutation

impairs G protein-coupling and internalisation (Gabilondo, Krasel, and Lohse 1996). It is suggested that the water mediating hydrogen bond between Tyr219$^{5.58}$ and Tyr326$^{7.53}$ is the equivalent opposite to the ionic lock, as Tyr219$^{5.58}$ - Tyr326$^{7.53}$ interaction is needed to stabilise the active state (Schneider et al. 2010).

During activation, the rotamer change of Ile121$^{3.40}$ side chain produce its inward movement towards TM6 to the space previously occupied by Phe282$^{6.44}$ in the inactive state. As a consequence of Ile121$^{3.40}$ inward movement, Phe282$^{6.44}$ is displaced away from TM3, tilting the intracellular end of TM6 away from TM3, breaking DRY ionic lock and creating an intracellular cavity for G protein binding (Dror et al. 2011) (Figure 5.7 and 5.8).

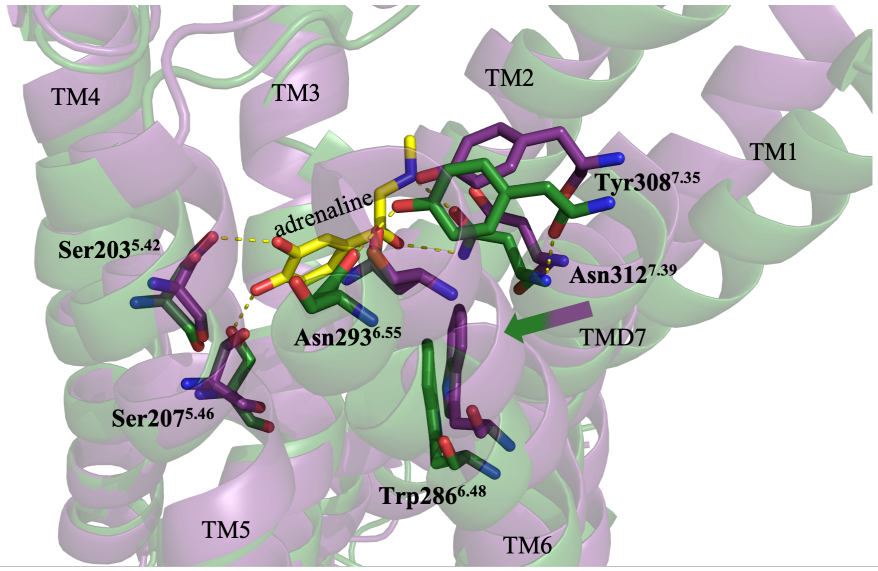


Figure 5.6. Lateral view of extracellular end overlaying initial inactive (purple) and resulting active (green) β2AR after 150 ns of targeted MD in complex with adrenaline ligand (yellow) in the final active state coordinates. TM1 and TM2 extracellular end were hidden. Stick representation have oxygen red, nitrogen blue and hydrogen bond (yellow dashed line) between Ser203$^{5.42}$ and Ser207$^{5.46}$ with adrenaline.

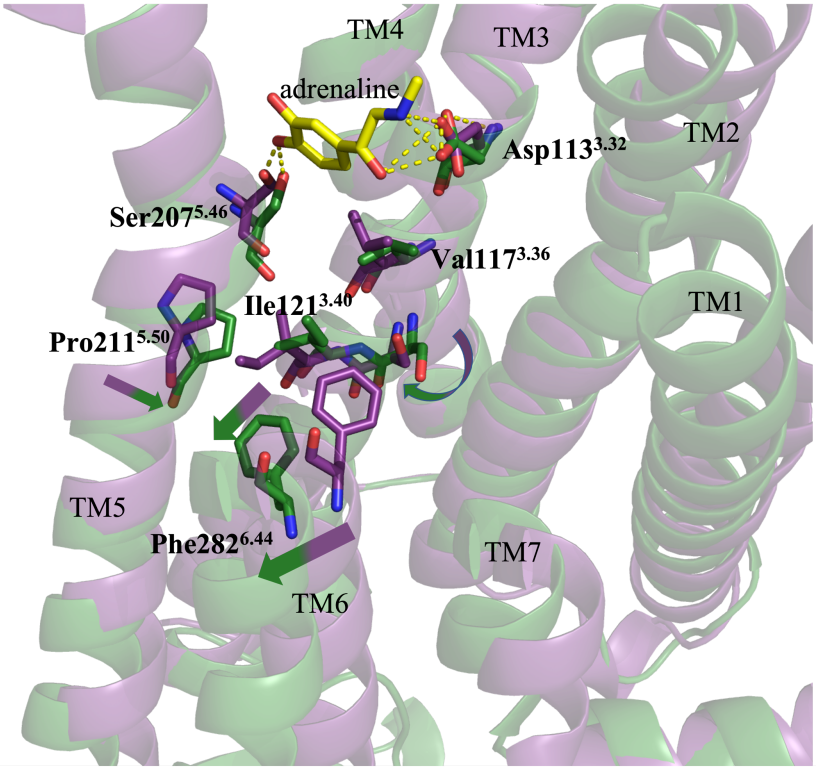


Figure 5.7. Lateral view overlaying initial inactive (purple) and resulting active (green) β2AR after 150 ns of targeted MD in complex with adrenaline ligand (yellow) in the final active state coordinates. The extracellular end of TM6 and TM7 were hidden. Stick representation have oxygen red, nitrogen blue and hydrogen bond (yellow dashed line) between adrenaline and Asp113$^{3.32}$ and Ser207$^{5.46}$. Rotamer change (curved arrow) in Val117$^{3.36}$ induce further Ser120$^{3.39}$ and Ile121$^{3.40}$ rotamer changes. Ile121$^{3.40}$ shift Phe282$^{6.44}$ resulting in TM6 outward movement.

Table 5.1. Percentage of hydrogen bond occupancy across B2AR trajectory

Region	Hydrogen bond	Segment						
		A	B	C	D	E	F	G
Extracellular	D113$^{3.32}$ - adrenaline	100	100	100	76	72	100	100
Extracellular	N312$^{7.39}$ - adrenaline	100	74	19	0	38	75	72
Extracellular	Y316$^{7.43}$ - D113$^{3.32}$	83	87	61	78	74	71	72
Extracellular	Y308$^{7.35}$ - N293$^{6.55}$	55	70	78	68	43	52	53
Extracellular	S203$^{5.42}$ - adrenaline	19	15	24	6	-	4	11
Extracellular	S207$^{5.46}$ - adrenaline	57	58	33	-	-	16	81
Core	D79$^{2.50}$ – S319$^{7.46}$	95	95	98	98	98	98	96
Core	D79$^{2.50}$ –N322$^{7.49}$	82	79	92	93	94	93	90
Core/Intracellular	D79$^{2.50}$ – Y326$^{7.53}$	91	96	21	0	0	0	0
Core/Intracellular	Tyr219$^{5.58}$ – Y326$^{7.53}$	0	0	0	0	0	0	51

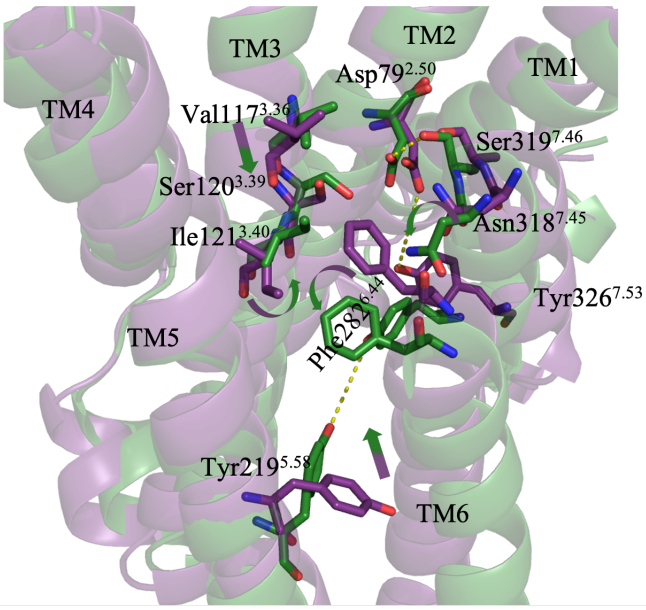


Figure 5.8. Lateral view overlaying initial inactive (purple) and resulting active (green) β2AR after 150 ns of targeted MD in complex with adrenaline ligand (yellow) in the final active state coordinates. TM7 and TM6 extracellular end were hidden. Stick representation have oxygen red, nitrogen blue and hydrogen bonds yellow dashed line. Rotamer change of Val117$^{3.36}$ (consequence of D113$^{3.32}$-ligand and shrinking of binding pocket –Figure 5.7). Change in Ile121$^{3.40}$ and F282$^{6.44}$ is facilitated by S207$^{5.46}$ - ligand interaction and TM5 inward movement and P211$^{5.50}$ movement which removes physical restraints for F282$^{6.44}$. The rotamer change in Tyr219$^{5.58}$ allows Tyr219$^{5.58}$ – Tyr326$^{7.53}$ hydrogen bond.

Although during the simulation hydrogen bonds occupancy varies. It is noticeable that hydrogen bonds at the extracellular end decrease during segments: C, in which RMSD values decrease, suggesting stabilisation of the receptor before TM6 outward movement; and segment D, in which RMSD achieve the lowest values but the distance between TM6-TM2 starts to increase, suggesting the formation of the intracellular cavity and activation.

Since targeted MDs only guide Cα atoms, gross changes in the structure are expected, as such the outward movement of TM6, however, during the simulation complementary changes in the sidechains associated to activation are seen.

Table 5.2. Percentage of chis angles through B2AR trajectory
Rotamer angles: 60º gauche + (g+), 180º trans (t) and -60º gauche - (g-)

		A	B	C	D	E	F	G
Residue	chi	%	%	%	%	%	%	%
D79$^{2.50}$	x_1 g+	0	0	0	0	0	0	0
	x_1 g-	100	100	99	100	100	100	100
	x_1 t	0	0	1	0	0	0	0
	x_2 g+	0	0	1	0	1	1	1
	x_2 g-	100	100	99	100	100	99	100
	x_2 t	0	0	0	0	0	0	0
V117$^{3.36}$	x_1 g+	0	0	13	6	3	5	0
	x_1 g-	0	0	0	0	0	0	0
	x_1 t	100	100	87	95	97	95	100
S120$^{3.39}$	x_1 g+	0	0	0	0	0	0	0
	x_1 g-	65	100	90	15	38	20	0
	x_1 t	35	0	10	85	62	80	100
I121$^{3.40}$	x_1 g+	0	0	0	0	0	0	0
	x_1 g-	0	0	0	0	0	0	22
	x_1 t	100	100	100	100	100	100	79
	x_2 g+	0	0	0	0	0	0	0
	x_2 g-	0	0	0	0	0	0	0
	x_2 t	0	0	0	0	0	0	0
L124$^{3.43}$	x_1 g+	0	0	0	0	0	0	0
	x_1 g-	100	100	100	100	100	48	59
	x_1 t	0	0	0	0	0	52	42
	x_2 g+	0	0	0	0	0	52	22
	x_2 g-	0	0	0	0	0	0	0
	x_2 t	100	100	100	100	100	48	78
P211$^{5.50}$	x_1 g+	75	71	76	70	57	72	34
	x_1 g-	25	29	24	30	43	28	66
	x_1 t	0	0	0	0	0	0	0
	x_2 g+	24	29	24	30	43	27	65
	x_2 g-	76	71	76	70	57	73	35
	x_2 t	0	0	0	0	0	0	0
M215$^{5.54}$	x_1 g+	0	4	8	1	12	18	16
	x_1 g-	0	0	0	0	0	13	40
	x_1 t	100	97	92	99	88	68	45
	x_2 g+	0	2	3	0	3	1	5
	x_2 g-	73	21	63	86	18	30	11
	x_2 t	27	77	34	14	79	70	85
	x_3 g+	26	74	29	13	69	45	41
	x_3 g-	73	22	63	86	20	35	37
	x_3 t	1	5	8	2	12	20	23
F282$^{6.44}$	x_1 g+	0	0	0	0	0	0	0
	x_1 g-	0	0	1	0	0	4	0
	x_1 t	100	100	99	100	100	96	100
	x_2 g+	0	0	0	0	0	0	0
	x_2 g-	100	100	100	100	100	100	89
	x_2 t	0	0	0	0	0	0	11
N318$^{7.45}$	x_1 g+	2	4	8	13	15	9	0
	x_1 g-	0	0	0	0	0	0	0
	x_1 t	98	96	93	87	85	91	100
	x_2 g+	19	25	73	85	84	58	78
	x_2 g-	2	4	5	11	12	7	0
	x_2 t	79	71	22	4	4	34	22
S319$^{7.46}$	x_1 g+	0	0	0	0	0	0	0
	x_1 g-	100	100	100	100	100	100	100
	x_1 t	0	0	0	0	0	0	0
N322$^{7.49}$	x_1 g+	0	0	0	0	0	0	0
	x_1 g-	100	100	100	100	99	100	100
	x_1 t	0	0	0	0	1	0	0
	x_2 g+	14	22	6	3	5	6	5
	x_2 g-	86	78	94	97	95	95	96
	x_2 t	0	0	0	0	0	0	0
Y326$^{7.53}$	x_1 g+	0	0	0	0	0	0	0
	x_1 g-	100	100	100	100	100	100	100
	x_1 t	0	0	0	0	0	0	0
	x_2 g+	0	0	0	0	0	0	63
	x_2 g-	100	100	100	100	100	100	22
	x_2 t	0	0	0	0	0	0	15

5.3.2. GLP-1R Targeted Molecular dynamics simulation

The inactive structure was guided towards the active state by applying restraints in Cα atoms over 150 ns. To facilitate the analysis of the trajectory in a chronological order, the resulting trajectory was divided in base of RMSD values of varied length into seven segments (A-G); segments with stable RMSD values were separated from segments with increase or changes in RMSD. Then each segment was analysed. Segments with rapid increases on RMSD reflect changes in the main chain while stable segments reflect mainly changes in sidechains (Figure 5.9).

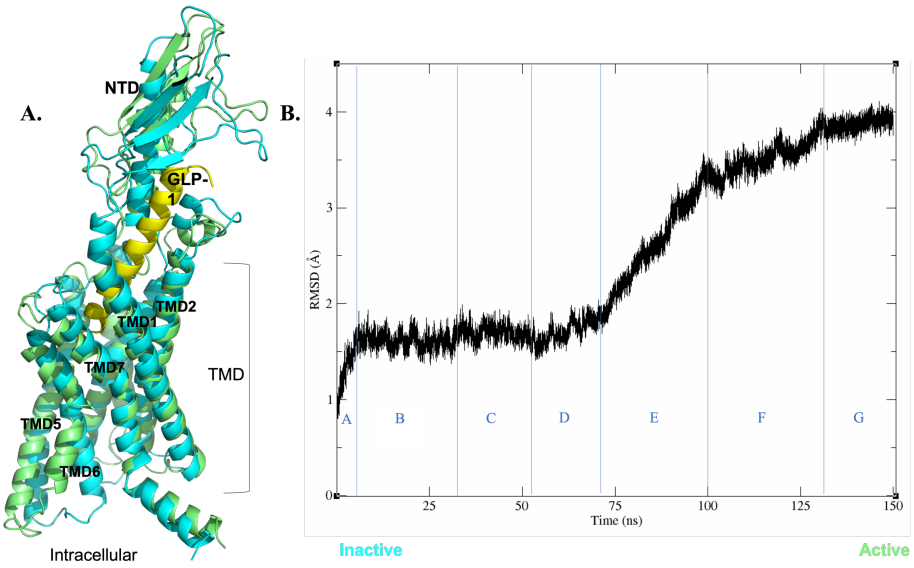


Figure 5.9. **A.** Overlaying initial inactive (cyan) and resulting active (green) GLP-1R after 150 ns of targeted MD. **B.** The total trajectory was divided into seven segments (A-G, blue lines) according to RMSD values. Each segment was analysed.

Binding of the ligand activates the receptor through conformational changes that translate to the intracellular side to activate intracellular proteins and produce a biological response. Simulation of the GLP-1 induced GLP-1R activation process started form the inactive state and was guided towards the active state, with both structures coming from MDFF simulation of models based on high resolution structures (Chapter4).

The starting structure show a GLP-1R in the inactive state in complex with GLP-1, providing information about the relative orientation, position, and interactions of the ligand prior to activation. At the extracellular side of the TMD the interaction between residues Arg190$^{2.60b}$, Asn240$^{3.43b}$, Arg310$^{5.40b}$, Glu364$^{6.53b}$, Gln394$^{7.49b}$ and Thr391$^{7.44b}$ in the inactive state close the binding pocket, maintaining the receptor a closed conformation (Chapter4). However, during binding, residues at the N-terminal of the ligand interact with residues of the extracellular side of the TMD triggering further conformational changes needed for activation.

At the extracellular side is Glu138$^{1.33b}$, guiding the ligand into the binding pocket via an interaction with Lys26* at the C-terminal end of the ligand through first a salt bridge (maintained until segment C) and a hydrogen bond (until segment D). This interaction between the ligand and the receptor, strengthened by Gln23* and Glu138$^{1.33b}$, appears to facilitate the ligand to enter via TM1-TM2 interface and to position the ligand's N-terminus towards TM5-TM6. In addition, the steric effect of the ligand itself through the trajectory to the binding pocket acts as a disruptor of the TM1-TM2 interface as Thr13* disrupts Tyr148$^{1.43b}$ - Lys197$^{2.67}$ /Asp198$^{2.628b}$ hydrogen bond allowing Thr13* to interact with the now free Lys197$^{2.67}$, which provides an anchor to the ligand by directing the N-terminus towards TM5-TM6 interface at the opposite side. Meanwhile, hydrophobic residues Tyr19* and Leu20* contribute to the disruption of TM1-TM2 interface allowing the movement of the ligand into the binding pocket resulting in TM1 shift towards TM7 (Figure 5.10, Table 5.4).

As the ligand accommodates in the binding pocket, Ala8* interact with Glu387$^{7.42b}$ at the extracellular end of helix 7 acting as an anchor facilitating the counter clock rotation of helix 7, and rotation of residues at the extracellular end of TMD7 in the direction of TM6. The TM7 rotation results in the sidechain of Lys383$^{7.37b}$ forming a hydrogen bond with Asp372^{ECL3} which acts as a hinge that maintains the active position but also limits the outward movement of TM6 (Figure 5.11, Table 5.4).

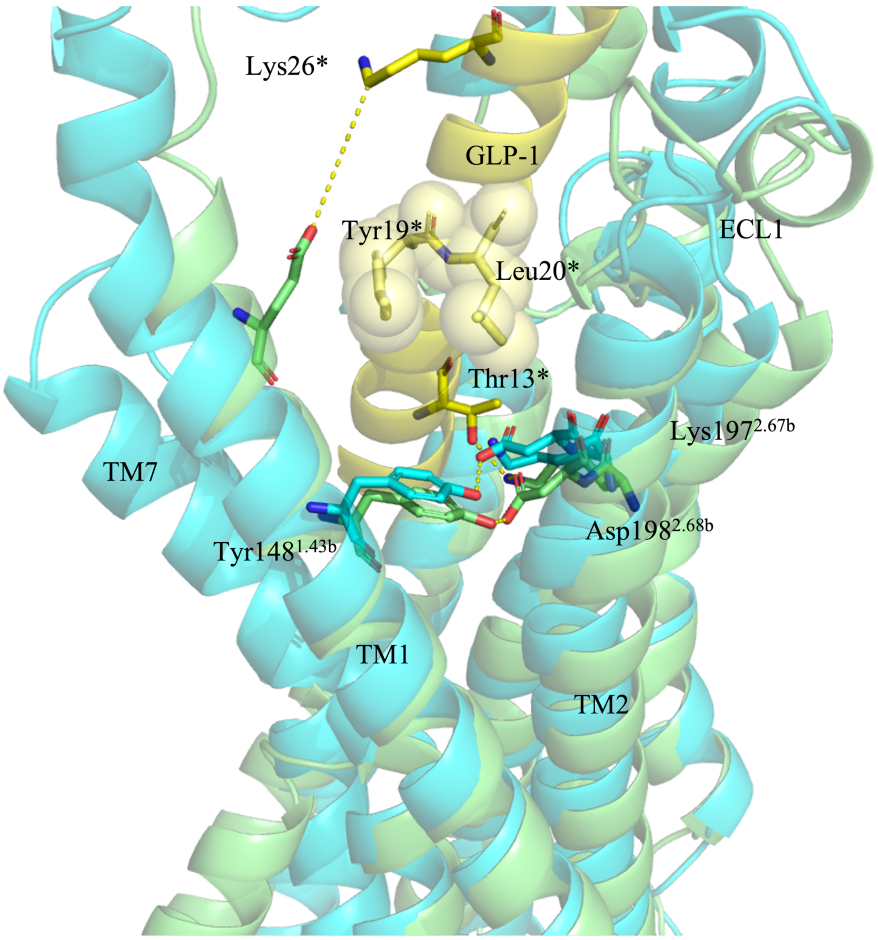


Figure 5.10. Lateral view of the extracellular TM1-TM2 interface displaying GLP-1R (green) – GLP-1 (yellow) in the active conformation and GLP-1R (cyan) in the inactive conformation. Stick representations have oxygen red and nitrogen blue. Hydrogen bonds in yellow dashed lines. Hydrophobic residues Tyr19* and Leu20* facilitate the movement of the ligand through the TM1-TM2 interface. Thr13* disrupts Tyr148$^{1.43b}$ – Lys197$^{2.67b}$ and Tyr148$^{1.43b}$ – Asp198$^{2.68b}$ hydrogen bonds allowing the ligand to be accommodate.

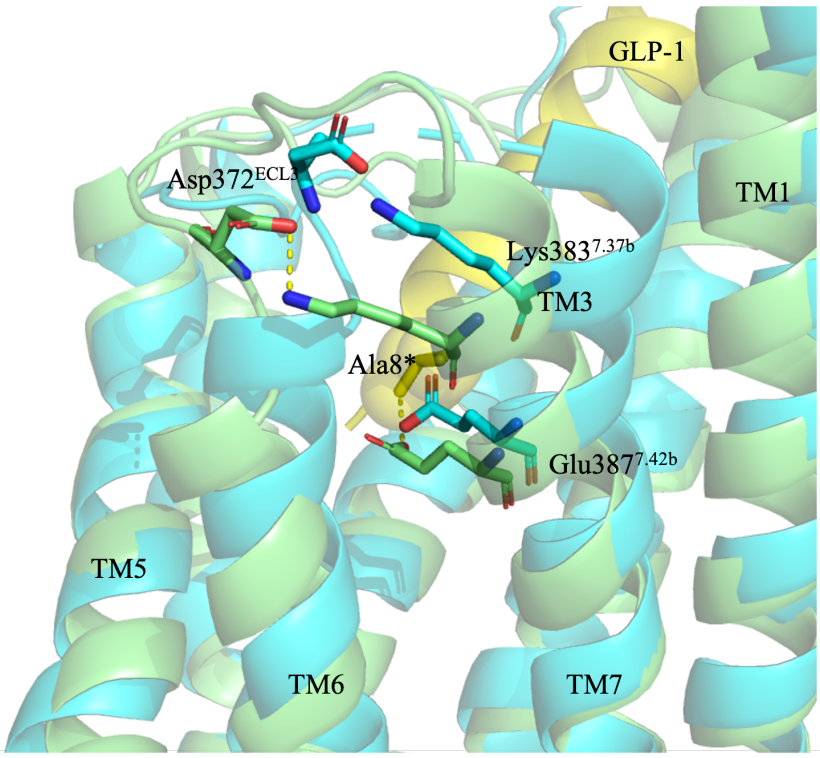


Figure 5.11. Lateral view overlaying GLP-1R (green) – GLP-1 (yellow) in the active conformation and inactive GLP-1R (cyan). Stick representations have oxygen red, nitrogen blue and hydrogen bond yellow dashed line. ECL3 and extracellular end of TM6 were hidden. Interaction between ligand and TM7 is used to accommodate the ligand into the binding pocket as it pushes TM7 towards TM6, facilitating TM6 rearrangement.

As the ligand moves into the binding pocket, the N-terminus faces the TM6-TM7 interface. In early stages His7* forms a hydrogen bond with Glu387$^{7.42b}$ facilitating helix 7 shift towards helix 6. Later, once the ligand has accommodated into the binding pocket, His7* disrupts Arg310$^{5.40b}$ – Glu364$^{6.53b}$ interaction by itself forming a hydrogen bond with Glu364$^{6.53b}$, strengthening Arg310$^{5.40b}$ – Glu364$^{6.53b}$ interaction while 'opening' TM5-TM6 interface in the extracellular half to accommodate the ligand but also allowing TM6 flexibility and rearrangement of residues needed for kink formation. By disrupting Arg310$^{5.40b}$ – Glu364$^{6.53b}$, His7* grants some flexibility to the TM6, facilitating clockwise rotation, triggering further conformational changes in helices 5 and 6, and the extracellular polar network. Although there is no direct interaction with the ligand, as the ligand's N-terminus accommodates at TM5-TM6 interface, it disrupts the hydrogen bond between Asn302^{ECL2} – Glu373^{ECL3} which is found holding the extracellular end of TM5 and

TM6 in the close conformation, however, disruption of this interaction and free TM6 extracellular end.

In helix 2 Arg190$^{2.60b}$ forms a hydrogen bond with Glu9*, this interaction is necessary for affinity and efficacy, as mutation of the positive charged residue Arg190$^{2.60b}$ in GLP-1R decreases GLP-1 efficacy (Δ Logτ_c = 0.53; Wootten, Simms, et al. 2013). Therefore, Arg190$^{2.60b}$ - Glu9* i) contributes to anchor the ligand into the binding pocket, ii) shrinks the binding pocket and iii) as a result of the mass effect of the ligand into the binding pocket shifts helix 3 in direction to TM2, improving the tightening of TM2-TM3 interface via interaction between residues Ser186$^{2.56b}$-Asn240$^{3.43b}$ (segments B, C, D) (Figure 5.12, Table 5.4). As a result, Asn240$^{3.43b}$ rearranges nearby hydrophobic residues.

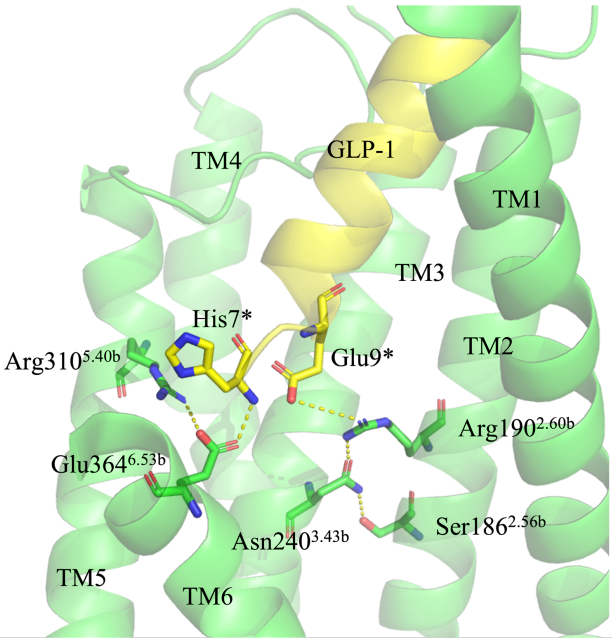


Figure 5.12. Extracellular view displaying GLP-1R (green) – GLP-1 (yellow) in the active conformation during Segment C. Stick representations have oxygen red, nitrogen blue and hydrogen bond as yellow dashed line. ECL3 and extracellular end of TM6 were hidden. Interaction between ligand and TM7 is used to accommodate the ligand into the binding pocket as it pushes TM7 towards TM6, facilitating TM6 rearrangement. Interaction between Glu9* and Arg190$^{2.60b}$ anchors the ligand pointing the ligand's N-terminus towards TM5-TM6 interface where His7* disrupts Arg310$^{5.40b}$ – Glu364$^{6.53b}$, by forming a hydrogen bond that confers TM6 some freedom allowing its rotation.

Table 5.3. Percentage of hydrogen bond occupancy and salt bridges across GLP-1R targeted MD trajectory

Interaction	Segment						
	A	B	C	D	E	F	G
Glu138$^{1.33b}$ - Lys26*	Salt bridge			21%	-	-	-
Tyr148$^{1.43b}$ - Lys197$^{2.67}$	33%	-	-	-	-	-	-
Tyr148$^{1.43b}$ - Asp198$^{2.628b}$	70%	-	-	-	-	-	-
Thr13* - Lys197$^{2.67}$	68%	60%	51%	59%	49%	51%	45%
Ala8* - Glu387$^{7.42b}$	90%	98%	98%	97%	97%	93%	92%
Lys383$^{7.37b}$ – Asp372^{ECL3}	52%	77%	74%	73%	55%	69%	58%
His7* main chain - Glu387$^{7.42b}$	80%	76%	78%	79%	74%	79%	74%
His7* - Glu364$^{6.53b}$	-	78%	78%	79%	83%	86%	76%
Glu364$^{6.53b}$ - Arg310$^{5.40b}$	60%	75%	83%	93%	86%	90%	43%
Ser186$^{2.56b}$- Asn240$^{3.43b}$	24%	31%	35%	33%	15%	18%	27%
Leu359$^{6.48b}$ main chain – Gln394$^{7.49b}$	-	-	-	-	-	10%	53%
Gln394$^{7.49b}$ – His363$^{6.52b}$	44%	51%	41%	32%	6%	6%	-
Gln394$^{7.49b}$ - Arg190$^{2.620b}$	36%	16%	5%	6%	-	-	-
Hydrogen bond occupancy between side chains, unless otherwise marked.							

In the middle region of the receptor, below the extracellular polar network, residues Leu183$^{2.53b}$, Leu244$^{3.47b}$, Leu356$^{6.45b}$, Leu360$^{6.49b}$, Val398$^{7.53b}$ form a hydrophobic network. In the inactive state, sidechains of these hydrophobic residues maintain the packing of the receptor core, but during activation these residues coordinate conformational changes between the extracellular polar network to the intracellular side allowing the rotation, outward moving and kink formation of TM6, while maintaining the packing of the core. Tightening of TM2-TM3 at the middle region of the core via Ser186$^{2.56b}$- Asn240$^{3.43b}$ facilitate Leu244$^{3.47b}$ rotamer pushing Leu360$^{6.49b}$ from the core of the receptor in the inactive state in direction to TM7 (table 5.5). Due to Leu244$^{3.47b}$ sidechain movement removes spatial restraints from Leu183$^{2.53b}$ as Leu244$^{3.47b}$ discreetly shift Leu183$^{2.53b}$ into helix 2. This allows i) the formation of a small cavity for Leu356$^{6.45b}$ and Leu360$^{6.49b}$ TM6 rotation and kink formation (segment D), ii) the stabilisation of helix 5 outward movement by forming a hydrogen bond with Asn320$^{5.50b}$ sidechain (segments F and G) and iii) shifting of Leu360$^{6.49b}$ which destabilises TM6, promoting TM6 kink formation and outward movement. Without restraints in the interior of the receptor and as consequence of the destabilisation of TM5, TM6 and TM7 initiated at their extracellular end with the binding of the ligand, the middle region of helix 6 formed forms a kink that allows

the outward movement of TM6's intracellular half. The formation of the kink is allowed by Gly361$^{6.50b}$, part of the conserved PxxG motif (Pro358$^{6.47b}$, Leu359$^{6.48b}$, Leu360$^{6.49b}$, Gly361$^{6.50b}$) and located a helical turn below Glu364$^{6.53b}$. As Leu244$^{3.47b}$ shifts Leu360$^{6.49b}$, which rotates clockwise due to Gly361$^{6.50b}$ low helical propensity. The TM6 kink resulting from TM6 outward movement is stabilised by the rearrangement of hydrophobic residues Leu183$^{2.53b}$, Leu244$^{3.47b}$, Leu360$^{6.49b}$ and Val398$^{7.53b}$ in the core of the receptor and by a shifting hydrogen bond between Gln394$^{7.49b}$ – His363$^{6.52b}$ during the first half of the simulation, and stabilised after the TM6 outward movement by a hydrogen bond between Leu359$^{6.48b}$ main chain and Gln394$^{7.49b}$ (Figure 5.13, Table 5.4).

In the intracellular side the polar ionic lock His180$^{5.50b}$, Glu247$^{3.50b}$, Thr353$^{6.42b}$ and Tyr402$^{7.57b}$, corresponding to HETx motif in class B and resembling DRY in class A maintains the receptor in the inactive state. During activation, as consequence of the rearrangement of residues in TM3 and in the intracellular side, and TM6 outward movement disrupt His180$^{5.50b}$- Glu247$^{3.50b}$ interaction (segment E to G), allowing receptor - G protein interaction. In the final stages of the simulation (segment F and G) residues Glu247$^{3.50b}$, Thr353$^{6.42b}$ and Tyr402$^{7.57b}$ are found interacting with residues of the G protein (Figure 5.14).

Table 5.4. Percentage of chis angles through targeted MD of GLP-1R

Residue	chi		A	B	C	D	E	F	G
His180$^{2.50b}$	x1	g+	0	0	0	0	0	0	0
		g-	100	100	100	100	100	100	100
		t	0	0	0	0	0	0	0
	x2	g+	15	6	6	3	2	3	0
		g-	85	94	94	97	97	97	0
		t	0	0	0	0	0	1	100
Leu183$^{2.53b}$	x1	g+	0	0	0	0	0	0	0
		g-	0	1	0	2	5	2	28
		t	100	99	100	98	95	98	72
	x2	g+	98	97	95	93	95	97	74
		g-	0	0	0	3	0	0	0
		t	2	3	5	4	5	3	26
Ser186$^{2.56b}$	x1	g+	100	100	100	100	10	33	88
		g-	0	0	0	0	0	0	0
		t	0	0	0	0	90	67	12
Arg190$^{2.60b}$	x1	g+	0	0	0	0	0	0	0
		g-	0	0	0	0	0	0	0
		t	100	100	100	100	100	100	100
	x2	g+	0	0	0	13	100	100	100
		g-	0	0	0	0	0	0	0
		t	100	100	100	87	0	0	0
	x3	g+	0	0	0	0	0	0	0
		g-	0	0	0	0	0	0	0
		t	100	100	100	100	100	100	100
	x4	g+	0	0	0	0	0	0	0
		g-	0	0	0	13	100	100	100
		t	100	100	100	87	0	0	0
Asn240^{b}	x1	g+	0	0	0	0	0	0	0
		g-	100	100	100	100	100	100	100
		t	0	0	0	0	0	0	0
	x2	g+	0	0	0	0	6	5	10
		g-	0	0	0	0	0	0	0
		t	100	100	100	100	94	95	90
Leu244	x1	g+	0	0	0	0	0	0	0
		g-	0	0	0	0	38	100	100
		t	100	100	100	100	62	0	0
	x2	g+	100	100	100	100	61	0	0
		g-	0	0	0	0	0	0	0
		t	0	0	0	0	39	100	100
Glu247	x1	g+	0	0	0	0	0	0	0
		g-	100	100	100	0	97	20	2
		t	0	0	0	100	3	80	98
	x2	g+	0	0	0	13	0	0	0
		g-	100	100	100	0	93	100	99
		t	0	0	0	87	7	0	1
	x3	g+	33	54	40	0	49	5	4
		g-	17	15	25	0	32	79	33
		t	50	32	35	100	18	16	63
Asn320	x1	g+	0	0	0	0	0	0	0
		g-	0	2	0	0	3	45	89
		t	100	98	100	100	97	55	11
	x2	g+	0	66	53	73	56	39	51
		g-	12	27	23	27	39	35	0
		t	88	7	24	0	5	26	49
Thr353	x1	g+	0	0	0	0	16	0	0
		g-	100	100	100	100	84	100	100
		t	0	0	0	0	0	0	0
Leu356	x1	g+	0	0	0	0	0	0	0
		g-	0	0	2	0	17	41	80
		t	100	100	98	100	83	59	20
	x2	g+	100	97	96	98	79	53	16
		g-	0	0	0	0	1	6	14
		t	0	3	4	2	20	41	70
Pro358	x1	g+	80	70	80	84	73	60	17
		g-	20	30	20	16	27	40	83
		t	0	0	0	0	0	0	0
	x2	g+	20	30	20	16	27	40	83
		g-	80	70	80	84	73	60	17
		t	0	0	0	0	0	0	0
Leu359	x1	g+	0	0	0	0	0	0	0
		g-	0	26	3	0	73	99	100
		t	100	74	97	100	27	1	0
	x2	g+	97	72	90	93	25	1	2
		g-	0	1	0	0	0	0	0
		t	3	27	10	7	75	99	98
Leu360	x1	g+	0	0	0	0	0	0	3
		g-	98	61	51	81	82	82	97
		t	2	39	49	19	18	18	0
	x2	g+	2	41	51	21	27	21	12
		g-	0	0	2	1	1	0	1
		t	98	59	46	78	72	79	87
His363	x1	g+	0	0	0	0	0	0	0
		g-	100	100	100	96	3	0	0
		t	0	0	0	4	97	100	100
	x2	g+	98	99	100	93	1	8	20
		g-	0	0	0	7	98	90	56
		t	2	1	0	0	0	1	24
Val398	x1	g+	32	0	0	43	2	3	4
		g-	10	10	0	0	0	0	0
		t	58	90	100	57	98	97	96
Tyr402	x1	g+	0	0	0	0	0	0	0
		g-	100	100	100	100	100	100	100
		t	0	0	0	0	0	0	0
	x2	g+	27	26	11	3	5	12	0
		g-	38	44	31	1	68	85	100
		t	35	29	58	96	27	3	0

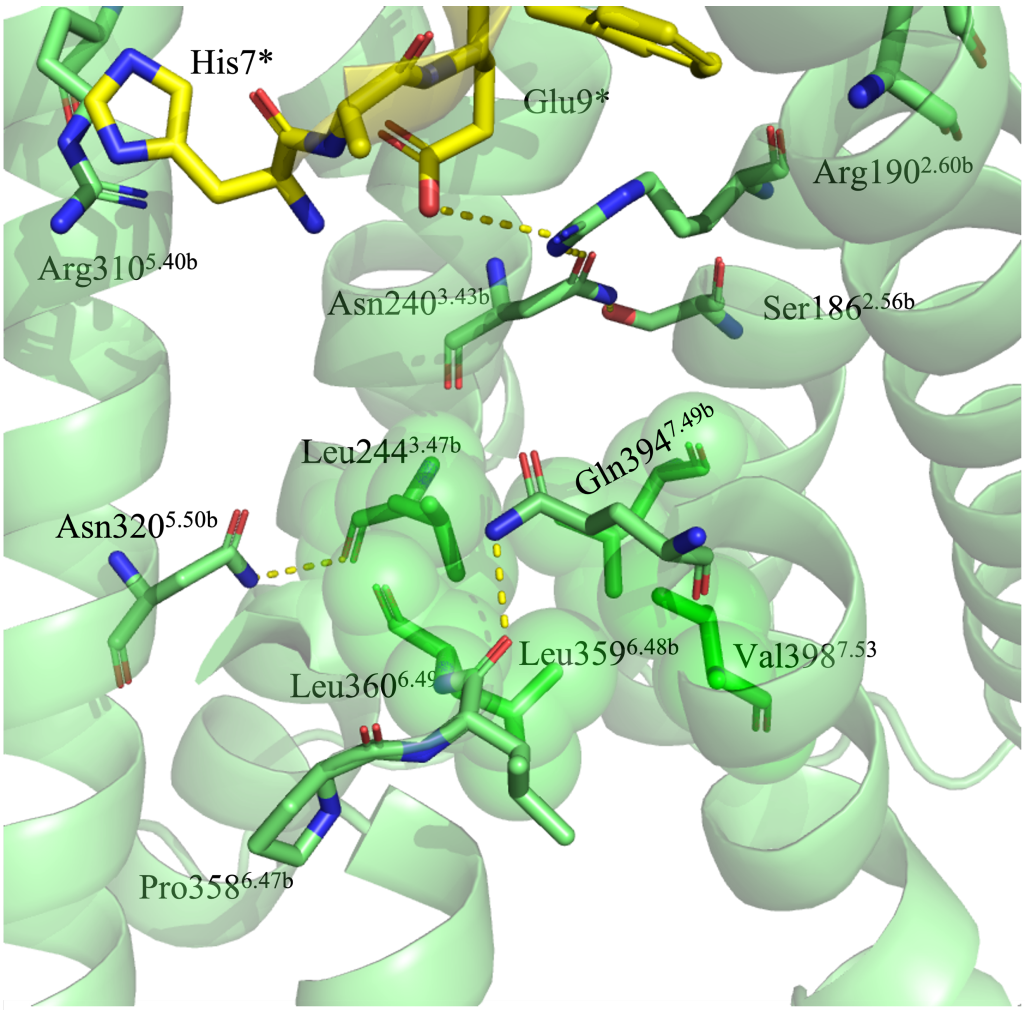


Figure 5.13. Lateral view of the intracellular residues in the active conformation, GLP-1R (green) and GLP-1 (yellow). Stick representations have oxygen red, nitrogen blue and hydrogen bonds as dashed yellow lines. Extracellular end of TM6 and TM7 were hidden. Hydrophobic residues (spheres) interaction lock the core of the receptor and Gln394$^{7.49b}$ and Leu359$^{6.48b}$ stabilise TM6 kink.

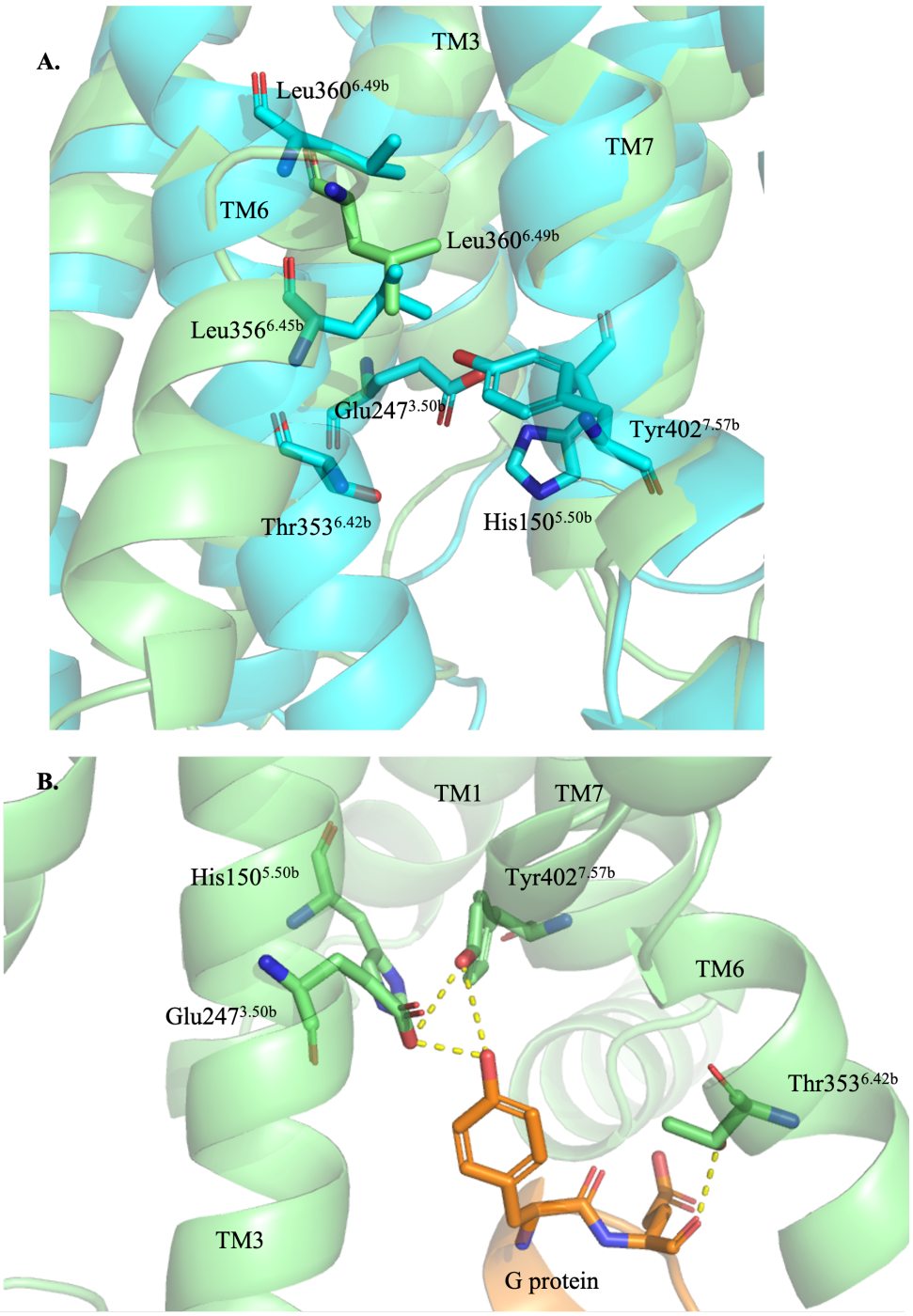


Figure 5.14. A. Overlay of inactive (cyan) and active (green) GLP-1R. TM6 outward movement, and Leu360$^{6.49b}$ contribute to the disruption of the intracellular ionic lock. **B.** Lateral view of the intracellular residues belonging to the ionic lock in the active conformation, GLP-1R (green) – G protein (orange) during segment G. Stick representations have oxygen red, nitrogen blue and hydrogen bonds as dashed yellow lines. ICL3 was hidden.

5.4. Discussion

The aim of this simulation was to identify the conformational changes as well the interactions leading to activation, therefore, starting with an inactive model and guiding it to an active state.

The first noticeable interaction within the movement of the ligand into the binding pocket is Glu138$^{1.33b}$ and Lys26*, as a salt bridge followed by a hydrogen bond. This interaction binds the C-terminus of the ligand to the NTD of the receptor, in accordance with the first step of the two-steps model. It appears that disruption of interactions between TM1-TM2 improves the binding of the ligand facilitating, as recently the binding of a PAM between TM1-TM2 interface allowed the full activation of a GLP-1R by GLP-1R(9-36) (Bueno et al. 2020), and might not be only related to allow the access of the ligand into the orthosteric pocket, but also highlights the importance of TM1 shift over TM7 and TM7 shift over TM6 resulting in the rearrangement of interactions that hold TM6 in the inactive state.

Once Glu9*-Arg190$^{2.60b}$ and Ala8* - Glu387$^{7.42b}$ provide some stability to the N-terminus of the ligand, His7* disrupts Arg310$^{5.40b}$ – Glu364$^{6.53b}$ allowing TM6 flexibility for rotation and outward movement by removing the restraints that holds the extracellular end in a close conformation, parallel to TM5.

It is unclear how TM6 kink is stabilised at TM6-TM7 interface. A hydrogen bond between Gln394$^{7.49b}$ sidechain and Pro358$^{6.47b}$ main chain stabilised TM6 kink during the active state of GLP-1R MDFF simulation (Chapter 4). However, during the course of the targeted simulation, interaction by a hydrogen bond between His363$^{6.52b}$ and Gln394$^{7.49b}$ is slowly overcome as TM6 rotation and outward movement and is stabilised by the now exposed Leu359$^{6.48b}$ main chain and Gln394$^{7.49b}$. This is interesting as His363$^{6.52b}$ mutation produces a significant decrease in efficacy in GLP-1 (Wootten, Reynolds, Koole, et al. 2016), but also affects oxyntomodulin and exendin-4 mediated cAMP signalling (Wootten, Simms, et al. 2013). Since in glucagon receptor, the extension of TM6 determines the selectivity towards the intracellular protein (Qiao et al. 2020), this would be the same for GLP-1R, and the importance of His363$^{6.52b}$ would rely on its imidazole ring determining the extension of TM6 outward movement.

Hydrophobic residues Leu183$^{2.53b}$, Leu244$^{3.47b}$ and Leu360$^{6.49b}$ stabilise the core in the inactive state; however, during activation Leu244$^{3.47b}$ triggers rotamer

changes, creating a small cavity allowing the rotation of TM6 - Leu359$^{6.48b}$ / Leu360$^{6.49b}$. However, this would not be possible without the Asn240$^{3.43b}$ sidechain since this residue packs TM2-TM3 middle region removing spatial restraints allowing Leu244$^{3.47b}$ rotamer change. Indeed, the Asn240$^{3.43b}$ mutation to alanine decreases GLP1-R efficacy but not the mutation to glutamine (Wootten et al. 2016b; Wootten, Simms, et al. 2013). Once Asn240$^{3.43b}$ packs TM2-TM3, hydrophobic residues are able to be rearranged and the kink is formed and stabilised by Gln394$^{7.49b}$. However, mutagenesis data show no effects on efficacy when only Gln394$^{7.49b}$ is mutated but significatively decreases when Asn240$^{3.43b}$ is mutated and even more with Asn240$^{3.43b}$ / Gln394$^{7.49b}$ double mutation (Wootten, Reynolds, Smith, et al. 2016). The absence of Asn240$^{3.43b}$ hinders Leu244$^{3.47b}$ rotation and rearrangement of hydrophobic residues and therefore TM6 outward movement; but in addition, the lack of Asn240$^{3.43b}$/Gln394$^{7.49b}$, not only results in an unstable arrangement of hydrophobic residues, but also removes Gln394$^{7.49b}$, a residue that stabilises TM6 kink. In addition, Leu360$^{6.49b}$ in GLP-1R resembles the β2AR residue Ile121$^{3.40}$, which rotamer change induce TM6 outward movement.

5.5. Conclusion

Targeted MD simulations were performed for the study of GLP-1R activation. Targeted MD simulations in GPCRs were verified using the β2AR as a control, since its activation mechanisms that are well characterised. Targeted MD of the β2AR showed the expected conformational changes that accompanied activation. Therefore, a targeted MD of the GLP-1R was performed. The resulting active state showed differences in the final conformation, but also confirmed i) the movement of the ligand via TM1-TM2 interface; ii) the importance of charged Arg190$^{2.60b}$ -Glu9* interaction in binding of the ligand; iii) the importance of TM7 being shift by Ala8* during activation; iv) role of hydrophobic residues triggering activation; and v) the interaction between the side chain of Gln394$^{7.49b}$ with the main chain of Pro358$^{6.47b}$ stabilises the TM6 kink.

Chapter 6: PTH1 receptor

6.1. Introduction

The PTH_1 receptor is of clinical interest mainly due to its role in calcium homeostasis and treatment of osteoporosis. Currently, two analogue drugs, teriparatide (PTH analogue) and abaloparatide (PTHrP analogue), are used for the treatment of osteoporosis and both significantly reduce the occurrence of fractures compared with other treatments (bisphosphonates, raloxifene, and monoclonal antibodies/RANK ligand inhibitors), despite their higher cost and limited usage due to subcutaneous administration (Fan et al. 2020; Reginster et al. 2019; Lewiecki 2006).

Before 2018 only ECD structures or NTD fragments (PDB codes 1BL1, 3C4M, 3H3G, 3L2J, 4Z8J were available (Pellegrini et al. 1998; Pioszak and Xu 2008; Pioszak et al. 2009; Pioszak et al. 2010; Clairfeuille 2015). In 2018, the thermostabilised crystal structure of a full-length PTH_1 receptor at 2.5 Å resolution PDB code 6FJ3 (Ehrenmann et al. 2018) became available, shortly followed in 2019 by the cryo-EM structure of a PTH_1 in the active state at 3.0, 3.5 and 4.0 Å resolution (PDB codes 6NBF, 6NBH and 6NBI; Zhao et al. 2019). Both structures display PTH_1 receptors in complex with a peptide agonist ligand, where the C-terminal end of the ligand is interacting with the NTD of the receptor and the N-terminal end of the ligand with the TMD, confirming PTH_1 receptor two-step ligand binding mechanism, characteristic of Class B (Bergwitz et al. 1996; Pioszak and Xu 2008).

Although the thermostabilised 6FJ3 and the cryo-EM structures 6NBF, 6NBH and 6NBI represent the PTH_1 receptor, there are conformational differences within the structures. The transmembrane region captured in the thermostabilised receptor 6FJ3 shows a receptor in an intermediate state, where the extracellular region is in transition towards activation, while the intracellular region remains in the inactive state (Ehrenmann et al. 2018). It is important to remember that the thermostabilised PTH_1 receptor was engineered and subject to multiple structural modifications in order to enhance thermostability through a combined approach of directed evolution in yeast and alanine scanning mutagenesis and the addition of the fusion protein *Pyrococcus abyssi* glycogen synthase (PGS) at the ICL3 (Ehrenmann et al. 2018); in

addition, during the process the ECD was separated from the TMD and residues S61-R104 were removed to prevent proteolytic cleavage (Klenk et al. 2010). Then, the TMD was re-joined with the ECD in complex with a PTH mimetic agonist (ePTH) which improved the stabilise the receptor. The resulting receptor had a slightly reduced affinity and potency from than the wild-type PTH(1-34) (Ehrenmann et al. 2018). Due to the multiple structural changes and modifications used to improve crystallisation the resulting receptor differ from the wild-type, for example the straight conformation of TM6 is likely to be the result of the PGS to favour crystallisation (Sutkeviciute et al. 2019). Finally, key interactions between the ligand $E4^{ePTH}$ and $Tyr195^{1.47b}$ and $Arg233^{2.60b}$ from the receptor are present but lead to the formation of a sparser polar network that fails to switch the receptor to the active state which might be the result of the thermostabilising mutations and the PGS insertion in the ICL3 (reviewed in Sutkeviciute et al. 2019). Therefore, the interactions found in this structure need to be carefully consider.

In contrast, cryo-EM PTH_1 receptor (PDB codes 6NBF, 6NBH and 6NBI) display a human PTH_1 receptor in complex with a long-acting PTH analogue (LA-PTH) and a G_s protein stabilised in the active state (Zhao et al. 2019). These structures show different positions of the ECD relative to the TMD, (Figure 6.1) resulting also in different position of the C-terminal of the ligand LA-PTH but similar interactions between the ligand LA-PTH N-terminal and the receptor's TMD (Zhao et al. 2019) showing the flexibility and dynamic of the ECD.

These structures confirm that the PTH_1 receptor shares Class B architecture while providing atomic information of the interaction between PTH_1 receptor and peptide agonist in different conformational states, increasing the knowledge of ligand binding mechanism and the understanding of Class B conserved mechanisms critical for activation. Similar to the recent oral GLP-1R agonist treatment for diabetes, opens the possibility of an orally available treatment for osteoporosis. Even more, these structures can be used for the further study of PTH_2 since there are no available structures but PTH_1 and PTH_2 receptors share almost 50% sequence identity PTH_2 receptor is endogenous activated by the tuberoinfundibular peptide of 39 residues (TIP39) and *in vitro* by PTH but not by PTHrP despite being able to bind PTH_2 receptor (Mann, Wigglesworth, and Donnelly 2008; Weaver et al. 2017; Usdin, Gruber, and Bonner 1995; Abou-Samra et al. 1992; Hoare, Gardella, and Usdin 2001; Sato et al. 2016b). Residue 5 from the ligand is the responsible, PTH Ile5* is able to

adopt different rotamers and bind in the region involving Leu289$^{3.37b}$ and Ile363$^{5.39b}$ in PTH_1 receptor as well Ile244$^{3.37b}$ and Tyr318$^{5.39b}$ from PTH_2 receptor, however, His5* in PTHrP is unable to bind, therefore resulting in the difference in activation. Since PTH_2 receptor is found in the central nervous system where it has been suggested to modulate affective behaviour and post-traumatic anxiety (Lee, Stephens, and Kuhn 2018; Kaouane et al. 2012) and in the skin where it may influence keratinocyte function (Sato et al. 2016a). Medical relevance for the treatment of post-traumatic anxiety (Kaouane et al. 2012) or skin disorders (Sato et al. 2016a).

6.1.1. Rationale for building and simulating cryo-EM based PTH_1 receptor models

At the time of this project, the crystal structure 6FJ3 and cryo-EM structure 6NBF were the better representation of a full-length human PTH_1 receptor in the inactive and active state respectively. Both structures show the ECD conformation similar to previous reported ECD and their respectively ligands adopt a helical conformation across the ECD and TMD, similar to other class B peptides (Ehrenmann et al. 2018; Liang et al. 2018; Liang et al. 2017; Zhang et al. 2017; Zhang et al. 2018; Zhao et al. 2019). No ECL1 was solved in either of the structures suggesting that the position of the ECD relative to TMD is defined by the ligand.

Although crystal and cryo-EM structures provided valuable information of PTH_1 receptor and engagement with agonist ligands, there are also differences with the wild-type receptor and endogenous ligand. To obtain the crystal structure 6FJ3, Ehrenmann et al. separated the ECD and TMD and individually subjected them to structural modifications to enhance their thermostability and overcome crystallisation issues before joining the ECD with the TMD. Residues belonging to the native signal peptide (Met1-Tyr23) and unstructured residues (Ser61-Arg104) were removed from the ECD to avoid proteolytic cleavage (Klenk et al. 2010). Then, the resulting ECD consisting of residues 24-177 was joined to the TMD, formed by residues 178-Ala480. In the same construct the previously reported PGS domain (Yin et al. 2015) was fused into the ICL between Lys388 ^{ICL3} and Asp398^{ICL3}. Finally, the construct was C-terminally truncated (at residue A480) before the final crystallisation construct.

A PTH mimetic agonist, ePTH, was engineered to improve the stability of the receptor, although this agonist had less affinity and potency than wild-type PTH(1-

34), it adopts a similar α-helical conformation connecting ECD and TMD, and therefore positioning the ECD relative to TMD similar to GLP-1 and glucagon ligands (Ehrenmann et al. 2018; Liang et al. 2018; Zhang, Qiao, et al. 2017; Zhang et al. 2018). Residues 1-14 from the ligand are found in binding pocket, parallel to TM2 and interacting with TM1-TM3, TM5, ECL2 and ECL3 (Ehrenmann et al. 2018). The ionic interaction between Arg20* and $Asp137^{ECD}$, essential for binding and activation and ligand C-terminal -ECD hydrophobic interactions (Weaver, Wigglesworth, and Donnelly 2014) are seen in the 6FJ3 structure. At the central polar network, agonist Glu4* forms a hydrogen bond with $Arg233^{2.60b}$, similar to equivalent residues in GLP-1R-GLP-1 (Glu9*-$Arg190^{.60b}$); and facilitates $Arg233^{2.60b}$ direct contact with $Gln451^{7.49}$. It is important to remember that this structure is found in a metastable state where the $Arg190^{.60b}$- $Gln394^{7.49b}$ interaction with $Pro358^{6.47b}$ is suspected to improve TM6 kink stability (Ehrenmann et al. 2018; Zhang et al. 2017). Unlike the glucagon receptor, no stalk is observed, although the B factors of the region corresponding to $Thr175^{1.27b}$-$Arg179^{1.31b}$ suggest higher flexibility and no interaction between the ligand and this region are seen (Ehrenmann et al. 2018). Unlike GLP-1R and GCGR where ECL1 stabilises ligand-bound ECD (Liang et al. 2018; Zhang, Qiao, et al. 2017; Zhang et al. 2018; Jazayeri et al. 2017), PTH_1 receptor long ECL1 is not resolved and might have no effect in ligand-ECD stability or activation (Ehrenmann et al. 2018; Yang et al. 2015b; Zhang, Qiao, et al. 2017). However, due to the thermostabilising process, point mutations, and the presence of the PGS in the ICL3, the interactions found in this structure need to be consider with caution.

The cryo-EM structure displays a human PTH_1 receptor in the active state in complex with LA-PTH ligand, G_s protein and a stabilising nanobody Nb35 (Zhao et al. 2019; Rasmussen, Choi, et al. 2011a). LA-PTH ligand is more potent than the endogenous peptide, and it is able to stabilise PTH_1 receptor either G protein free or bound states (Maeda et al. 2013; Hattersley et al. 2016; Shimizu et al. 2016). With cryo-EM data suggesting multiple simultaneous conformations these were subdivided into three different subclasses according to the position of the ECD relative to the TMD, (Figure 6.1) resulting in different position of ECD and LA-PTH C-terminal but not LA-PTH N-terminal as the interaction with the receptor's TMD is similar in the three structures (Zhao et al. 2019). Two conformations show a similar rearrangement of the ECD and LA-PTH C-terminal rotated 15º against each other, however, in the third one LA-PTH C-terminal is bent losing interaction with the ECD and with less

resolved density map; the three structures were solved at different resolutions (3.0, 3.5 and 4.0 Å; Figure 6.11) (Zhao et al. 2019).

The PTH_1 receptor structures, both crystal and cryo-EM, show the position of the ECD relative to the TMD as being slightly different to GLP-1R. In the human PTH_1 receptor, ECL1 is not solved, while in GLP-1R ECL1 adopts a helical conformation within distance of the GLP-1 C-terminal possible affecting the ligand and its binding (Jazayeri et al. 2017; Zhang et al. 2017; Zhao et al. 2019). Similarly, in Class B GPCRs active receptors a kink around PxxG motif in TM6 is seen allowing, with TM6 outward movement, the formation of an intracellular cavity for G protein binding.

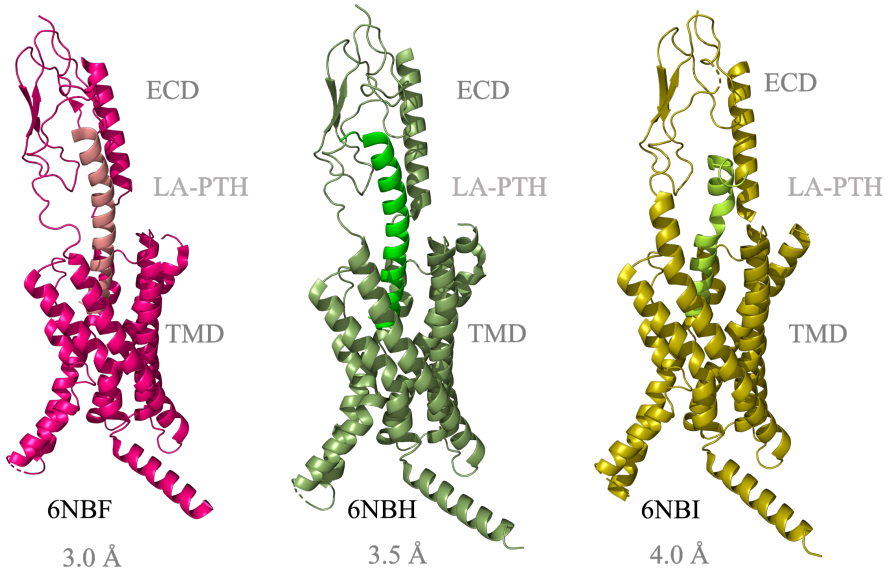


Figure 6.1. Cryo-EM structures of the human PTH_1 receptor in the active state at different resolution. Different position of the ECD and ligand LA-PTH relative to the ECD suggest three different activation sub-states.

6.1.2. Aims and objectives

Although having different physiological effect, Class B receptors share an architecture and possible similar activation mechanisms. In order to identify conserved interactions in the active and inactive state, as well to understand activation mechanisms in Class B GPCRs, models of an active and inactive human PTH_1 receptor model in complex with human PTH(1-34) ligand (and G_s protein in the active model) were built and simulated using electron density maps; then, the resulting

structures were used to investigate conformational changes leading to activation using targeted MD.

6.2. Methods

Model building manipulations, alignments and *in silico* mutations were done using tools from PyMOL software (Schrödinger 2015) unless otherwise stated.

6.2.1. Homology modelling

6.2.1.1. Building inactive PTH_1

Cryo-EM structure 6FJ3 was used as a template to build the inactive state of human PTH_1 receptor (Ehrenmann et al. 2018). The structure represented in 6FJ3 represents a thermostabilised structure in an intermediate state, however its intracellular TMD is in the inactive conformation, and since the structure was solved in complex with a peptide agonist ligand (ePTH) it provides valuable information about early ligand-receptor interactions.

The 6FJ3 structure was used to build the model of the PTH1 receptor in the inactive state. Since this receptor lacks the ECL1, the ECL1 segment of the glucagon 5YQZ structure was used to build the missing residues of the ECL1. (i) All residues not belonging to the receptor were removed from 6FJ3 structure and only the receptor (ii) The glucagon receptor PDB 5YQZ was aligned to the PTH_1 receptor. Homologous residues from the ECL1 of 5YQZ were used to build PTH_1 receptor ECL1; 5YQZ segment including R201-K205 and V221-D208 corresponding to PTH_1 receptor G247-D251 and Y278-A265 respectively (Figure 6.2). The merged segments were renumbered according to PTH_1 receptor sequence (iii) SWISS-MODEL homology modelling server (Waterhouse et al. 2018) was used to build the human PTH_1 receptor using the construct model and PTH_1 receptor sequence (UniProt Q03431). During the process, missing residues from ECL1 (E252-I264), ICL3 (A394-T399) and missing side chains were built (iv) the resulting receptor was saved into a new file. (v) ePTH was *in silico* mutated to PTH(1-34) (Table 6.1) (vi) The

inactive state receptor model was aligned to the receptor, while PTH(1-34) was aligned using ePTH coordinates in 6FJ3 structure. (vii) PTH(1-34) and PTH_1 receptor complex was saved into a single file (ix) the model was subjected energy minimisation through steepest descent method (50 steps), followed by Newton Raphson methods (50 steps) to optimise minor steric clashes and poor bond/torsion geometries using NAMD (Bernardi et al. 2016).

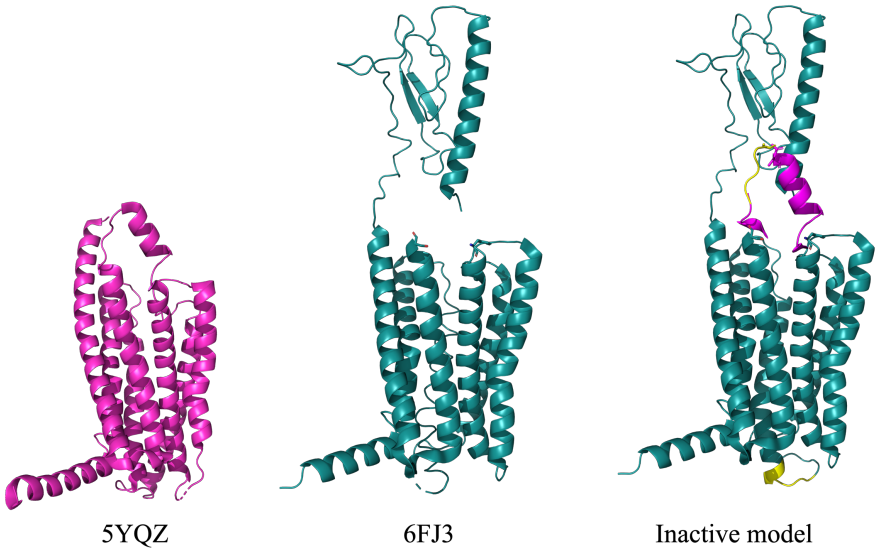


Figure 6.2. Building of the inactive PTH_1 receptor model. To build the inactive PTH_1 receptor the 6FJ3 structure (green) was used as a template. PTH_1 receptor has a long and flexible ECL1; due to the flexibility of the region, this was not resolve in the 6FJ3 structure. Therefore, the homologous residues of ECL1 in the glucagon receptor 5YQZ (pink) were used. However, the ECL1 in the PTH_1 receptor is longer than in the GCGR 5YQZ, and the missing residues from ECL1 and ICL3 were built using SWISS-MODEL (yellow). ECD of 5YQZ and ligand were hidden.

6.2.1.2. Building active PTH_1

Crystal structure 6NBF was used as a template to build the active state of human PTH_1 receptor (Zhao et al. 2019). (i) All residues not belonging to the receptor were removed from 6NBF structure and was saved into a file. (ii) ECL1 from the inactive PTH_1 receptor was used to build missing ECL1 in 6NBF. (iii) SWISS-MODEL homology modelling server (Waterhouse et al. 2018) was used to model human PTH_1 receptor in the active state and human PTH_1 receptor sequence (UniProt Q03431). During the process missing residues and side chains were built (iv) the

resulting receptor was saved into a new file. (v) LA-PTH was *in silico* mutated to PTH(1-34) (Table 6.1). Because LA-PTH is shorter than PTH(1-34) missing first and last residues were added using PTH(1-34) ligand found in the inactive model. PTH(1-34) ligand from the inactive model was aligned to mutated LA-PTH and only first and last residues were added. The new PTH(1-34) ligand was renumbered (vi) the active model was aligned to the receptor, PTH(1-34) was aligned to LA-PTH coordinates and α-helical G_s protein was added according to 6NBF structure (vii) PTH(1-34), PTH_1 receptor and G protein complex was saved into a single file (ix) the model was subjected to energy minimisation through steepest descent method (50 steps), followed by Newton Raphson methods (50 steps) to optimise minor steric clashes and poor bond/torsion geometries using NAMD (Bernardi et al. 2016).

Table 6.1. Comparison of some of class B peptide ligands

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
PTH(1-34)	S	V	S	E	I	Q	L	M	H	N	L	G	K	H	L	N	S	M	E	R	V	E	W	L	R	K	K	L	Q	D	V	H	N	F	-		
PTHrP(1-36)	A	V	S	E	H	Q	L	L	H	D	K	G	K	S	I	Q	D	L	R	R	R	F	F	L	H	H	L	I	A	E	I	H	T	A	E	I	
TIP39	S	L	A	L	A	D	D	A	A	F	R	E	R	A	R	L	L	A	A	L	E	R	R	H	W	L	N	S	Y	M	H	K	L	L	V	L	D	A	P	39
ePTH	Ac5c	V	Aib	E	I	Q	L	M	H	Q	Hrg	A	K	W	L	N	S	Nle	E	R	V	E	W	L	R	K	K	L	Q	D	V	H	N	Y	-		
LA-PTH	A	V	A	E	I	Q	L	M	H	Q	R	A	K	W	I	Q	D	A	R	R	R	A	F	L	H	K	L	I	A	E	I	H	T	-			

GLP-1	H	A	E	G	T	F	T	S	D	V	S	S	Y	L	E	G	Q	A	A	K	E	F	I	A	W	L	V	K	G	R	G					
	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37						

6.2.2. Molecular Dynamics Flexible Fitting

6.2.2.1. System building

Models were orientated with respect to the membrane using OPM database (Lomize et al. 2012). Models were fitted into density maps using 'Colores' program from 'Situs' package (Wriggers 2012; Wriggers and Birmanns 2001). The active model was fitted into 6NBF density map (Zhao et al. 2019) and the inactive model into 6FJ3 density map (Ehrenmann et al. 2018). Previously, MDFF Utilizer using CCP4i Software (McNicholas et al. 2011; Potterton et al. 2003) was used to convert 6FJ3 DSN6 format into ccp4. Files containing oriented models in the membrane and models fitted into density maps were submitted into CHARMM-GUI server (Jo et al. 2008) where MDFF Utilizer (Qi et al. 2017) was used to generate the input files for

simulation. Disulphide bonds were defined, and each model was embedded into a 90 x 90 $Å^2$ heterogeneous lipid membrane (Cholesterol 25%, POPC 26%, POPE 20%, POPS 5%, POPI 5%, PSM 19%) before adding neutralizing ions (0.15 M NaCl) and a water/solvent box (TIP3P). Simulations were set at 303.15 K temperature and 1 atm pressure, using Langevin dynamics with a damping coefficient of 1-ps to control the temperature and a Nose-Hoover Langevin piston to control the pressure.

6.2.2.2. System simulation

Each system underwent equilibration for 70 ns and a total of 350 ns production using all-atoms molecular dynamics flexible fitting (MDFF) method (Trabuco et al. 2008; Trabuco et al. 2009). Simulations were done using CHARMM36 force-field (Huang and MacKerell 2013) and NAMD software (Bernardi et al. 2016) at a 2 fs timestep, using Particle Mesh Ewald to account for long-range electrostatics. In both trajectories, harmonic restraints obtained from the density map potential were applied to all Cα atoms. Restraints where steadily removed from the TMD and ligand during the first 100 ns. After 60 ns of simulation restraints from Cα atoms of the TMD in the active and inactive models were removed. Restraints on the ligand were removed after 100 ns in the active and inactive model. During the following 250 ns, corresponding to the segment of 150-350 ns from the total trajectory, restraints were only applied to Cα atoms of the NTD but TMD and ligand were simulated without restraints.

6.2.2.3. Trajectory analysis

The resulting trajectories were analysed using WORDOM (Caflisch et al. 2007) and VMD (Humphrey, Dalke, and Schulten 1996). From the total output trajectory for the active and inactive models the segment between 150 to 350 ns was in detail analysed. VMD was used to estimate the fraction of time in which all residue-residue hydrogen bonds were formed defined as heavy atom distance of 3.5 Å or less, with an angle of 25° or less.

6.2.3. Targeted Molecular Dynamics

6.2.3.1. System building

To study activation, the inactive state was set as the starting structure and the active state set as the 'target' structure. Structures obtained from the final frame of MDFF simulation of active and inactive MDFF of PTH_1 receptor were used for targeted MDs as 'target' and 'initial' structures respectively. Since this method requires the same number of atoms, G protein α5-helix was added to the inactive file but was manually moved towards the intracellular space away from the receptor. Therefore, the new file contain contains PTH_1 receptor in the inactive state in complex with PTH(1-34) and the newly added G protein α5-helix.

The inactive file was orientated with respect to the lipid bilayer utilising OPM database and server (Lomize et al. 2012) and submitted into CHARMM-GUI server (Jo et al. 2008; Jo et al. 2009; Lee et al. 2016) where membrane builder option was used to generate the input for targeted MD simulation with NAMD (Jo, Kim, and Im 2007; Jo et al. 2009; Wu, Cheng, et al. 2014; Schlitter, Engels, and Kruger 1994). The active structure was set as 'target' for targeted molecular dynamics simulation.

6.2.3.2. Targeted molecular dynamics simulation

To study activation, the inactive state was set as the starting structure and the active state set as the 'target' structure. The simulation was performed using CHARMM36 force-field (Huang and MacKerell 2013) and NAMD software (Bernardi et al. 2016) at a 2 fs timestep, using Particle Mesh Ewald to account for long-range electrostatics. The system underwent equilibration for 70 ns and was set for a total of 150 ns production using targeted MDs. Steering forces guiding the initial inactive structure towards the final 'target' active structure were only applied to Cα atoms using an elastic constant for targeted MD of 200 kcal/mol/$Å^2$ (Schlitter, Engels, and Kruger 1994).

6.2.3.3. Trajectory analysis

The resulting trajectory was analysed using WORDOM (Caflisch et al. 2007) and VMD hydrogen bonds plugin (Luo 2007) and Bio3D module in R to calculate rotamer states (Grant et al. 2006; Skjærven et al. 2014). The trajectories were divided into segments according to values in RMSD to analyse rotamer changes and hydrogen bonds.

VMD hydrogen bonds plugin (Luo 2007) was used to estimate the fraction of time in which all residue-residue hydrogen bonds were formed defined as heavy atom distance of 3.5 Å or less, with an angle of 25° or less.

Rotameric angles were obtained using torsion.pdb Bio3D package from R. In order to get rotameric angles, first, water and lipid bilayer were removed from the trajectory, then the trajectory was converted into PDB, files using WORDOM and alanine and Glycine residues were removed from the file. Then, using torsion.pdb Bio3D package from R (Grant et al. 2006), each frame was analysed and rotameric angles were saved into a txt file. Since the frames were analysed in chronological order, it was possible to identify segments with changes in rotameric angles. Special attention was paid to residues with a solvent accessible surface area (SASA) with a 20 percent difference between the inactive and active state.

SASA for each residue in each state was first calculated using "Accessible surface area and accessibility calculation for protein" version 1.2 web server from the Centre for Informational Biology, Ochanomizu University (Center for Informational Biology 2012). If there was a 20% difference in the SASA value of a residue in the inactive – active state, then the rotamer angles for this residue was analysed.

Residues with a difference of 20 percent in SASA were analysed. The trajectories were divided into segments according to values in RMSD to analyse rotamer changes and hydrogen bonds.

6.2. Results and Discussion

6.3.1. Molecular Dynamics Flexible Fitting

6.3.1.1. Trajectory analysis

RMSD values of the transmembrane domain in both models simulated with restraints and unrestrained fall in acceptable values for simulations (Theoretical and Computational Biophysics Group. 2016 ; Humphrey, Dalke, and Schulten 1996). On average the RMSD value of the TMD relative to starting models in the segment analysed, 150-350 ns, was 1.11 Å in the active model and 0.94 Å in the inactive model (Figure 6.3).

Active GPCRs are characterised by the outward movement of the intracellular half of TM6 which creates an intracellular cavity allowing the binding of G protein (Rasmussen et al. 2011). Superposition of simulated inactive and active models shows conformational differences. The most characteristic, TM6 outward movement in the active model is maintained through the simulation of the active model; as such the distance between Cα atoms of Arg219$^{2.46b}$ TM2 and Lys405$^{6.37b}$ TM6 is maintained at an average of 23.08 Å, at the same time the α5-helix of G protein remains in the intracellular cavity. In contrast, in the inactive model TM6 remained parallel to other helices with the receptor in a close conformation; the average distance between Cα atoms of Arg219$^{2.46b}$ and Lys405$^{6.37b}$ was 11.75 Å (Figure 6.3). Besides TM6 outward movement there are other structural differences between inactive and active receptor.

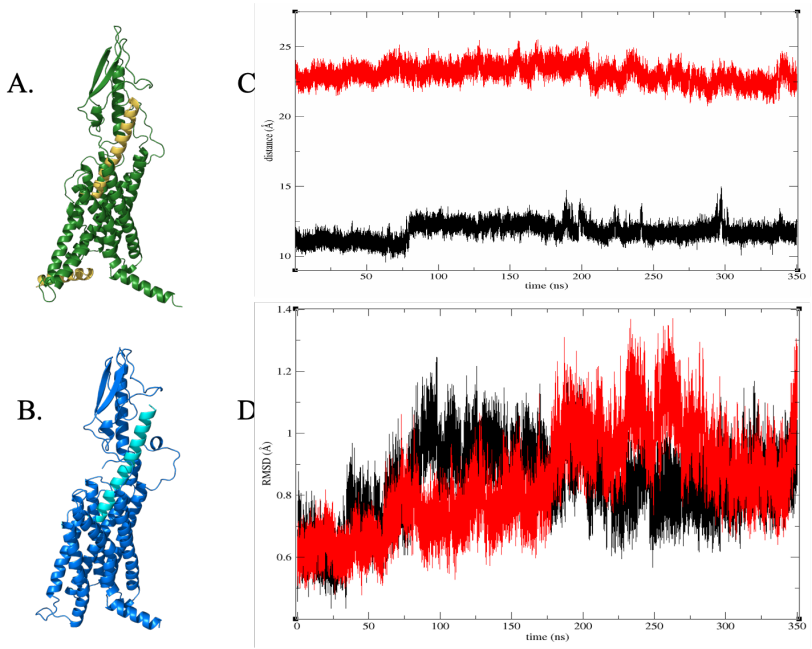


Figure 6.3. A. PTH_1 receptor (green) in complex with PTH(1-34) (yellow) in the active state. **B**. PTH_1 receptor (blue) in complex with PTH(1-34) (cyan) in the inactive state. **C.** Distance between the C atoms of $Arg219^{2.46b}$ and $Lys405^{6.37b}$ in the inactive (black) and active (red) model. **D.** RMSD of the TMD region of inactive (black) and active (red), due to its flexibility ECL1 was excluded from RMSD calculations.

Superposition of active and inactive PTH_1 receptor models show differences between states. Unlike GLP-1R where ECL1 engages with the C-terminus end of the ligand or GCGR where ECL1 has a critical role in ligand binding (Zhang et al. 2018) in PTH_1 receptor this function appears to be dependent of hydrophobic residues at TM2 and TM1 extracellular tips. Comparison of the active PTH_1 receptor model with the inactive state show the movement of TM2 extracellular end towards the core of the receptor while hydrophobic interactions between receptor $Leu244^{2.68b}$ –Leu11* and $Leu244^{2.68b}$ –Leu15*, position the ligand into the binding pocket (Figure 6.4). Experimental data show that by introducing bigger side-chains at position 11 the potency of shorter PTH analogue PTH(1-11) was enhanced (Shimizu et al. 2001), suggesting that the function of this area, either ligand position 11 or extracellular end of TM2, would be to position the ligand into the binding pocket. In this way shorter ligands with extended side chains in position 11* wouldn't need to rely on NTD for fully positioning the ligand (Figure 6.4). At the same time, the extended side chain of residue 11 of the ligand avoids the movement back of the extracellular end of TM1

parallel to TM2, as in the inactive state; facilitating TM1 extracellular movement towards TM7 by spatially restraining the interaction, such as Leu11* and Phe184$^{1.36b}$.

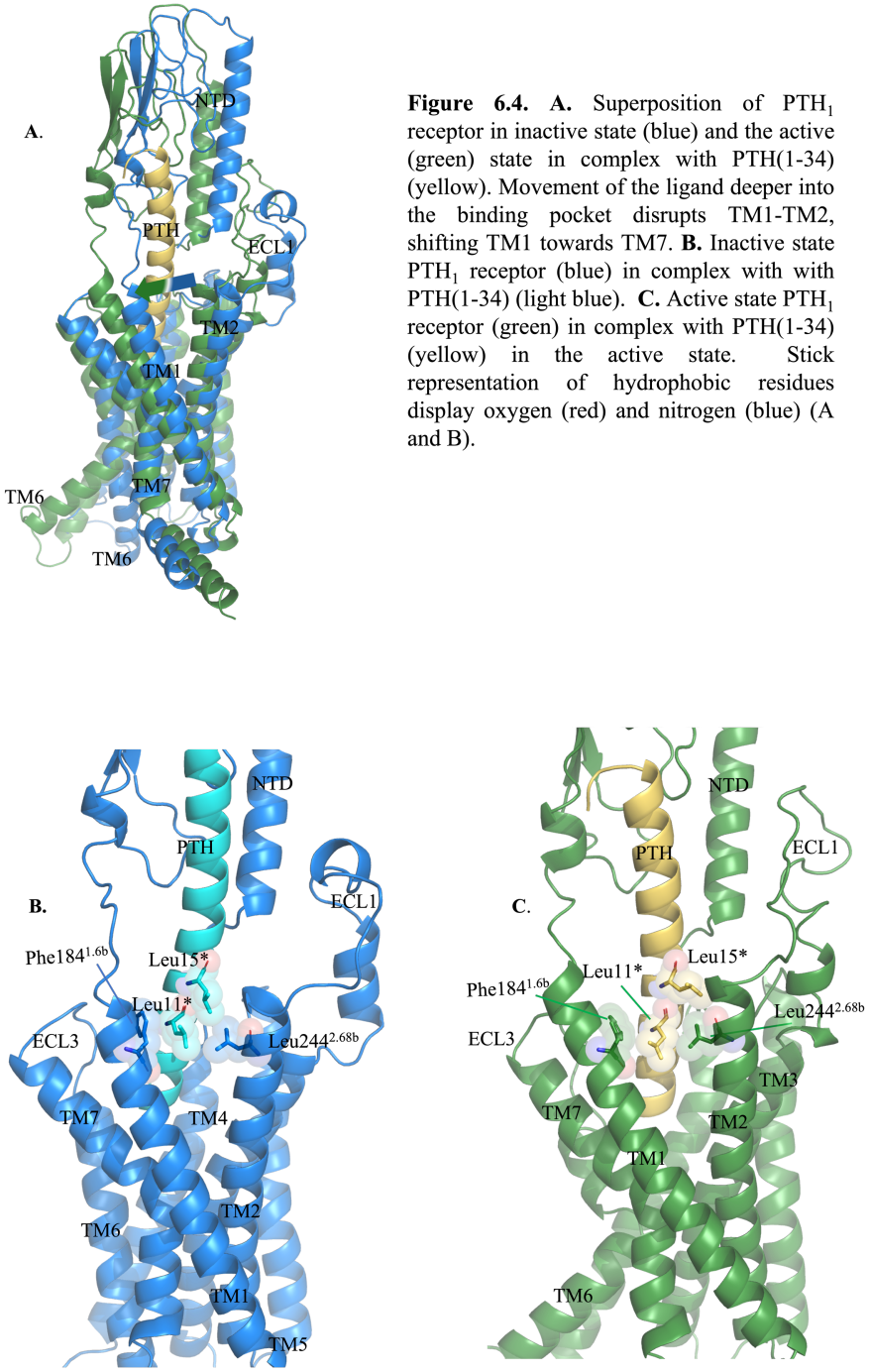


Figure 6.4. A. Superposition of PTH_1 receptor in inactive state (blue) and the active (green) state in complex with PTH(1-34) (yellow). Movement of the ligand deeper into the binding pocket disrupts TM1-TM2, shifting TM1 towards TM7. **B.** Inactive state PTH_1 receptor (blue) in complex with with PTH(1-34) (light blue). **C.** Active state PTH_1 receptor (green) in complex with PTH(1-34) (yellow) in the active state. Stick representation of hydrophobic residues display oxygen (red) and nitrogen (blue) (A and B).

6.3.1.2. Inactive PTH_1 receptor

At the extracellular end of the inactive receptor conserved residues form two polar clusters surrounding the binding pocket, unlike the inactive GLP-1R where extracellular polar residues form a single polar cluster 'closing' the binding pocket (Dods and Donnelly 2015); (Chapter 4). The first cluster connecting TM3-TM2-TM1-TM7 is formed by hydrogen bonds between $Asn295^{3.43b}$, $Arg233^{2.60b}$, $Gln451^{7.49b}$ and $Tyr195^{1.47b}$; the second cluster linking TM5-TM6 is formed by the hydrogen bond between $Gln364^{5.40b}$ - $Tyr421^{6.53b}$ (Table 6.2 and Figure 6.5) maintaining the extracellular portion of the helices in a close state. Since the inactive PTH_1 receptor model was built based on a PTH_1 receptor captured in a transitional state, and the rearrangement of the extracellular polar network is necessary for activation the lack of connection between clusters would be the earlier signs of activation, as well suggesting that changes at the extracellular polar network would be the consequence of changes in the NTD and not only by the direct interaction of the ligand into the binding pocket.

Table 6.2. Comparison between active and inactive state hydrogen bond occupancy

Hydrogen bonds	Active (%)	Inactive (%)
$Arg233^{2.60b}$ - $Asn295^{3.43b}$	93	96
$Arg233^{2.60b}$- $Gln451^{7.49b}$	0	38
$Tyr195^{1.47b}$ - $Arg233^{2.60b}$	14	82
$Tyr421^{5.53b}$– $Gln364^{5.40b}$	0	63

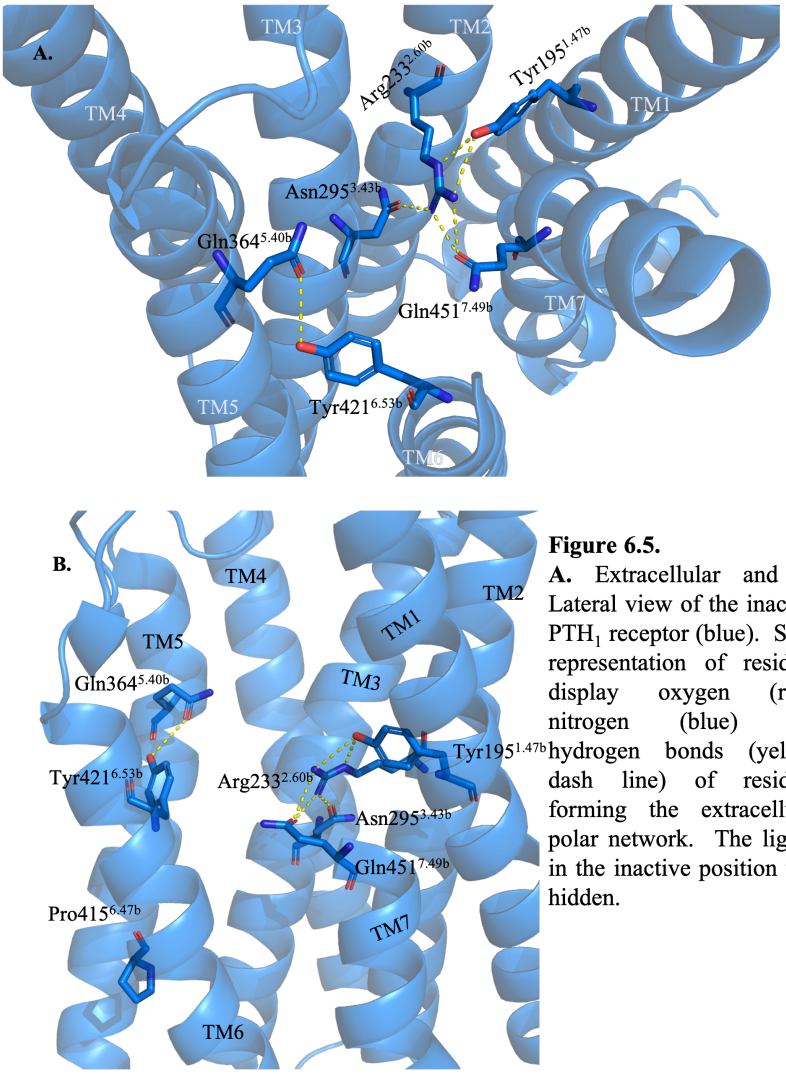


Figure 6.5.
A. Extracellular and **B.** Lateral view of the inactive PTH_1 receptor (blue). Stick representation of residues display oxygen (red), nitrogen (blue) and hydrogen bonds (yellow dash line) of residues forming the extracellular polar network. The ligand in the inactive position was hidden.

In the centre of the transmembrane region sidechains of hydrophobic residues Leu226$^{2.53b}$, Phe230$^{2.57b}$, Ile299$^{3.47b}$, Leu413$^{6.45b}$, Phe417$^{6.49b}$ and Val455$^{7.53b}$ close the core of the receptor (Figure 6.6). A similar arrangement of homologous hydrophobic residues is seen in the inactive GLP-1R simulation (Chapter 4). Two asparagine residues, Asn295$^{3.43b}$ and Asn374$^{5.50b}$, and Gln451$^{7.49b}$ connect hydrophobic residues with the extracellular polar network. When comparing active and inactive states, rotamer changes in these asparagine residues appear to facilitate the translation of conformational changes from the extracellular side towards the centre as both residues join the extracellular polar network with hydrophobic residues.

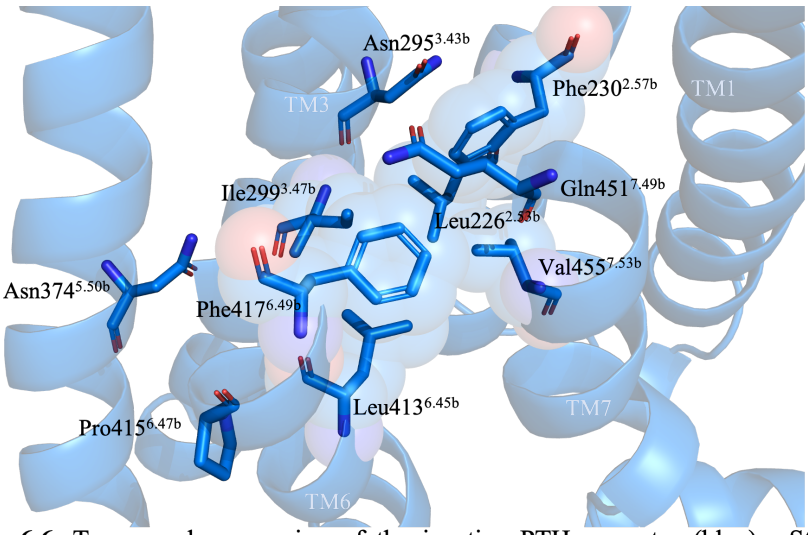


Figure 6.6. Transmembrane region of the inactive PTH_1 receptor (blue). Stick representation of residues display oxygen (red), nitrogen (blue) and hydrogen bonds (yellow dash line) of residues forming the extracellular polar network, hydrophobic residues surrounded by spheres. The ligand in the inactive position was hidden.

Similar to Class A, interaction of polar residues at the intracellular side of the inactive PTH_1 receptor maintain a close conformation. Class A well-known E/DRY motif from TM3 and TM6 form the cytoplasmic "ionic lock" and strengthen by the salt-bridge interaction with the neighbouring arginine residue in helix 3 (Rasmussen et al. 2007). In the intracellular side of the inactive PTH_1 receptor, polar residues His223$^{2.50b}$, Glu302$^{3.50b}$, Thr410$^{6.42b}$ and Tyr459$^{7.57b}$ constitute class B ionic lock HETx, similar to E/DRY. In the same way, this polar cluster is strengthened by Arg219$^{2.46b}$ and Ser409$^{6.41b}$. In addition to the ionic lock, hydrogen bonds between Arg214^{ICL1} – Glu465$^{7.63b}$– Lys405$^{7.36b}$ may contribute to maintain TM6 in the close conformation in the inactive state and later interact with the intracellular protein in the active state (Figure 6.7 and Table 6.3).

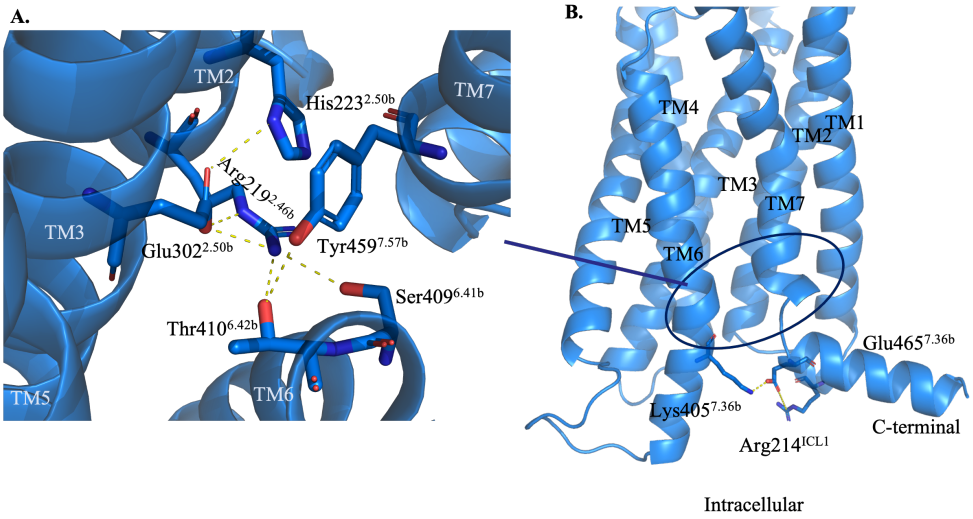


Figure 6.7. Intracellular end of the inactive PTH_1 receptor (blue) with stick representation of conserved residues displaying oxygen (red), nitrogen (blue),and hydrogen bonds (dotted yellow lines) between residues. **A.** Family B ionic lock **B.** Circle showing the location of ionic lock in the receptor (A) and in sticks residues from second polar cluster strengthening the ionic lock

Table 6.3. Comparison between active and inactive state hydrogen bond occupancy

Hydrogen bonds	Active (%)	Inactive (%)
$Tyr459^{7.57b}$ - $Thr410^{6.42b}$	0	67
$Tyr459^{7.57b}$ – $Glu302^{2.50b}$	0	25
$Glu302^{2.50b}$ – $Arg219^{2.46b}$	0	94
$Glu302^{2.50b}$ – $Thr410^{6.42b}$	0	55
$Glu302^{2.50b}$ – $His223^{2.50b}$	26	0
$Arg219^{2.46b}$ – $Ser409^{6.41b}$	0	73
$Glu465^{7.63b}$ – $Arg214^{ICL1}$	0	34
$Glu465^{7.63b}$ - $Lys405^{7.36b}$	0	15

6.3.1.3. Active PTH_1 receptor

Although activation is not seen in our simulation, the presence of the ligand in the active state disrupts polar networks causing differences between inactive and active state. Comparison of PTH_1 receptor–PTH complex with GLP-1R-GLP-1 simulation (Chapter 4) show some similarities between Class B.

The extracellular polar cluster formed by Asn295$^{3.43b}$ - Arg233$^{2.60b}$ - Gln451$^{7.49b}$ - Tyr296$^{3.44b}$ in the inactive state (Figure 6.5) is modified as Glu4* disrupts Arg233$^{2.60b}$ - Gln451$^{7.49b}$ interaction to form a hydrogen bond with Arg233$^{2.60b}$, triggering further changes. The hydrogen bond between ligand Glu4* and Arg233$^{2.60b}$ and the effect of the ligand into the binding pocket in the active state displace Arg233$^{2.60b}$ sidechain in direction to the intracellular/core of the receptor (Figure 6.8A, 6.8B). This triggers: i) the rotamer change of conserved Asn295$^{3.43b}$, which its new position affects hydrophobic residues bellow (Figure 6.6, 6.8, and 6.9) favouring rotamer changes of their side chains; ii) ease restraints on TM7, favouring TM7 extracellular end movement towards TM6. The now released Gln451$^{7.49b}$ will be able to form a hydrogen bond with Pro415$^{6.47b}$ to stabilise TM6 kink in the active state (Figure 6.8). Similar changes are seen in GLP-1R-GLP-1. Residue Glu4* from PTH(1-34) is highly conserved in class B ligands and a similar interaction is seen between homologous residues in GLP-1R-GLP-1: Glu9*, and GLP-1R Arg190$^{2.60b}$ highlighting the importance of Glu9* (Wootten et al. 2016a; Wootten et al. 2016b); (Chapter 3). Furthermore, homologous residues to Arg190$^{2.60b}$, Asn240$^{3.43b}$ and Gln394$^{7.49b}$ in GLP-1R have been identify in altering GLP-1R signalling (Wootten, Simms, et al. 2013), suggesting a role of these residues in G protein signalling.

The second polar cluster form by a hydrogen bond between Gln364$^{5.40b}$ and Tyr421$^{6.53b}$, also found in GLP-1R between homologous residues R310$^{5.40b}$ and Glu364$^{6.53b}$, is seen during the inactive state stabilising the extracellular tips of TM5 and TM6 until the N-terminus residue of the ligand breaks this interaction. Within binding of the ligand, the ligand's N-terminus S1* forms a hydrogen bond with Gln364$^{5.40b}$, and the now released Tyr421$^{6.53b}$ forms a hydrogen bond with the also released Gln451$^{7.49b}$ contributing to stabilise the TM6 kink (Figure 6.8 and Table 6.4).

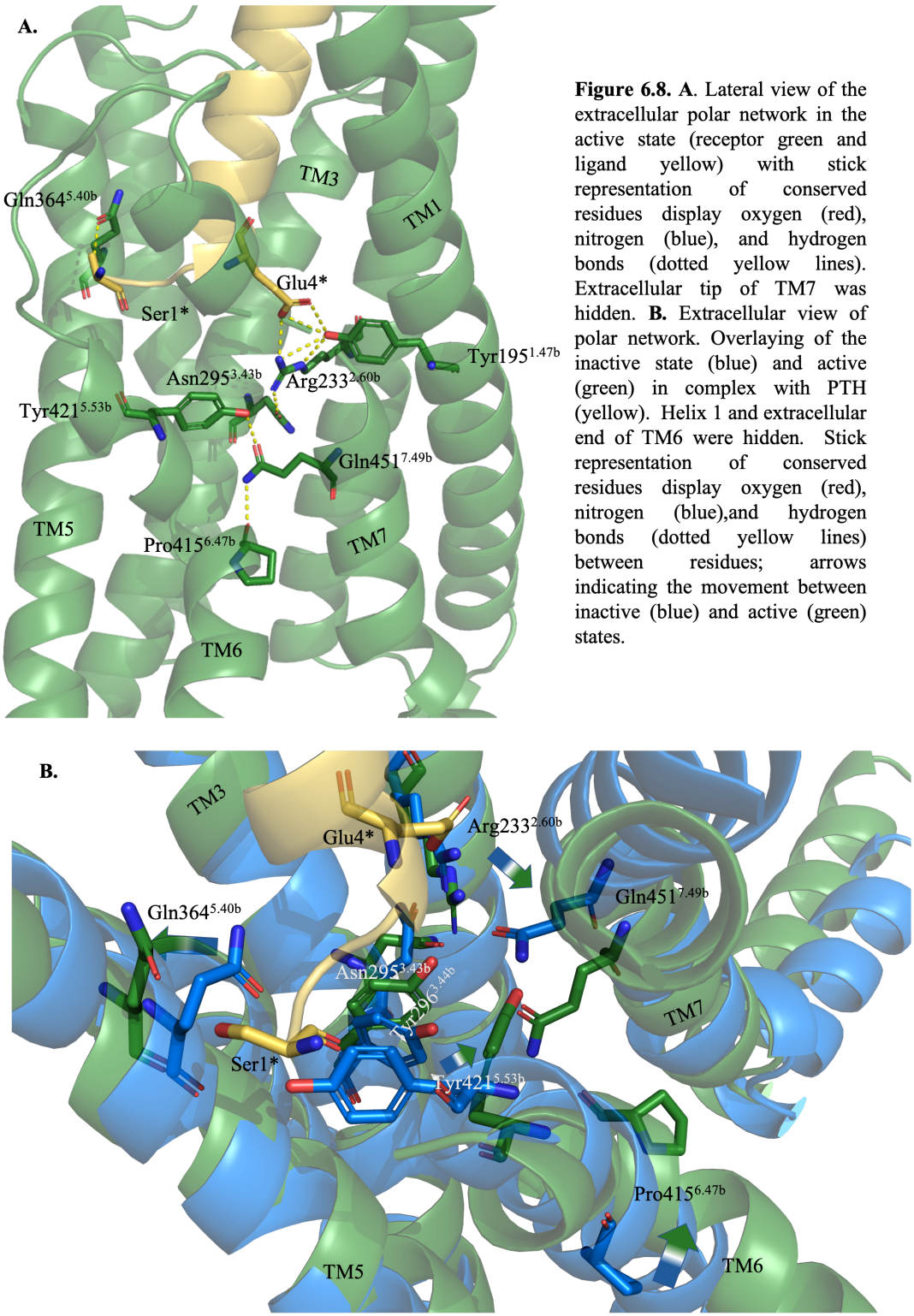


Figure 6.8. A. Lateral view of the extracellular polar network in the active state (receptor green and ligand yellow) with stick representation of conserved residues display oxygen (red), nitrogen (blue), and hydrogen bonds (dotted yellow lines). Extracellular tip of TM7 was hidden. **B.** Extracellular view of polar network. Overlaying of the inactive state (blue) and active (green) in complex with PTH (yellow). Helix 1 and extracellular end of TM6 were hidden. Stick representation of conserved residues display oxygen (red), nitrogen (blue),and hydrogen bonds (dotted yellow lines) between residues; arrows indicating the movement between inactive (blue) and active (green) states.

Table 6.4. Comparison between active and inactive state hydrogen bond occupancy		
Hydrogen bonds occupancy	Active (%)	Inactive (%)
Arg233$^{2.60b}$– Glu4*	94	0
Arg233$^{2.60b}$– Asn295$^{3.43b}$	93	96
Arg233$^{2.60b}$- Gln451$^{7.49b}$	0	38
Arg233$^{2.60b}$- Tyr296$^{3.44b}$	27	0
Gln451$^{7.49b}$–Tyr296$^{3.44b}$	0	11
Gln451$^{7.49b}$ – Tyr421$^{5.53b}$	85	0
Gln451$^{7.49b}$– Pro415$^{6.47b}$ main chain	71	0
Gln364$^{5.40b}$- Tyr421$^{5.53b}$	0	63
Gln364$^{5.40b}$– Ser1*	53	0
Hydrogen bond occupancy between side chains, unless otherwise marked.		

There are differences in the sidechains of hydrophobic residues in the active and inactive state (Figure 6.9 and Table 6.5). Although not seen during the simulation, activation involves the binding of the ligand in the orthosteric binding pocket at the extracellular end while creating an intracellular cavity to accommodate the G protein. In order to "create" two cavities, hydrophobic residues at the core of the receptor work as a gear by changing sidechains rotamers to translate changes from the extracellular end into the intracellular end while maintaining the core stability of the receptor. This is possible due to conserved Asn295$^{3.43b}$, Asn374$^{5.50b}$ and Gln451$^{7.49b}$ which translate the signal from the extracellular to the intracellular side in what appear to be checkpoints/switches for activation; while the bulkier sidechains of hydrophobic residues contribute to the changes and stability of the receptor while at the same time of rotating, they remove steric restraints in the middle region of the receptor and PxxG motif allowing the kink formation. The expected order of events would be binding of ligand E4* to Arg233$^{2.60b}$ and movement towards the intracellular side (Figure 6.8), these residues displace Asn295$^{3.43b}$ towards the intracellular, reducing the space at the core of the receptor and forcing Ile299$^{3.47b}$ rotamer change, which removes spatial restraints and allows Phe417$^{6.49b}$ - part of the PxxG motif, rotation. Mutagenesis data show the importance of conserved Asn295$^{3.43b}$; mutation of its homologue residue in GLP-1R decreases receptor activation (Wootten, Simms, et al. 2013). Together Asn295$^{3.43b}$ and Gln451$^{7.49b}$ contribute to TM6 kink formation by restraining the space at the centre of the receptor to favour rotamer changes at the side chain of

hydrophobic residues and the formation of a hydrogen bond between Gln451$^{7.49b}$ and Pro415$^{6.47b}$. Mutagenesis data show that simultaneously mutation of Asn240$^{3.43b}$ and Gln394$^{7.49b}$, homologue residues in GLP-1R, has a bigger effect in activation than the solely mutation of Asn240$^{3.43b}$ (Wootten, Simms, et al. 2013).

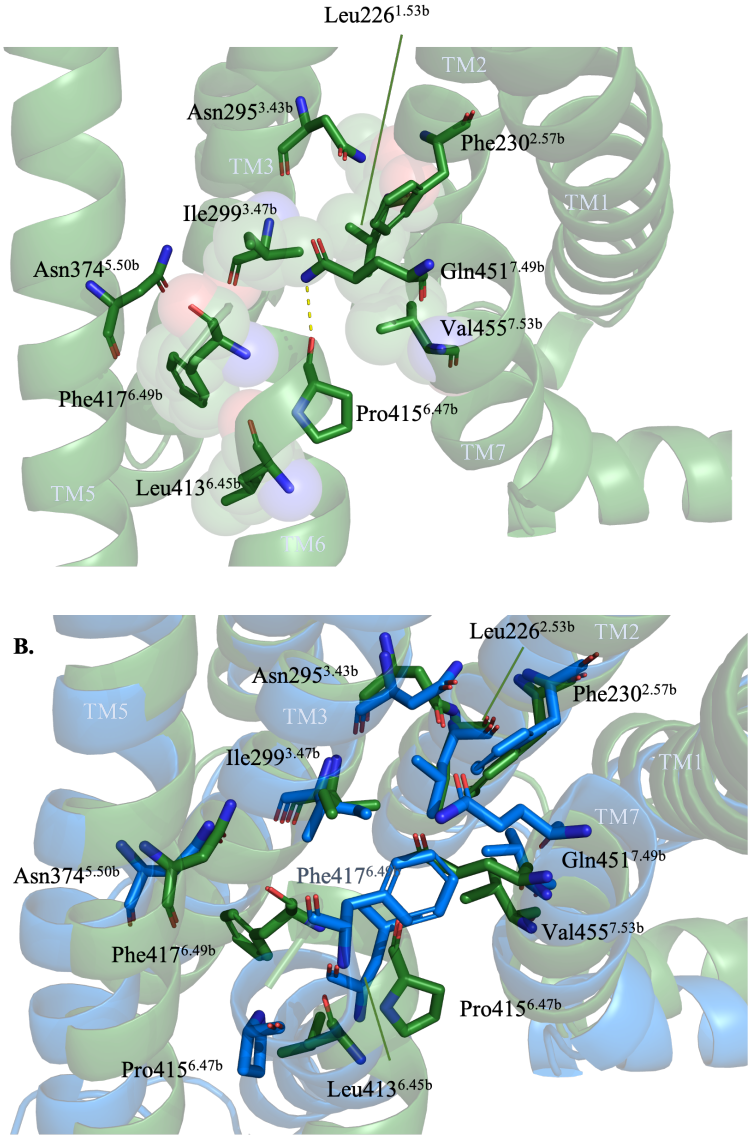


Figure 6.9. A. Extracellular view of residues in the core of PTH_1 receptor in the active state. Hydrophobic residues represented as stick/spheres and residues connecting extracellular polar with hydrophobic network represented as only sticks. Stick representation display oxygen (red), nitrogen (blue) and hydrogen bond as dashed yellow line. **B.** Extracellular view of residues in the core of the receptor overlaying active (green) and inactive receptor showing changes in the sidechain position. Stick representation display oxygen (red), nitrogen (blue).

Table 6.5. Comparison between active and inactive state hydrogen bond occupancy		
Hydrogen bonds occupancy	Active (%)	Inactive (%)
Asn295$^{3.43b}$- Arg233$^{2.60b}$	93	96
Asn295$^{3.43b}$– Ser229$^{2.56b}$	8	40
Asn295$^{3.43b}$ main chain – Ile299$^{3.47}$ main chain	85	91
Asn295$^{3.43b}$ main chain – Phe291$^{3.39b}$ main chain	54	66
Ans374$^{5.50b}$ main chain – Phe378$^{5.54b}$ main chain	95	95
Ans374$^{5.50b}$ main chain – Ser370$^{5.46b}$ main chain	74	91
Ans374$^{5.50b}$– Met414$^{6.46b}$	0	18
Ans374$^{5.50b}$– Ile299$^{3.47b}$	18	0
Gln451$^{7.49b}$ main chain – Val455$^{7.53b}$ main chain	98	90
Gln451$^{7.49b}$ – Tyr421$^{5.53b}$	85	0
Gln451$^{7.49b}$ – Pro415$^{6.47b}$ main chain	71	0
Gln451$^{7.49b}$ main chain – Asn448$^{7.44b}$ main chain	0	79
Gln451$^{7.49b}$ – Arg233$^{2.60b}$	0	38
Gln451$^{7.49b}$ – Tyr296$^{3.44b}$	0	11

At the intracellular end of the active receptor disruption the polar interaction involving His223$^{2.50b}$, Glu302$^{3.50b}$, Thr410$^{6.42b}$, Tyr459$^{7.57b}$, Arg219$^{2.46b}$ and Ser409$^{6.41b}$ allows TM6 outward movement and further interaction with the intracellular G protein (Figure 6.7, 6.10; Table 3).

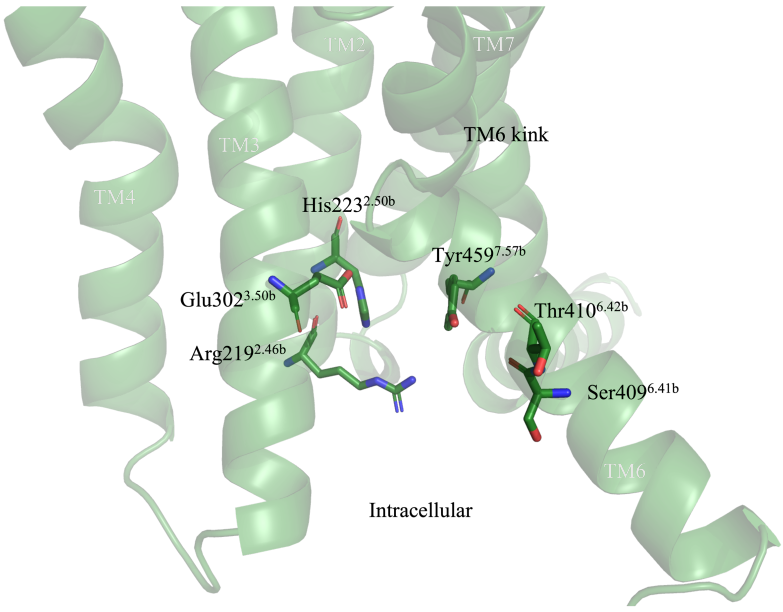


Figure 6.10. Lateral view of the intracellular ionic lock in the active state of PTH_1 receptor. Stick representation display oxygen (red) and nitrogen (blue). TM5 and G protein were hidden.

Comparison of the inactive with the active state show that within changes in the extracellular polar network during activation and the bulky movement of hydrophobic residues reflected in rotamer changes that translate conformational changes from the extracellular to the intracellular side at the same time as they stabilise/lock the position of polar residues at the extracellular end in the active conformation favouring G protein binding. The shift of TM7 over TM6 and rearrangement of residues in helix 7, plus rotation of hydrophobic residues side chains towards the intracellular side, specially Val455$^{7.53b}$, disrupts the ionic lock formed by Thr410$^{6.42b}$ -Tyr459$^{7.57b}$ - Glu302$^{3.50b}$. As the now free Tyr459$^{7.57b}$ forms a hydrogen bond with Val412$^{6.44}$ from helix 6, maintaining the correct size of the cavity for G protein binding or determining receptor specificity towards intracellular protein. (Figure 6.11).

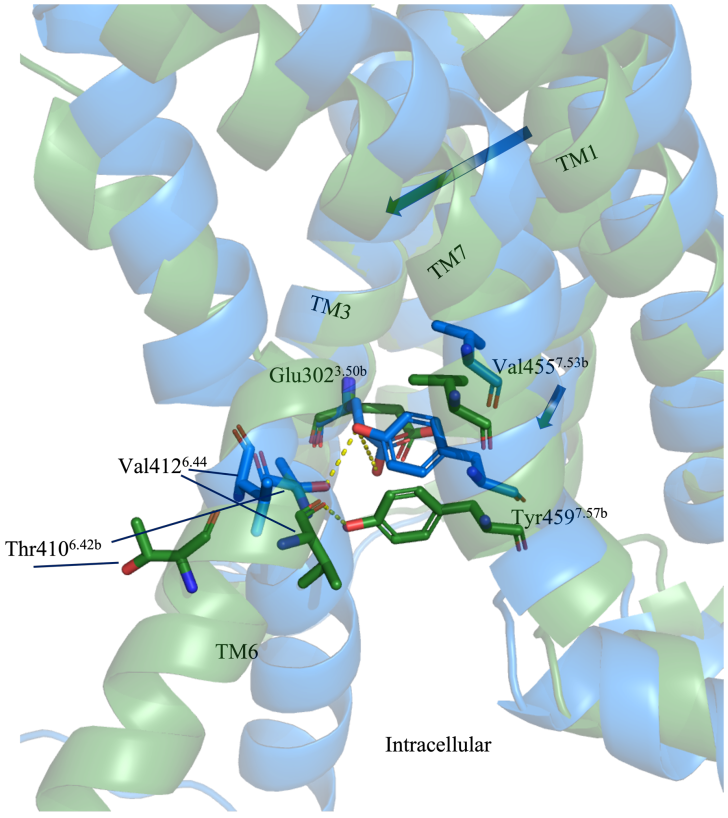


Figure 6.11. Overlapping of active (green) and inactive (blue) PTH_1 receptor. Stick representation display oxygen (red), nitrogen (blue) and hydrogen bond as dashed yellow line. Arrows show the movement from inactive (blue) to active (green).
In the image residues Glu302$^{3.50}$, Thr410$^{6.42b}$ and Tyr459$^{7.57b}$ belonging to the ionic lock interact between them via hydrogen bonds (yellow dashed line) in the inactive conformation (blue). Within activation, there are changes in hydrophobic residues at the core, one of which is Val455$^{7.53b}$, which its position in the active state (green) shifts and locks Tyr459$^{7.57b}$, in the active position stabilising TM6 by forming a hydrogen bond with Val412$^{6.44b}$ in helix 6.

In the same way, Glu4*- Arg233$^{2.60b}$ interaction in the active state (Table 6.4) and comparison between inactive and active structures (Figure 6.8 and 6.12) shows the shift of Arg233$^{2.60b}$ in the direction of the intracellular side, while maintaining Arg233$^{2.60b}$- Asn295$^{3.43b}$ hydrogen bond (Table 6.4). Comparison of inactive and active models show Asn295$^{3.43b}$ rotamer change (Figure 6.9 and 6.12). The difference in Asn295$^{3.43b}$ jeopardize the space at the core of the receptor and in order to avoid spatial restraints the change of Asn295$^{3.43b}$ acts as a switch translating changes from the extracellular polar network to the core of the receptor, where it forces Ile299$^{3.47b}$ rotamer change. In the active state, the rotamer adopted by Ile299$^{3.47b}$ removes spatial restrains from the sidechain creating a small cavity in the core allowing residues in the middle of TM6 movement.

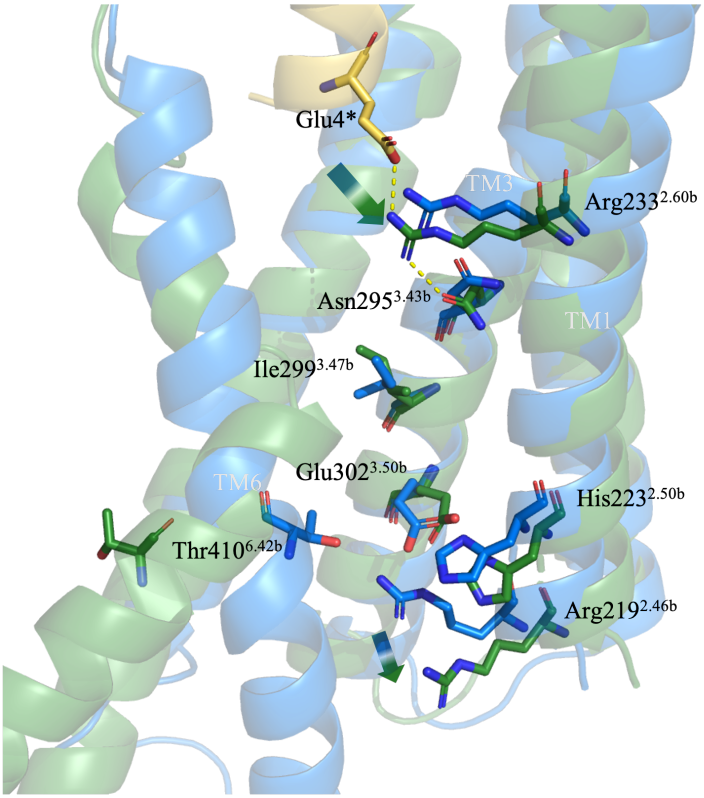


Figure 6.12. Lateral view showing the overlaying of active (green) and inactive (blue) PTH_1 receptor and PTH (yellow) from the active model simulation. Stick representation display oxygen (red), nitrogen (blue) and hydrogen bond as dashed yellow line. Arrows show the movement from inactive (blue) to active (green). Ligand binding and interaction of Glu4*-Arg233$^{2.60b}$ trigger further changes: as the orthosteric pocket shrinks, this forces Asn295$^{3.43b}$ rotamer change affecting Ile299 rotamer, shifting in the process Glu302$^{3.50b}$ and contributing to the disruption of the intracellular network allowing further interaction of intracellular residues with G protein.

6.3.2. Targeted Molecular Dynamics

6.3.2.1. Trajectory analysis

Starting inactive and the active 'target' structure were obtained from the inactive and active MDFF simulation respectively. The inactive complex was guided towards the active state by applying restraints in Cα atoms for over 150 ns. To facilitate the analysis of the trajectory in a chronological order, the resulting trajectory was divided in base of RMSD values resulting into eight segments (A to H) of varied length; segments with stable RMSD values were separated from segments with increase or changes in RMSD. Then each segment was analysed. (Figure 6.13).

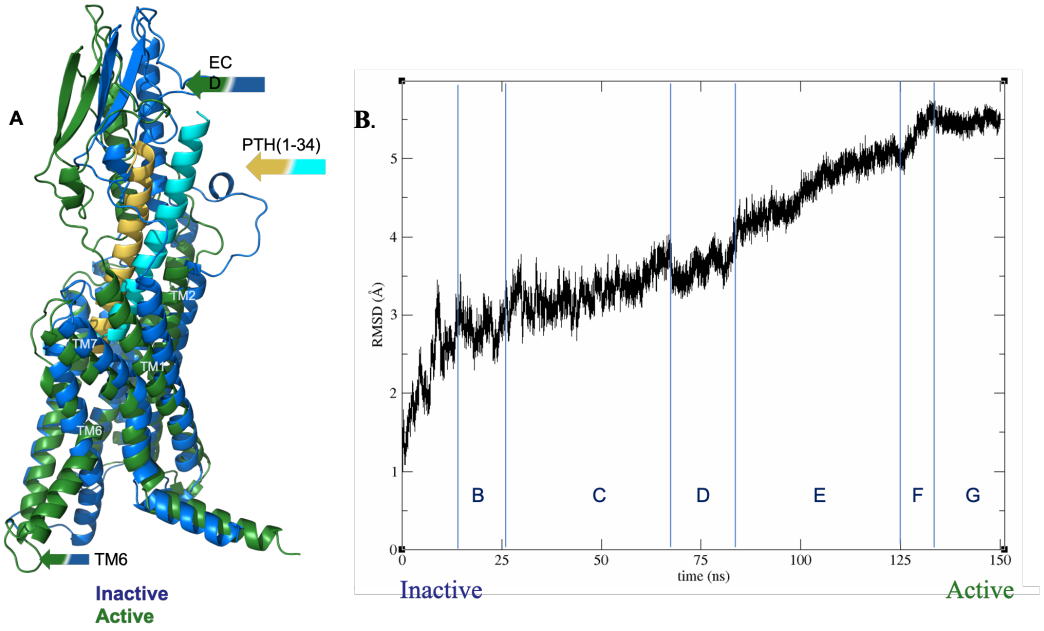


Figure 6.13. Overlaying of active (green) and inactive (navy) PTH_1R and PTH in active state (yellow) and inactive state (cyan). During the trajectory the ligand accommodates into the binding pocket, TM1 moves over TM7 and TM7 towards TM6. **B.** RMSD of PTH1 receptor over 150 ns of targeted MDs. The trajectory was divided into segments (A-G - blue lines) according to RMSD values. Each segment was later analysed.

6.3.3. Results

Simulation of the receptor activation using targeted MD showed conformational changes as expected during activation.

At the extracellular side Asp137^{ECD} guides the ligand into the binding pocket via an interaction with Arg20* through first a salt bridge (maintained until segment C) and then by a hydrogen bond until the end of the simulation. This interaction is essential for full affinity (Weaver, Wigglesworth, and Donnelly 2014) and showing that the ligand remained bound into the receptor during activation as movement of the N-terminus end of the ligand towards the binding pocket is seen during the simulation. This interaction resembles GLP-1R K26*-Glu138^{ECD} (Chapter 5) which also anchors the ligand's middle region directing it to the binding pocket and facilitating the second step in class B ligand binding.

Binding of the ligand into the binding pocket rearranges the extracellular polar clusters Arg233$^{2.60b}$, Tyr195$^{1.47b}$, Asn295$^{3.43b}$, Gln451$^{7.49b}$; Gln364$^{5.40b}$- Tyr421$^{6.53}$, seen in the inactive conformation. As the ligand moves into the binding pocket via TM1-TM2, hydrophobic residues Leu11* and Leu15* separate TM1 and TM2 extracellular ends as Glu4* temporarily disrupts the Tyr191$^{1.43}$ - Lys240$^{2.67b}$ interaction (segment B). Then, as the ligand goes deeper into the binding pocket, Glu4* forms hydrogen bonds with Arg233$^{2.60b}$, Tyr191$^{1.43b}$, Tyr195$^{1.47b}$. These interactions resemble the disruption of Tyr148$^{1.43b}$ – Lys197$^{2.67b}$ in GLP-1R and Glu9* Arg190$^{2.60b}$ interaction suggesting Class B ligand enters via the TM1-TM2 interface and rely on positive charged TM2 residue Arg$^{2.60b}$ to anchor the ligand into the binding pocket, facilitating the extracellular polar rearrangement and positioning the N-terminus towards TM5-TM6 interface (Figure 6.14).

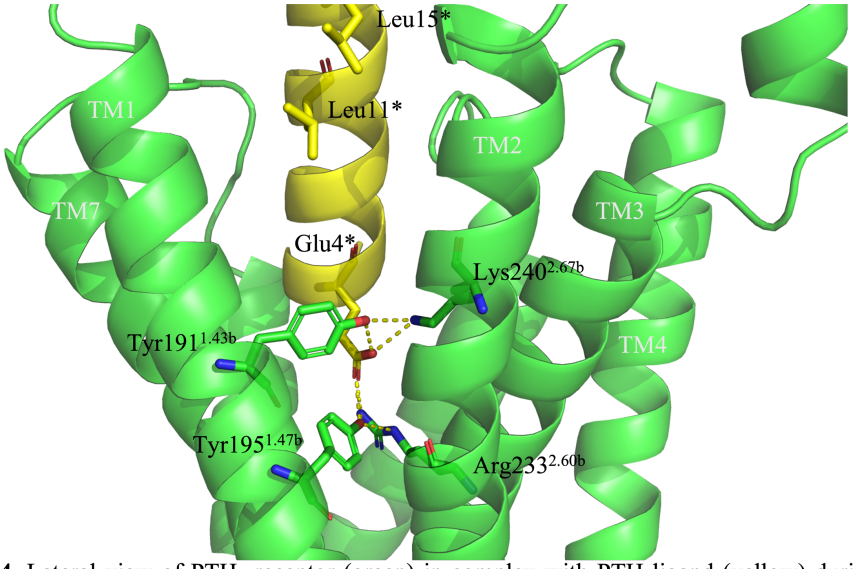


Figure 6.14. Lateral view of PTH_1 receptor (green) in complex with PTH ligand (yellow) during segment A. Stick representation display oxygen (red), nitrogen (blue) and hydrogen bond as dashed yellow line. Movement of the ligand via the TM1 - TM2 interface towards the binding pocket. In the trajectory Glu4* interacts with residues of the extracellular polar network as hydrophobic Leu11* and Leu15* facilitates the opening of the TM1 - TM2 interface. The mass effect of the ligand shifts TM1 towards TM7.

Following the movement of the ligand into the binding pocket, S3* interacts with Glu444$^{7.42b}$; similar to homologous residues Glu387$^{7.42b}$ and Ala8^{GLP-1} in GLP-1R. Mutagenesis of Glu387$^{7.42b}$ impairs GLP-1R efficacy (Dods and Donnelly 2015). This suggest that Class B ligand interaction with the extracellular portion of TM7 shift TM7 towards TM6, contributing to TM6 rearrangement and positioning the ligand's N-terminal towards the TM5-TM6 interface. In the inactive state, TM5-TM6 extracellular end remain parallel via Gln364$^{5.40b}$ and Tyr421$^{6.53}$interaction. Once the ligand's N-terminus is facing TM5-TM6, S1* joins Gln364$^{5.40b}$ - Tyr421$^{6.53}$ interaction (Figure 6.15) during the first stages of targeted MD simulation (segments A to B) and remains within binding distance for the rest of the simulation. Residues Gln364$^{5.40b}$ - Tyr421$^{6.53}$ and S1* resemble GLP-1R Arg310$^{5.40b}$ and Glu364$^{6.53b}$ interaction with His7^{GLP-1}. However, in GLP-1R Arg310$^{5.40b}$ - Glu364$^{6.53b}$ has bigger role in activation as mutation of these residues impairs GLP-1R efficacy (Dods and Donnelly 2015; Wootten, Reynolds, Koole, et al. 2016).

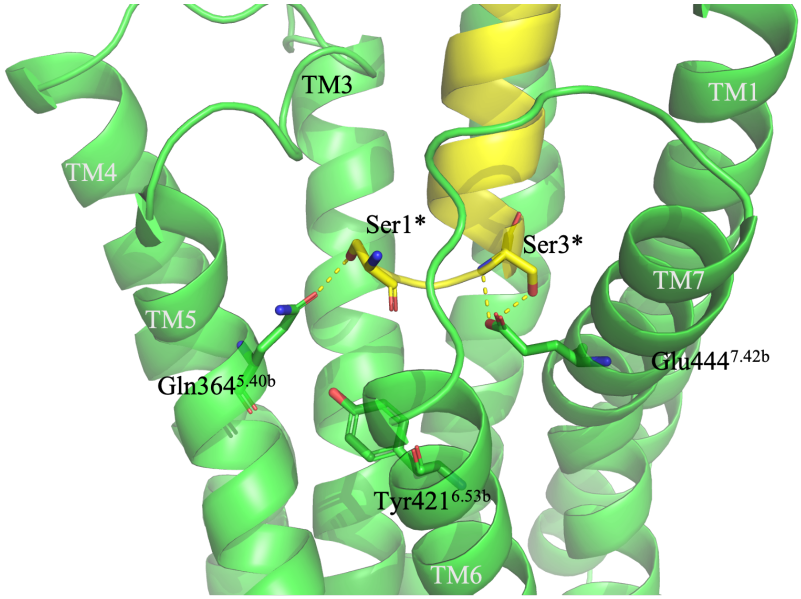


Figure 6.15. Lateral view of PTH_1 receptor (green) in complex with PTH ligand (yellow). Stick representation display oxygen (red), nitrogen (blue) and hydrogen bond as dashed yellow line. Ser1* disrupts Gln364$^{5.40b}$ – Tyr421$^{6.53b}$ interaction, and Ser3* uses Glu444$^{7.42b}$ as an anchor facilitating the movement of the ligand into the binding pocket and TM7 movement towards TM6.

Comparison of PTH_1 and PTH_2 receptors show that the movement of the ligand's N-terminus into PTH_1 receptor binding pocket is possible in the absence of physical restraints. PTH_1 receptor residues Leu289$^{3.37b}$ and Ile363$^{5.39b}$ sidechains are able to change their rotamers allowing Ile5$^{PTH(1-34)}$ or His5^{PTHrP} to accommodate, in contrast Ile244$^{3.37b}$ and Tyr318$^{5.39b}$ in PTH_2 receptor 'block' the movement of the ligand into the binding pocket acting as a selectivity mechanism during the second-step (Weaver et al. 2017; Figure 6.16).

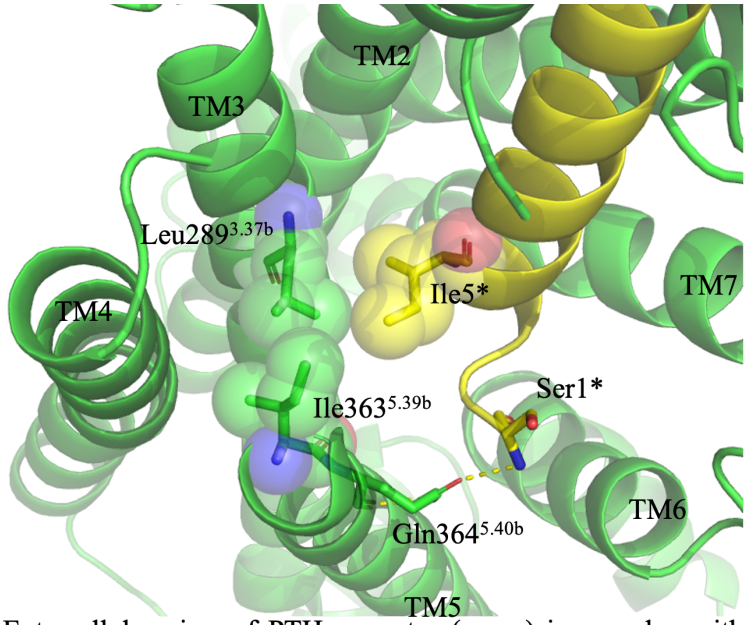


Figure 6.16. Extracellular view of PTH_1 receptor (green) in complex with PTH ligand (yellow). Stick representation display oxygen (red), nitrogen (blue) and hydrogen bond as dashed yellow line. Ser1* interaction with Gln364$^{5.40b}$ is possible as Ile5* is able to move into the binding pocket.

Binding of the ligand, movement of TM7 over TM6 and disruption of TM5-TM6 facilitate rotamer changes at the core of the receptor: i) the mass effect of ligand's N-terminus shifts the now free Tyr421$^{6.53}$towards the core of the receptor ii) this shifts Tyr296$^{3.44b}$ towards the core iii) triggering Ile299$^{3.47b}$ rotamer change, which removes restraints from the core of the receptor and iv) pushes Phe417$^{6.49b}$ towards the lipid bilayer facilitating TM6 outward movement.

Interaction between the ligand's Glu4* and Arg233$^{2.60b}$ anchors the ligand to the receptor, and as the ligands moves deeper into the binding pocket Arg233$^{2.60b}$ sidechain subtly moves into the core driving Asn295$^{3.43b}$ sidechain towards TM3 and removing any possible constraints from the centre of the receptor. Unlike MDFF, the resulting active state and interactions stabilising the kink are different. For example, after TM6 rotation and outward movement, Gln451$^{7.49b}$ forms a hydrogen bond with Arg233$^{2.60b}$ stabilising the core of the receptor (Figure 6.17; Table 6.6).

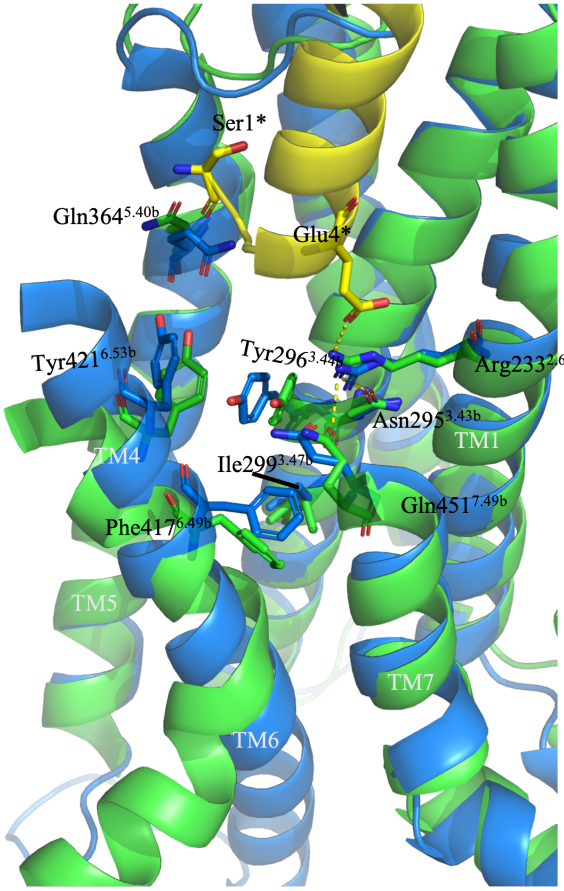


Figure 6.17. lateral view overlapping inactive (blue) and active (green) PTH_1 receptor in complex with PTH ligand (yellow) in the active state. TM1 and ECL3 were hidden.
Ser1* and the N-terminal end of the ligand push Tyr421$^{6.53b}$ towards the core, which at the same time push Tyr296$^{3.44b}$ and Phe417$^{6.49b}$ allowing TM6 outward movement. Interaction between Arg233$^{2.60b}$ - Asn295$^{3.43b}$ and Gln451$^{7.49b}$ stabilise the core of the receptor while TM6 rotation and outward movement.
Stick representation display oxygen (red), nitrogen (blue) and hydrogen bond as dashed yellow line.

Table 6.6. Percentage of hydrogen bond occupancy across targeted MD of PTH_1 receptor

Hydrogen bond	Segment							
	A	B	C	D	E	F	G	H
Arg20* - Asp137^{ECD}	99%	98%	90%	96%	98%	97%	94%	94%
Glu4* - Arg233$^{2.60b}$	-	90%	93%	88%	87%	80%	81%	80%
Glu4* - Tyr191$^{1.43b}$	95%	89%	98%	93%	92%	99%	99%	99%
Glu4* - Tyr195$^{1.47b}$	95%	73%	98%	91%	91%	98%	98%	96%
Gln451$^{7.49b}$ - Arg233$^{2.60b}$	-	9%	6%	7%	40%	48%	63%	40%
Thr410$^{6.42b}$- Glu302$^{3.50b}$	-	84%	66%	-	23%	17%	-	-
Glu302$^{3.50b}$ – Tyr459$^{7.57b}$	25%	90%	78%	90%	86%	87%	88%	-

Table 6.7. Percentage of chis angles throughout targeted MD PTH1 receptor

Residue	chi		A	B	C	D	E	F	G	H
His223$^{2.50b}$	x1	g+	0	0	0	0	0	0	0	0
		g-	100	100	100	100	100	100	100	100
		t	0	0	0	0	0	0	0	0
	x2	g+	19	17	17	3	3	5	0	0
		g-	81	83	83	97	97	81	0	0
		t	0	0	0	0	0	14	100	100
Leu226$^{2.53b}$	x1	g+	0	0	0	0	0	0	0	0
		g-	0	0	0	0	0	0	2	0
		t	100	100	100	100	100	100	98	100
	x2	g+	98	100	98	97	99	100	98	100
		g-	0	0	0	0	0	0	0	0
		t	2	0	2	3	1	0	2	0
Phe230$^{2.57b}$	x1	g+	0	0	0	0	0	0	0	0
		g-	43	64	72	63	100	100	100	100
		t	57	36	28	37	0	0	0	0
	x2	g+	0	0	0	0	0	0	0	0
		g-	100	100	100	100	100	100	100	100
		t	0	0	0	0	0	0	0	0
Arg233$^{2.60b}$	x1	g+	0	0	0	0	3	0	0	1
		g-	0	0	0	0	0	0	5	0
		t	100	100	100	100	97	62	95	99
	x2	g+	1	10	29	0	0	0	0	0
		g-	0	0	0	0	0	0	0	0
		t	99	90	71	100	100	62	100	100
	x3	g+	0	0	0	0	0	0	5	0
		g-	0	0	0	0	3	0	0	1
		t	100	100	100	100	97	62	95	99
	x4	g+	0	0	0	0	0	0	0	0
		g-	1	10	29	0	0	0	0	0
		t	99	90	71	100	100	62	100	100
Asn295$^{3.43b}$	x1	g+	0	0	0	0	0	0	0	0
		g-	100	100	84	11	0	0	0	0
		t	0	0	16	89	100	100	100	100
	x2	g+	1	0	0	0	0	0	0	0
		g-	0	0	2	12	16	22	20	22
		t	99	100	98	88	84	78	80	78
Ile299$^{3.47b}$	x1	g+	0	2	0	0	0	0	0	0
		g-	100	92	100	100	100	100	100	100
		t	0	6	0	0	0	0	0	0
	x2	g+	0	0	0	0	0	0	0	0
		g-	0	0	0	0	0	0	0	0
		t	0	0	0	0	0	0	0	0

Table 6.7. Percentage of chis angles throughout targeted MD PTH1 receptor

Residue	chi		A	B	C	D	E	F	G	H
Glu302$^{3.50b}$	x1	g+	1	0	0	0	0	0	0	0
		g-	2	31	58	0	41	82	0	1
		t	97	69	42	100	59	18	100	99
	x2	g+	0	14	0	2	0	0	0	0
		g-	3	26	58	0	41	100	100	100
		t	97	60	42	98	59	0	0	0
	x3	g+	44	28	55	2	5	8	13	26
		g-	35	50	15	62	53	0	0	0
		t	21	22	31	36	41	92	87	74
Gln364$^{5.40b}$	x1	g+	0	0	0	0	0	0	0	0
		g-	99	100	99	10	6	0	0	0
		t	1	0	1	90	94	100	100	100
	x2	g+	2	1	3	90	87	100	100	100
		g-	29	83	38	0	0	0	0	0
		t	69	16	59	9	13	0	0	0
	x3	g+	5	25	31	88	82	80	85	80
		g-	26	65	52	9	8	1	0	0
		t	69	10	17	3	10	19	15	20
Asn374$^{5.50b}$	x1	g+	0	0	0	0	0	0	0	0
		g-	100	71	2	0	0	0	0	0
		t	0	29	98	100	100	100	100	100
	x2	g+	4	29	58	25	77	70	344	20
		g-	88	64	20	42	13	20	279	38
		t	7	7	22	33	10	11	377	42
Thr410$^{6.42b}$	x1	g+	0	0	16	0	0	1	2	23
		g-	100	99	82	99	100	43	0	16
		t	0	1	2	1	0	55	98	61
Leu413$^{6.45b}$	x1	g+	0	0	0	0	0	0	0	0
		g-	67	62	100	100	98	100	100	20
		t	33	38	0	0	2	0	0	80
	x2	g+	23	31	0	0	3	0	0	79
		g-	1	0	0	0	0	0	0	0
		t	77	69	100	100	97	100	100	21
Pro415$^{6.47b}$	x1	g+	85	93	60	31	15	7	47	50
		g-	15	7	40	69	85	93	53	50
		t	0	0	0	0	0	0	0	0
	x2	g+	14	6	40	70	86	93	52	49
		g-	86	94	60	30	14	7	48	51
		t	0	0	0	0	0	0	0	0

Table 6.7. Percentage of chis angles throughout targeted MD PTH1 receptor

Residue	chi		A	B	C	D	E	F	G	H
Phe417$^{6.49b}$	x1	g+	0	1	0	0	0	0	0	0
		g-	100	68	100	100	100	100	100	100
		t	0	31	0	0	0	0	0	0
	x2	g+	19	36	11	0	0	7	15	29
		g-	25	42	18	0	0	12	59	62
		t	56	22	71	100	100	81	26	9
Tyr421$^{6.53b}$	x1	g+	0	0	0	0	0	0	0	0
		g-	0	0	0	0	0	0	0	0
		t	100	100	100	100	100	100	100	100
	x2	g+	0	0	0	0	0	84	59	0
		g-	94	99	94	95	98	16	34	93
		t	6	1	6	5	2	0	7	7
Gln451$^{7.49b}$	x1	g+	7	100	38	14	59	27	0	39
		g-	5	0	9	13	6	9	33	9
		t	88	0	53	72	35	64	67	51
	x2	g+	0	1	1	0	0	0	0	0
		g-	0	0	0	0	0	0	0	0
		t	100	99	99	100	100	100	100	100
	x3	g+	70	0	38	64	33	61	77	49
		g-	10	62	25	13	40	20	18	22
		t	20	38	38	23	27	18	5	29
Val455$^{7.53b}$	x1	g+	0	1	0	0	2	1	0	1
		g-	0	0	4	58	0	0	100	0
		t	100	99	94	42	98	99	0	99
Tyr459$^{7.57b}$	x1	g+	0	0	0	0	0	0	0	0
		g-	100	100	99	100	100	100	100	100
		t	0	0	1	0	0	0	0	0
	x2	g+	2	7	36	11	11	3	12	16
		g-	98	93	41	89	73	1	67	51
		t	0	1	23	0	16	96	21	32

At the intracellular side, the ionic lock is broken as result of TM6 outward movement, specially by direct action of Met414$^{6.46b}$ which disrupts Thr410$^{6.42b}$-Glu302$^{3.50b}$, and Glu302$^{3.50b}$ – Tyr459$^{7.57b}$ interaction resulting in an intracellular cavity for G protein binding.

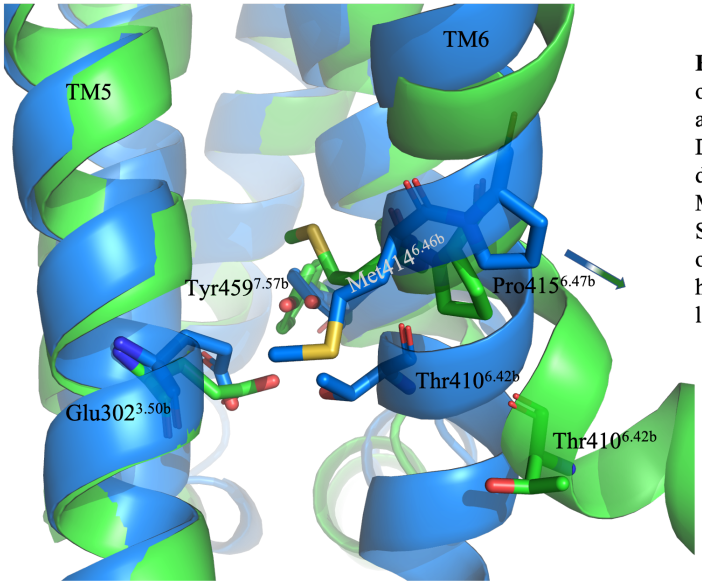


Figure 6.18. Lateral view overlapping inactive (blue) and active (green) PTH_1 receptor. Disruption of the ionic lock as direct consequence of Met414$^{6.46b}$ and TM6 rotation. Stick representation display oxygen (red), nitrogen (blue) and hydrogen bond as dashed yellow line.

6.4. Discussion

Recently full-length PTH_1 receptor structures became available showing the interactions between the peptide ligand and the receptor (Ehrenmann et al. 2018; Zhao et al. 2019). Despite differences in size, peptides used in the resolution of the receptors (ePTH, LA-PTH) and PTH used in the simulations, the peptide- PTH_1 receptor appear in a similar orientation as the peptide in GLP-1R in agreement with the two-steps activation model. Comparison between PTH_1 receptor inactive and active state and with GLP-1R show common activation mechanisms in Class B GPCRs. For instance, the ionic interaction between ligand Arg20* and extracellular residue Asp137^{ECD}, essential for full affinity (Weaver, Wigglesworth, and Donnelly 2014) look like GLP-1R K26*-Glu138^{ECD} (Chapter 5) and suggesting similar orientation and activation mechanisms, even when activated by peptide agonists (Ehrenmann et al. 2018).

Binding of the ligand into the orthosteric binding pocket involved the interaction between Glu4* and charged Arg233$^{2.60b}$. This interaction, characteristic in Class B receptor-ligand interactions resembling Glu9* - Arg190$^{2.60b}$ in GLP-1 GLP-1R complex, appear to improve the stability of the ligand and help to accommodate the ligand's N-terminus towards TM6 (Zhao et al. 2019).

Comparison of PTH_1 and PTH_2 receptors show residues Leu289$^{3.37b}$ and Ile363$^{5.39b}$ sidechains in PTH_1 receptor are able to change their rotamers. Therefore, able to accommodate Ile5$^{PTH(1-34)}$ or His5^{PTHrP}. In contrast, Ile244$^{3.37b}$ and Tyr318$^{5.39b}$ in PTH_2 receptor are unable to perform such rotamer changes, blocking the binding of the N-terminus and activation (Weaver et al. 2017).

The simulations showed interaction between the N-terminus of the ligand with residues in TM6, allowing the extracellular outward movement of helix 6, as well in PTH_1 receptor the unwinding TM6 extracellular end trigger by the N-terminus (Zhao et al. 2019) However, after TM6 outward movement there are differences in the contacts stabilising the kink. For instance, in the active PTH_1 receptor 6NBF the kink is stabilised by His420$^{6.52b}$ – Gln451$^{6.47b}$ (Zhao et al. 2019), but in the targeted MD Gln451$^{7.49b}$ forms a hydrogen bond with Arg233$^{2.60b}$ and TM6 outward movement appears to be stabilised by Phe417$^{6.49b}$. In the GLP-1R targeted MD TM6 outward movement is triggered by the shift of Leu360$^{6.49b}$ towards the lipid bilayer. This would suggest different activation states, since this model was built using an agonist, would be that the extension of TM6 outward movement determines the intracellular selectivity. Like in glucagon receptor, in which the extension of TM6 determines the selectivity towards the intracellular protein (Qiao et al. 2020).

6.5. Conclusion

The study of PTH_1 receptor show conserved residues and a common activation mechanism in Class B GPCRs. There are differences between the resulting final states between MDFF and targeted MD. However, overall the movement of the extracellular end of TM7 over TM6, a highly conserved extracellular polar network, the role of hydrophobic residues to stabilise inactive and active conformation and the role of asparagine residues and glutamine as switches translating changes in the

extracellular region to the intracellular polar network and residues Leu289$^{3.37b}$ and Ile363$^{5.39b}$ at the extracellular end of TM3 and TM5 physically regulating selectivity between ligands after first step from Class B ligand binding.

Chapter 7: General Discussion

GPCRs mediate a wide and important range of physiological functions making them therapeutic targets. Although a relatively small subset of all GPCRs, Class B GPCRs are involved in validated therapeutic approaches to treating diseases such T2DM, obesity, migraine and osteoporosis, as well having potential in neurodegenerative diseases, cardiovascular disease, NASH and mood disorders like anxiety and depression. However, their peptide nature hinders the development of orally-administrated peptide agonists; until recently when an orally available semaglutide (Rybelsus®) became the first approved GLP-1R agonist for oral use (Bucheit et al. 2020).

Due to Class B biological relevance and therapeutic potential, in the last years the number of solved receptor structures have dramatically increased. Receptor mutations and ligands have been used to stabilise Class B receptors in diverse states and overcome the challenges derived from the native membrane environment to obtain high quality structural data and provide valuable receptor contacts and ligand-receptor interaction useful for understanding Class B activation. Knowledge and understanding of Class B receptor ligand binding and agonist-mediated activation may allow the development of synthetic or peptide agonists to target specific signalling pathways and treat a wide range of endocrinological diseases with greater efficacy, reduced side effects and more convenient administration and storage.

The aim of this PhD project was to build a reliable Class B GPCR-agonist complexes in the active and inactive state and identify key interactions characterising each state. After defining these interactions, the second aim was to identify the changes that lead to activation. Class B share structural characteristics and many ligands also for similar helical conformations, hence there is a possible shared activation mechanism. Hence the project compared two different Class B GPCRs, GLP-1R and PTH_1.

7.1 GLP-1R and PTH_1 receptor

The NTD has a critical role in peptide binding and receptor activation, as this dynamic structure facilitates and guides the movement of the peptide ligand into the orthosteric binding pocket, agreeing with the two-domain-binding mechanism. Indeed, inactive full-length receptors in the absecene of the peptide ligand, (PDB 5XF1; Zhang, Qiao, et al. 2017) are incompatible with the two-steps activation model. Differences in the position of the NTD relative to the TMD, due to the dynamism of the NTD, were found suggesting different states during binding. Still NTD-ligand interactions were found during the simulations. In GLP-1 and PTH_1 receptors models, a persistent interaction seen first as a salt bridge and continued as a hydrogen bond between ligand's residues, Lys26* in GLP-1 and Arg20* in PTH, and charged residues from the NTD, $Glu138^{1.36b}$ in GLP-1R and $Asp137^{NTD}$ in PTH_1 receptor was seen in agreement with the two-steps activation model and guiding the ligand into the binding pocket. Indeed, Arg20*- $Asp137^{NTD}$ interaction is essential for affinity (Weaver, Wigglesworth, and Donnelly 2014) and since Class B homology and shared activation mechanisms an analogous interaction would be expected for GLP-1R. It is important to remember that Class B peptide ligands are largely disordered structures in aqueous solution but adopt an α-helical conformation when binding to the receptor (Runge et al. 2008; Underwood et al. 2010; Parthier et al. 2009; Sun et al. 2007; Pioszak and Xu 2008; Grace et al. 2010; Inooka et al. 2001). This early analogous salt bridge interaction could promote the ligand to adopt an α-helical conformation while guiding the rest of the ligand towards the orthosteric binding pocket.

The two-steps model suggests that once the ligand's C-terminus has bound to the receptor NTD, the N-terminus of the ligand needs to interact with the TMD of the receptor. Available full-length Class B receptors in complex with peptide ligands (PDB code 5NX2, 5VAI, 6NBF, 6NBH, 6NBI, 5UZ7, 6X18, 5YQZ, 6M1I, 6WPW, 6PB1, 6P9X, 6P9Y, 6WZG, 6LML, 6LMK, 7CZ5; Ehrenmann et al. 2018; Jazayeri et al. 2017; Zhang, Qiao, et al. 2017; Zhang et al. 2017; Dong et al. 2020; Qiao et al. 2020; Wang et al. 2020; Ma, Shen, et al. 2020; Hilger, Kumar, Hu, Pedersen, O'Brien, Giehm, Jennings, Eskici, Inoue, Lerch, et al. 2020; Liang et al. 2020; Zhou et al. 2020) show the ligand with the N-terminus towards the orthosteric binding pocket and the C-terminus between TM1-TM2 interface; still unclear if the first interactions are with

residues at the TM5-TM6 and orthosteric binding pocket or residues in TM1-TM2. In the targeted MD simulation, early interactions were found between GLP-1 T13* - GPL-1R Tyr148$^{1.43b}$ – Lys197$^{2.67b}$ and PTH Glu4* - PTH_1 receptor Tyr191$^{1.43}$ - Lys240$^{2.67b}$ as hydrophobic residues from the ligand (Tyr19* and Leu20* in GLP-1, and Leu11* and Leu15* PTH_1 receptor) disrupt TM1-TM2 facilitating the entrance of the ligand into the binding pocket, accommodating and stabilising the ligand towards TM5-TM6. In GLP-1R, mutation of Lys197$^{2.67b}$ to alanine reduced GLP-1 affinity (Xiao, Jeng, and Wheeler 2000; Coopman et al. 2011), while hydrophobic interactions would increase the helicity of the peptide ligand. Recently a GLP-1R-GLP-1(9-36) and a PAM (LSN3160440) was able to fully activate the GLP-1R by binding of the PAM in the TM1-TM2 interface, allowing access of the ligand (Bueno et al. 2020). This shows the movement of the ligand through TM1-TM2 and the importance of the hydrophobic environment to facilitate the entrance of the ligand into the binding pocket, as well to stabilise the N-terminus of the ligand for later interaction with TM5-TM6, as a 'shorter' ligand would adopt multiple rotamers before interacting with TM5-TM6 (Bueno et al. 2020). Since peptide ligand structures are medium-dependent, hydrophobic interactions would encourage an α-helical conformation of the ligand (Andersen et al. 2002).

Following the interactions of GLP-1R and PTH_1 receptor in complex with their respectively ligand, the targeted MD simulations suggested two constant interactions with residues Arg$^{2.60b}$ and Glu$^{7.42b}$: Arg190$^{2.60b}$ - Glu9* and Glu387$^{7.42b}$ - Ala8* interaction in GLP-1R-GLP-1, and Arg133$^{2.60b}$ - Glu4*, and Glu444$^{7.42b}$ - Ser3* in PTH_1 receptor-PTH. Mutagenesis data show that removal of the positive charged residue Arg190$^{2.60b}$ in GLP-1R decreases efficacy (Wootten, Simms, et al. 2013) since arginine mutation to lysine had no effect in efficacy (Yang et al. 2016) but mutation to alanine significatively decrease GLP-1 efficacy ($\Delta\log\tau_c = 0.53$) (Wootten, Simms, et al. 2013). In the same way mutation of Glu387$^{7.42b}$ to alanine decrease GLP-1 efficacy ($\Delta\log\tau_c = 0.52$) (Dods and Donnelly 2015) but mutation to aspartic acid had no effect (Yang et al. 2016), supporting the importance of electrostatic interactions. Arg$^{2.60b}$ and Glu$^{7.42b}$ are located in TM2 and TM7 respectively, opposite to each other, interacting with the ligand to stabilise it. In addition, targeted MD shows that binding of the ligand with Arg$^{2.60b}$ and Glu$^{7.42b}$ translate further conformational changes, specially Arg$^{2.60b}$ to Asn$^{3.43b}$.

Comparison between active and inactive models shows that TM3-TM2 region becomes more stable in the active state compared to the flexible or less helical nature in the inactive state. It has been reported in the GCGR Asn238$^{3.43b}$ stabilises the inactive conformation via interactions with the extracellular polar network (Mattedi et al. 2020), therefore disruption of the extracellular polar network frees Asn238$^{3.43b}$. However, it is still unclear how changes from the extracellular polar network translate to the intracellular side of the receptor. The simulation of the PAC1 receptor shows that the Asn$^{3.43b}$ distance relative to R$^{2.60b}$ increases during activation and decreases during inactive state (Liao et al. 2021). However, the opposite happened during the simulation of GLP-1R and PTH_1 receptor, as the distance between R$^{2.60b}$ and Asn$^{3.43}$ decreased during activation and residues in the area become more packed. The packing of TM3-TM2 removes spatial restraints from TM3 residue side chains from the core towards TM2 resulting in a small cavity in the core of the receptor. As the result of TM3-TM2 packing and Asn240$^{3.43}$, the Leu244$^{3.47b}$ rotamer change shifts Leu360$^{6.49b}$ and neighbouring hydrophobic residues Leu183$^{2.53b}$, Leu356$^{6.45b}$, Leu360$^{6.49b}$, Val398$^{7.53b}$, stabilising the new conformation. The subsequent TM6 outward movement is produced as a consequence of the Leu244$^{3.47b}$ rotamer change shifting Leu360$^{6.49b}$, in addition to the low helical propensity allowed by Gly361$^{6.50b}$, part of the conserved PxxG motif (Pro358$^{6.47b}$, Leu359$^{6.48b}$, Leu360$^{6.49b}$, Gly361$^{6.50b}$). Although packing of TM3-TM2 by the ligand (Glu9*- Arg190$^{2.60b}$ - Asn240$^{3.43}$) disagrees with PAC1 receptor (Liao et al. 2021) the Leu244$^{3.47b}$ shifting of Leu360$^{6.49b}$ in GLP-1R resembles the well characterised rotamer change of Ile121$^{3.40}$ in β_2AR which shifts towards TM6 into the space occupied by Phe282$^{6.44}$.

At the opposite side, the interaction between the ligand Ala8* - GLP-1 and Ser3* PTH with Glu$^{7.42b}$ contributes to the movement of TM7 towards TM6 and facilitates the interaction between Gln394$^{7.49b}$ – and the Leu359$^{6.48b}$ main chain, that stabilises TM6 kink. The shift of TM7 towards TM6 contributes to the rearrangement of TM6 residues resulting in TM6 flexibility. Interaction between the ligand N-terminus and Class B TM5-TM6, (His7* GLP-1 or modify GLP-1R Arg310$^{5.40b}$-Glu364$^{6.53b}$ or Ser1* PTH Gln364$^{5.40b}$ – Tyr421$^{6.53b}$ PTH_1 receptor) provides flexibility to TM6 for the outward movement. Arg310$^{5.40b}$ in GLP-1R is crucial for GLP-1 but also for ExP5 to promote Gα_s signalling; Arg310$^{5.40b}$ mutation to alanine abolished cAMP accumulation (Dods and Donnelly 2015) showing the role of TM5 in GLP-1R cAMP-mediated signalling, and the role of the ligand acting as a bridge

between TM5 and TM6 allowing TM6 flexibility and outward movement but linking it to TM5 and maintaining the stability of the receptor. Indeed, in the 5NX2 structure (Jazayeri et al. 2017) the peptide agonist shift TM7 in direction to TM6, allowing the outward movement of TM6 creating an intracellular cavity but fails to interact with residues at TM6 extracellular end, resulting in an inner continuous cavity (Chapter 3).

7.2 Strengths and Limitations

Without doubts homology modelling and MD simulations have become valuable tools to obtain detailed information about GPCR conformation, interaction with the ligand and signalling pathways with applications to drug discovery. However, like any other method, homology modelling and MD simulations have strengths and limitations.

The main limitation of homology modelling is that the resulting model directly depends on the quality of the template. Errors in sidechains, loops or secondary structures need to be considered. However, the models used during this project were constructed using high quality structures at atomic resolution, incorporating all available experimental data and using the best available data at the time, resulting in high quality models. The available experimental data was carefully considered when building missing sidechain or segments. In most cases, the original GPCR structure underwent mutations or structural modifications to facilitate the structure determination, therefore, the structures in which the models were based were different from the human wild-type receptor or have missing segments. Homology modelling and available mutagenesis and experimentally data was used to change the sequences to the human wild-type and to build the missing segments. To avoid steric clashes resulting from homology modelling, all models underwent energy minimization. Then the resulting models used in this project where simulated and the resulting structures revealed important interactions previously shown by mutagenesis, experimental data and in agreement the new structures that followed.

GPCR are dynamic structures unlike the static representation form the reported experimentally determined structure in the PDB or model. To overcome this issue, the receptor was simulated using different MD methods under force fields parameters that have been validated on experimentally obtained data from proteins

and lipids molecules. All-atoms classic MD simulation was used for the simulation of GLP-1R presented in Chapter 3, and then a variation of all-atoms MD, MDFF, simulation was used for simulations of GLP-1R, β2AR and PTH1 receptor (Chapters 4, 5 and 6). All-atoms MD and MDFF method enables the study at atomic detail of GPCR, allowing the receptor to find a stable state. In Chapters 4 to 6, MDFF method was used to improve the simulation and results. MDFF methods adds the information contained in density map as external forces proportionally to the gradient of the density map to a molecular dynamic simulation. This allows to add as much experimentally obtained data, even to segments with poor resolution, and avoids the use of reduced representations or a single PDB structure improving the simulations allowing to add more experimental data. In this way, high resolution areas drive atoms into high-density regions, low resolution area allows the flexibility of the structure while since MDFF fitting is performed locally, while MD force fields and harmonic restraints avoid structural distortion keeping the integrity of the structure. However, an important limitation is that during MD simulations, classic all-atoms, MDFF or Targeted MD, covalent bonds are not form or broken.

GPCRs are membrane receptors, the models were embedded into a lipid environment to simulate membrane composition. Most MD simulations of GPCR use a POPC homogeneous lipid bilayer. The first simulation of GLP-1R presented in Chapter 3 used a homogeneous POPC membrane. However, lipids have a physiological role in the function, dynamics, and integrity of the membrane as well of the receptor function. I consider that a strength of the following simulations, presented in Chapter 4 to 6, was the use of a heterogeneous membrane with an asymmetrical lipid distribution similar to the mammalian membranes. The heterogeneous membrane composition considers the physiological effect of the different lipids in the activation of GPCRs.

One of the main limitations of MD simulations used during this project was the timescales. Most relevant conformational changes and activation of GPCRs occur in milliseconds, or higher timescales, which despite the increasing computational power still timescales represent one of the major limitations in the study of GPCRs where large conformational changes and activation become impossible to observe in classic all-atoms MD. To overcome the timescale limitation, Targeted MD simulation method was used to study activation of GLP-1R, β2AR and PTH1 receptor (Chapters 5 and 6). Since the activation mechanism of β2AR has been widely studied, this

receptor was used as a control to verify the accuracy of Targeted MD, the results were in accordance to known data. Targeted MD is not often used for the study of GPCRs, however, this method proved to be reliable to study GPCR activation and to overcome the timescale limitation. Targeted MD calculates the transition pathways between two known structures, guiding a subset of atoms, in this case the main chain, towards the final structure without affecting the mobility of other atoms or sidechains but instead allowing enough flexibility to explore plausible pathways and accommodate.

7.3 Similarities and differences to other studies

The growing number of experimentally determined GPCR structures have increased the understanding of GPCRs. Previously, the accuracy of models built using homology modelling method was limited by the low homology percentage between the available templates and the sequence identity, however, the growing number of high-quality experimentally determined GPCRs structures has improve the quality and accuracy of the models (reviewed in Bender, Marlow and Meiler, 2020). In addition, MD simulations have been successfully used to obtain GPCR structural information and the different conformational ensembles that could be impossible to obtain by experimental methods (Dror et al. 2011; Hauser et al. 2017; Sabbadin and Moro, 2014). This project relied on the new experimentally obtained GLP-1R structures. The best available data at the time was used during this project, and as new experimentally determined GLP-1R structures became available there were improvements in the models. Classical all-atoms MD, MDFF and targeted MD simulations methods were used with the aim to incorporate as much experimentally obtained data to the simulations. After being simulated, the resulting complexes showed stable receptor structures interacting with their respectively ligands. However, there is the possibility that the resulting structures are not found at the global minimum energy conformation due to the time-step limitation, or the existence of more than one stable structure; still the resulting structures provide information of GLP-1R activation mechanisms and order of events that would be useful for the design of novel drugs and the reduction of side effects.

In the first simulation described in Chapter 3, a homogeneous POPC membrane was used. MD simulations in the literature have use homogeneous POPC

lipid bilayers as these matches well with experimental data (reviewed in Zhuang et al. 2014; Wootten et al. 2016). However, one of the main differences in this project was the membrane composition used in the simulations described in Chapters 4 to 6. In these simulations a heterogeneous lipid bilayer was built to increase the realism of the mammalian membrane and to account for the effect of the lipids in the receptor. During the simulation no detection of lipid penetration was found.

A second difference were the MD methods used. Most of simulations of proteins are performed using all-atoms or coarse-grain MD. The coarse grain model method has been used to describe GPCR dynamics (reviewed in Periole, 2017). This method allows longer timescales, however, the main limitation is that in some cases the model can be too rough to identify certain interactions. Unlike the coarse grain method, all-atoms MD simulation can better identify detailed protein-protein interactions, however, the length of the simulation is limited (reviewed in Periole 2017). To get a better look at the interactions the all-atoms MD method and variations of this mehtod were used during this project. One of the variations was the use of the MDFF method that incorporates the electron density into the simulation, adding experimentally obtained data into the simulation. The use of MDFF simualtions, described in Chapters 4 to 6, was a difference between this study and published GLP-1R simulations at the time.. Although cryo-EM is used to determine GPCR structures, until 2020 there were no GLP-1R – GLP-1 complex simulated using MDFF. The density gradient obtained from experimentally determined structures is used to drive the model structures to high density areas while maintaining the flexibility of structures reported by low density gradiente. In this way, it is possible to incorporate as much experimentally obtained data into the simulation. MDFF method was used for the study of class B GPCRs because of two main reasons: 1) to incorporate the density gradient into the simulations and 2) to avoid the 'bending' of the NTD parallel to the lipid bilayer in longer simulations (data not shown), which hindered the study of ligand-receptor interactions. Therefore, GLP-1R – GLP-1 complex in the active and inactive states were simulated using MDFF; the density gradient was applied and later removed from the TMD and ligands. The resulting active and inactive models showed some differences with previously classic all-atoms simulated models and provided information about peptide agonists engagement with GLP-1R. In concordance with available data, the MDFF simulated models showed 1) the importance of the hydrophobic interactions between GLP-1 and TMD1-TMD2, 2) the

disruption of TMD1-TMD2 interface facilitates the entrance of the ligand into the bidning pocket (Bueno et al. 2020), and 3) triggers further conformational changes as the presence of the ligand between TMD1-TMD2 facilitates the the shift of TMD1 over TMD7, and then the movement of TMD7 over TMD6 which improves the binding of the ligand and closing the binding pocket; the characteristic interaction between class B GPCR TMD2 residue, Arg190$^{2.60b}$, and the ligand Glu9* promotes rotamer changes in the middle region of TMD2 below Arg190$^{2.60b}$ which creates a small cavity allowing TMD6 rotation and kink formation. Unlike previously observed in the active all-atoms classic MD, where the N-terminal residue His7* interacts with Glu364$^{6.53b}$, in the active MDFF simulation His7* interacts with Glu387$^{7.42b}$ contributing to the shift of TMD7 extracellular end towards TMD6 and showing how Glu387$^{7.42b}$ contributes to GLP-1 ligand binding by interacting with His7*. Indeed, Glu387$^{7.42b}$ is necessary for GLP-1 efficacy (Dods and Donnelly, 2015), but not for ExP5 (Liang et al. 2018). Another variation of the all-atoms MD simulations that was used during this project was Targeted MD. Despite the advances in computational power, spontaneous conformational changes such activation occur in in greater timescales that could be almost impossible to observe using classic all-atoms MD. Targeted MD guide atoms from an initial conformation to the targeted one (Schlitter, Engels and Kruger 1994). For these reasons Targeted MD was used to study GLP-1R activation. In addition the last frames obtained from the MDFF simulation of the active and inactive states where used as these represented high quality models of the GLP-1R in complex with GLP-1, simulated using the density gradiente and after the removal of the restraints from the TMD and ligand, the structures remained stable. At the time Targeted MD had been successfully used for the study of protein unfolding, the search of pathways of conformational transitions,identification of biologically relevant drug molecules, allosteric mechanisms of calmodulin (Schlitter, Engels and Kruger 1994; Ferrara, Apostolakis and Caflisch, 2000; Rumpf et al 2015; Jones et al. 1997; Liang et al. 2017;), but not larger systems, such GLP-1R. To validate the use of targeted MD simulations for GLP-1R, the β_2AR was used since its activation mechanism are well known. The resulting simulation confirmed the movemnt of the ligand via TMD1-TMD2 interface, the importance of Arg190$^{2.60b}$ – Glu9* interaction, the packing of TMD2 towareds TMD3 creating a small inner cavity allowing the formation of TMD6 kink,

the movment of TMD7 towards TMD6 during activation and the role of Gln394$^{7.49b}$ stabilising TMD6 kink.

Despite the number of class B GPCR structures available and studies, there is not a concensus about the position of the NTD in the active and inactive states. Zhang et. al investigated conformational transition of the ECD using MD and Markov state model (Zhang et al. 2019) in the bound and apo GLP-1R state for a period of 1.6 μs finding that while in the bound state, GLP-1R and GLP-1 binds closely to the ECD, avoiding the movement of the ECD, in the apo-state ECL1 and ECL2 lock the binding pocket keeping the ECD away (Zhang et al. 2019). However, their approach differs from the ones used during this project as they rely on the Markov state model, which sample and divide conformational space into states showing the transfer rate between conformations. One of the main limitations of the Markov models time, as Markov models is inappropriate for short lengths of time as the individual displacement are determined on time. Although Markov methods can be very effective, all-atom simulation provide more accurate and realistic information (reviewed in He, Paul, and Roux 2021). Class B the position of the NTD depends on the ligand; different position of the NTD relative to the TMD are observed in GLP-1R thermostabilised 5NX2, cryo-EM 5VAI and 6B3J or in PTH_1 receptor thermostabilised 6FJ3, and cryo-EM 6NBH, 6NBF, 6NBI. To overcome the 'falling' from the NTD and study NTD-ligand interactions, the NTD was restraint to the density map, allowing the study of the ligand movement into the binding pocket, but maintain receptor-ligand interactions. Still differences in individual residues and interactions were found. When simulated the GLP-1-GLP-1R bound receptor, similar to the reported Markov model simulation, the ligand remained bound to the receptor but different contacts were found. Zhang et al found that residues Val30^{NTD}, Leu32^{NTD}, TRP39^{NTD} maintained the interaction with residues in the C-terminal of GLP-1 Ala24*, Lys26*, Phe28* and Leu32* (Zhang et al. 2019), in concordance with the two-steps model and these interactions can't be ignored as may represent a different state of GLP-1R - GLP-1. However, in the targeted MD GLP-1R simulation Lys26* was found interacting with the extracellular end of TM1. Overlay and comparison from the result of active state MDFF simulation and active from targeted MD with GLP-1R in complex with ExP5 (PDB 6B3J) or GLP-1 (PDB 5VAI) show differences in the position of the TM1, the extracellular end of TM7, the outward movement of TM6 and the outward movement of the intracellular half of TM6. All structures, simulated

and experimentally show the shift of TM1 over TM7, followed by TM7 movement towards TM6. From these structures, 6B3J showed the largest movement, with the N-terminal pointing towards TM6. TM6-ECL3-TM7 region is involved in biased agonism, and changes in the region are determinant for G protein-biased signalling. (Liang et al. 2018; Dods and Donnelly 2015; Wootten and Miller 2020; Wootten et al. 2016; Wootten et al. 2016a; Wootten et al. 2016b). Targeted MD showed His7* interacting with Glu364$^{6.53b}$ which simultaneously interact with Arg310$^{5.40b}$ allowing TM6 extracellular end outward movement. ExP5-GLP-1R, E1* interaction with Arg310$^{5.40b}$ is crucial for cAMP accumulation, (Liang et al. 2018). In addition, recently improvement of 5VAI modelling and MD simulation show His7* interaction with Arg310$^{5.40b}$, Gn2343$^{.37b}$, Val237$^{3.40b}$, Trp306$^{5.36b}$ and Ile313$^{5.43b}$ and, through structural water with TM5 (Zhang, Qiao, et al. 2017; Zhang et al. 2020b).

7.4 Summary

The aim of this project was to gain an understanding GLP-1R's molecular mechanism of action and the key interactions leading to agonist-mediated activation. At the start of this project (February 2017) there were no experimentally determined GLP-1R structures available, and hence the initial approach was to use experimentally determined structures of other family B GPCRs for homology modelling of GLP-1R. However, soon after, GLP-1R structures in active and TMD of inactive state became available (Jazayeri et al. 2017; Zhang et al. 2017; Song et al. 2017). These high-quality structures, and the ones that followed, provided valuable information about GLP-1R interactions with peptide agonists, but not with the endogenous human GLP-1 peptide. Therefore GLP-1R-GLP-1 complex models in the active and inactive state were built and simulated using MD simulations with the aim of studying active and inactive state as well understanding GLP-1R activation. After defining active and inactive state, a second class B GPCRs, the human PTH_1 receptor in complex with PTH model in the active and inactive state was built to compare characteristics of active and inactive state in Class as well activation mechanisms.

Simulation of GLP-1R and PTH_1 receptor in complex with their respectively ligands show that the ligands disrupt TM1-TM2 interface, relying in hydrophobic residues to facilitate the movement into the binding pocket and charged residues

Lys$^{2.67b}$. At the extracellular end of TM7, the movement of the ligand into the binding pocket shift TM7 towards TM6 contributing to the rearrangement of TM6. In addition, interaction between the ligand's charged residue in the N-terminal (His7* in GLP-1, Ser1* PTH). Charged residues in helix 2, Arg$^{2.60b}$, - contribute to the binding and stabilisation of the ligand as well to translate conformational changes in the middle region of TM3-TM2, where Asn$^{3.43b}$ remove spatial restraints from the core of the receptor. The removal of spatial restrains is seen as a packing of TM3-TM2 that allows Leu$^{3.47b}$ rotamer change. Rotamer change in hydrophobic residue in helix 3 (GLP-1R Leu244$^{3.47b}$ / PTH_1 receptor) push hydrophobic residue in helix 6 (GLP-1R Leu360$^{6.49b}$ / Phe417$^{6.49b}$ PTH_1 receptor) to produce TM6 outward movement. In consequence, surrounding hydrophobic residues (Leu$^{2.53b}$, Leu$^{6.48b}$, Leu$^{6.45b}$, Val$^{7.53b}$) change rotamers to stabilise the core of the receptor in the active conformation. In the extracellular end of TM5-TM6, the charged residue of the ligand's N-terminal (His7* in GLP-1, Ser1* PTH) modify GLP-1R Arg310$^{5.40b}$-Glu364$^{6.53b}$ or Ser1* PTH Gln364$^{5.40b}$ – Tyr421$^{6.53b}$ PTH_1 receptor allowing TM6 flexibility and outward movement while maintaining the stability of the receptor.

References

Abou-Samra, Abdul-Badi, H Jüppner, Thomas Force, Mason W Freeman, Xiang-Fu Kong, Ernestina Schipani, Pablo Urena, Jennifer Richards, Joseph V Bonventre, and John T Potts. 1992. 'Expression cloning of a common receptor for parathyroid hormone and parathyroid hormone-related peptide from rat osteoblast-like cells: a single receptor stimulates intracellular accumulation of both cAMP and inositol trisphosphates and increases intracellular free calcium', *Proceedings of the National Academy of Sciences*, 89: 2732-36.

Ackermann, Theodor. 1990. 'C. L. Brooks III, M. Karplus, B. M. Pettitt. Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics, Volume LXXI, in: Advances in Chemical Physics, John Wiley & Sons, New York 1988. 259 Seiten, Preis: US $ 65.25', 94: 96-96.

ADA. 2018. 'Economic Costs of Diabetes in the U.S. in 2017', *Diabetes Care*: dci180007.

Adcock, Stewart A., and J. Andrew McCammon. 2006. 'Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins', *Chem Rev*, 106: 1589-615.

Adelhorst, K., B. B. Hedegaard, L. B. Knudsen, and O. Kirk. 1994a. 'Structure-activity studies of glucagon-like peptide-1', *J Biol Chem*, 269: 6275-8.

Adelhorst, Kim, BB Hedegaard, L Bo Knudsen, and O Kirk. 1994b. 'Structure-activity studies of glucagon-like peptide-1', *Journal of Biological Chemistry*, 269: 6275-78.

Alexiadou, K., Anyiam, O. & Tan, T.2019. 'Cracking the combination: Gut hormones for the treatment of obesity and diabetes ' J Neuroendocrinol 31, e12664, doi:10.1111/jne.12664

Al-Sabah, S., and D. Donnelly. 2003a. 'A model for receptor-peptide binding at the glucagon-like peptide-1 (GLP-1) receptor through the analysis of truncated ligands and receptors', *Br J Pharmacol*, 140: 339-46.

Al-Sabah, Suleiman, and Dan Donnelly. 2003b. 'The positive charge at Lys-288 of the glucagon-like peptide-1 (GLP-1) receptor is important for binding the N-terminus of peptide agonists', *FEBS Lett*, 553: 342-46.

Alberts B, Johnson A, Lewis J. 2002. 'The Lipid Bilayer', Garland Science. https://www.ncbi.nlm.nih.gov/books/NBK26871/.

Alvarez-Curto, Elisa, Asuka Inoue, Laura Jenkins, Sheikh Zahir Raihan, Rudi Prihandoko, Andrew B Tobin, and Graeme Milligan. 2016. 'Targeted elimination of G proteins and arrestins defines their specific contributions to both intensity and duration of G protein-coupled receptor signaling', *Journal of Biological Chemistry*, 291: 27147-59.

Andersen, Niels H, Yan Brodsky, Jonathan W Neidigh, and Kathryn S Prickett. 2002. 'Medium-dependence of the secondary structure of exendin-4 and glucagon-like-peptide-1', *Bioorganic & medicinal chemistry*, 10: 79-85.

Andre De Lean, Jeffrey Stadel, Robert Lefkowitz. 1980. 'A ternary complex model explains the agonist-specific binding properites of the adenylate cyclase-couple beta-adrenergi receptor', *J Biol Chem*, 255: 7108-17.

Archbold, J. K., J. U. Flanagan, H. A. Watkins, J. J. Gingell, and D. L. Hay. 2011. 'Structural insights into RAMP modification of secretin family G protein-coupled receptors: implications for drug development', *Trends Pharmacol Sci*, 32: 591-600.

Ariens, E. J. 1954. 'Affinity and intrinsic activity in the theory of competitive inhibition. I. Problems and theory', *Arch Int Pharmacodyn Ther*, 99: 32-49.

Asmar, Meena, Jens J %J Current Opinion in Endocrinology Holst, Diabetes, and Obesity. 2010. 'Glucagon-like peptide 1 and glucose-dependent insulinotropic polypeptide: new advances', 17: 57-62.

Ast, Julia, Anastasia Arvaniti, Nicholas H. F. Fine, Daniela Nasteska, Fiona B. Ashford, Zania Stamataki, Zsombor Koszegi, Andrea Bacon, Ben J. Jones, Maria A. Lucey, Shugo Sasaki, Daniel I. Brierley, Benoit Hastoy, Alejandra Tomas, Giuseppe D'Agostino, Frank Reimann, Francis C. Lynn, Christopher A. Reissaus, Amelia K. Linnemann, Elisa D'Este, Davide Calebiro, Stefan Trapp, Kai Johnsson, Tom Podewin, Johannes Broichhagen, and David J. Hodson. 2020. 'Super-resolution microscopy compatible fluorescent probes reveal endogenous glucagon-like peptide-1 receptor distribution and dynamics', *Nat Commun*, 11: 467.

Attwood, T. K., and J. B. C. Findlay. 1994. 'Fingerprinting G-protein-coupled receptors', *Protein Engineering, Design and Selection*, 7: 195-203.

Austin, Robert Hamilton, KW Beeson, L Eisenstein, H Frauenfelder, and IC Gunsalus. 1975. 'Dynamics of ligand binding to myoglobin', *Biochemistry*, 14: 5355-73.

Baggio, L. L., and D. J. Drucker. 2014. 'Glucagon-like peptide-1 receptors in the brain: controlling food intake and body weight', *J Clin Invest*, 124: 4223-6.

Baggio, L. L., B. Yusta, E. E. Mulvihill, X. Cao, C. J. Streutker, J. Butany, T. P. Cappola, K. B. Margulies, and D. J. Drucker. 2018. 'GLP-1 Receptor Expression Within the Human Heart', *Endocrinology*, 159: 1570-84.

Baggio, L. L., and D. J. Drucker. 2021. 'Glucagon-like peptide-1 receptor co-agonist for treating metabolic disease', *Molecular Metabolism,* 46:101090.

Bailey, R. J., and D. L. Hay. 2007. 'Agonist-dependent consequences of proline to alanine substitution in the transmembrane helices of the calcitonin receptor', *Br J Pharmacol*, 151: 678-87.

Ballesteros, Juan A, and Harel Weinstein. 1995. '[19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors.' in, *Methods in neurosciences* (Elsevier).

Bang, Injin, and Hee-Jung Choi. 2015. 'Structural features of β2 adrenergic receptor: crystal structures and beyond', *Molecules and cells*, 38: 105-11.

Bank, Protein Data. 2020. 'PDB Statistics: Overall Growth of Released Structures Per Year', Accessed February 23th 2020. https://www.rcsb.org/stats/growth/overall.

Barth, Eric, Krzysztof Kuczera, Benedict Leimkuhler, and Robert D Skeel. 1995. 'Algorithms for constrained molecular dynamics', *J Comput Chem*, 16: 1192-209.

Bavec, Aljoša, Mattias Hällbrink, Ülo Langel, and Matjaž Zorko. 2003. 'Different role of intracellular loops of glucagon-like peptide-1 receptor in G-protein coupling', *Regul Pept*, 111: 137-44.

Bender BJ, Marlow B, Meiler J. Improving homology modeling from low-sequence identity templates in Rosetta: A case study in GPCRs. PLoS Comput Biol. 2020;16(10):e1007597. Epub 2020/10/29. pmid:33112852; PubMed Central PMCID: PMC7652349.

Benovic, JL, H Kühn, I Weyand, J Codina, MG Caron, and RJ Lefkowitz. 1987. 'Functional desensitization of the isolated beta-adrenergic receptor by the beta-adrenergic receptor kinase: potential role of an analog of the retinal protein

arrestin (48-kDa protein)', *Proceedings of the National Academy of Sciences*, 84: 8879-82.

Berg JM, Tymoczko JL, Stryer L. 2002. *Chapter 3, Protein Structure and Function.* (W H Freeman: New York).

Bergwitz, C., T. J. Gardella, M. R. Flannery, J. T. Potts, Jr., H. M. Kronenberg, S. R. Goldring, and H. Juppner. 1996. 'Full activation of chimeric receptors by hybrids between parathyroid hormone and calcitonin. Evidence for a common pattern of ligand-receptor interaction', *J Biol Chem*, 271: 26469-72.

Berne, C. 1975. 'The metabolism of lipids in mouse pancreatic islets. The biosynthesis of triacylglycerols and phospholipids', *Biochem J*, 152: 667-73.

Bharat, T. A., L. R. Castillo Menendez, W. J. Hagen, V. Lux, S. Igonet, M. Schorb, F. K. Schur, H. G. Kräusslich, and J. A. Briggs. 2014. 'Cryo-electron microscopy of tubular arrays of HIV-1 Gag resolves structures essential for immature virus assembly', *Proc Natl Acad Sci U S A*, 111: 8233-8.

Bienert, Stefan, Andrew Waterhouse, Tjaart A P de Beer, Gerardo Tauriello, Gabriel Studer, Lorenza Bordoli, and Torsten Schwede. 2017. 'The SWISS-MODEL Repository—new features and functionality', *Nucleic Acids Research*, 45: D313-D19.

Bill, R. M., P. J. Henderson, S. Iwata, E. R. Kunji, H. Michel, R. Neutze, S. Newstead, B. Poolman, C. G. Tate, and H. Vogel. 2011. 'Overcoming barriers to membrane protein structure determination', *Nat Biotechnol*, 29: 335-40.

Black, J. W., and P. Leff. 1983. 'Operational models of pharmacological agonism', *Proc R Soc Lond B Biol Sci*, 220: 141-62.

Bloch, Konrad. 1991. 'Chapter 12 Cholesterol: evolution of structure and function.' in Dennis E. Vance and Jean E. Vance (eds.), *New Comprehensive Biochemistry* (Elsevier).

Blumer, Joe B., Mary J. Cismowski, Motohiko Sato, and Stephen M. Lanier. 2005. 'AGS proteins: receptor-independent activators of G-protein signaling', *Trends Pharmacol Sci*, 26: 470-76.

Bockaert, J., and J. P. Pin. 1999a. 'Molecular tinkering of G protein-coupled receptors: an evolutionary success', *Embo j*, 18: 1723-29.

———. 1999b. 'Molecular tinkering of G protein-coupled receptors: an evolutionary success', *Embo j*, 18: 1723-9.

Bockenhauer, S., A. Furstenberg, X. J. Yao, B. K. Kobilka, and W. E. Moerner. 2011. 'Conformational dynamics of single G protein-coupled receptors in solution', *J Phys Chem B*, 115: 13328-38.

Bologna, Zuzana, Jian-peng Teoh, Ahmed S. Bayoumi, Yaoliang Tang, and Il-man Kim. 2017. 'Biased G Protein-Coupled Receptor Signaling: New Player in Modulating Physiology and Pathology', *Biomolecules & Therapeutics*, 25: 12-25.

Bond, Aaron. 2006. 'Exenatide (Byetta) as a novel treatment option for type 2 diabetes mellitus', *Proceedings (Baylor University. Medical Center)*, 19: 281-84.

Bringhurst, FR, H Juppner, J Guo, P Urena, JT Potts Jr, HM Kronenberg, AB Abou-Samra, and GV Segre. 1993. 'Cloned, stably expressed parathyroid hormone (PTH)/PTH-related peptide receptors activate multiple messenger signals and biological responses in LLC-PK1 kidney cells', *Endocrinology*, 132: 2090-98.

Brooks, B. R., C. L. Brooks, 3rd, A. D. Mackerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu,

W. Yang, D. M. York, and M. Karplus. 2009. 'CHARMM: the biomolecular simulation program', *J Comput Chem*, 30: 1545-614.

Bucheit, J. D., L. G. Pamulapati, N. Carter, K. Malloy, D. L. Dixon, and E. M. Sisson. 2020. 'Oral Semaglutide: A Review of the First Oral Glucagon-Like Peptide 1 Receptor Agonist', *Diabetes Technol Ther*, 22: 10-18.

Buchwald, Peter. 2019. 'A Receptor Model With Binding Affinity, Activation Efficacy, and Signal Amplification Parameters for Complex Fractional Response Versus Occupancy Data', *Frontiers in Pharmacology*, 10.

Bueno, Ana B., Bingfa Sun, Francis S. Willard, Dan Feng, Joseph D. Ho, David B. Wainscott, Aaron D. Showalter, Michal Vieth, Qi Chen, Cynthia Stutsman, Betty Chau, James Ficorilli, Francisco J. Agejas, Graham R. Cumming, Alma Jiménez, Isabel Rojo, Tong Sun Kobilka, Brian K. Kobilka, and Kyle W. Sloop. 2020. 'Structural insights into probe-dependent positive allosterism of the GLP-1 receptor', *Nat Chem Biol*, 16: 1105-10.

Burke, B., and J. Ellenberg. 2002. 'Remodelling the walls of the nucleus', *Nat Rev Mol Cell Biol*, 3: 487-97.

Caflisch, Amedeo, Francesco Rao, Giovanni Settanni, Marco Cecchini, and Michele Seeber. 2007. 'Wordom: a program for efficient analysis of molecular dynamics simulations', *Bioinformatics*, 23: 2625-27.

Cahill, Thomas J, Alex RB Thomsen, Jeffrey T Tarrasch, Bianca Plouffe, Anthony H Nguyen, Fan Yang, Li-Yin Huang, Alem W Kahsai, Daniel L Bassoni, and Bryant J Gavino. 2017. 'Distinct conformations of GPCR–β-arrestin complexes mediate desensitization, signaling, and endocytosis', *Proceedings of the National Academy of Sciences*, 114: 2562-67.

Caillon, Lucie, Olivier Lequin, and Lucie Khemtémourian. 2013. 'Evaluation of membrane models and their composition for islet amyloid polypeptide-membrane aggregation', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1828: 2091-98.

Carafoli., A. Scarpa and E. 1979. 'Calcium transport and cell function, edited by New York Academy of Sciences, ' *Journal of Neurobiology*, 10: 201-01.

Carpenter, Byron, Rony Nehmé, Tony Warne, Andrew G. W. Leslie, and Christopher G. Tate. 2016. 'Structure of the adenosine A2A receptor bound to an engineered G protein', *Nature*, 536: 104.

Case, D. A., T. E. Cheatham, 3rd, T. Darden, H. Gohlke, R. Luo, K. M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods. 2005. 'The Amber biomolecular simulation programs', *J Comput Chem*, 26: 1668-88.

Castro, Marián, Viacheslav O. Nikolaev, Dieter Palm, Martin J. Lohse, and Jean-Pierre Vilardaga. 2005. 'Turn-on switch in parathyroid hormone receptor by a two-step parathyroid hormone binding mechanism', *Proc Natl Acad Sci U S A*, 102: 16084-89.

CDC. 2017. "National Diabetes Statistics Report." In. Atlanta, GA: Centers for Disease Control and Prevention.

Cegla, J., B. J. Jones, J. V. Gardiner, D. J. Hodson, T. Marjot, E. R. McGlone, T. M. Tan, and S. R. Bloom. 2017. 'RAMP2 Influences Glucagon Receptor Pharmacology via Trafficking and Signaling', *Endocrinology*, 158: 2680-93.

Center for Informational Biology, Ochanomizu University. 2012. "Accessible Surface Area and Accessibility Calculation for Protein " In.

Chan, H. C. Stephen, Slawomir Filipek, and Shuguang Yuan. 2016. 'The Principles of Ligand Specificity on beta-2-adrenergic receptor', *Scientific Reports*, 6: 34736.

Chan, J. M., E. B. Rimm, G. A. Colditz, M. J. Stampfer, and W. C. Willett. 1994. 'Obesity, fat distribution, and weight gain as risk factors for clinical diabetes in men', *Diabetes Care*, 17: 961-9.

Chang, Rulue, Xin Zhang, Anna Qiao, Antao Dai, Matthew J. Belousoff, Qiuxiang Tan, Lijun Shao, Li Zhong, Guangyao Lin, Yi-Lynn Liang, Limin Ma, Shuo Han, Dehua Yang, Radostin Danev, Ming-Wei Wang, Denise Wootten, Beili Wu, and Patrick M. Sexton. 2020. 'Cryo-electron microscopy structure of the glucagon receptor with a dual-agonist peptide', *Journal of Biological Chemistry*, 295: 9313-25.

Chapman, M. S. 1995. 'Restrained real-space macromolecular atomic refinement using a new resolution-dependent electron-density function', 51: 69-80.

Che, Tao, Susruta Majumdar, Saheem A. Zaidi, Pauline Ondachi, John D. McCorvy, Sheng Wang, Philip D. Mosier, Rajendra Uprety, Eyal Vardy, Brian E. Krumm, Gye Won Han, Ming-Yue Lee, Els Pardon, Jan Steyaert, Xi-Ping Huang, Ryan T. Strachan, Alexandra R. Tribo, Gavril W. Pasternak, F. Ivy Carroll, Raymond C. Stevens, Vadim Cherezov, Vsevolod Katritch, Daniel Wacker, and Bryan L. Roth. 2018. 'Structure of the Nanobody-Stabilized Active State of the Kappa Opioid Receptor', *Cell*, 172: 55-67.e15.

Cheloha, Ross W, Samuel H Gellman, Jean-Pierre Vilardaga, and Thomas J Gardella. 2015. 'PTH receptor-1 signalling—mechanistic insights and therapeutic prospects', *Nature Reviews Endocrinology*, 11: 712.

Chen, L. F., E. Blanc, M. S. Chapman, and K. A. Taylor. 2001. 'Real space refinement of acto-myosin structures from sectioned muscle', *J Struct Biol*, 133: 221-32.

Cheng, Y., N. Grigorieff, P. A. Penczek, and T. Walz. 2015. 'A primer to single-particle cryo-electron microscopy', *Cell*, 161: 438-49.

Cherezov, Vadim, Daniel M. Rosenbaum, Michael A. Hanson, Søren G. F. Rasmussen, Foon Sun Thian, Tong Sun Kobilka, Hee-Jung Choi, Peter Kuhn, William I. Weis, Brian K. Kobilka, and Raymond C. Stevens. 2007. 'High-Resolution Crystal Structure of an Engineered Human β₂-Adrenergic G Protein–Coupled Receptor', 318: 1258-65.

Cheung, WY. 1980. 'Calmodulin plays a pivotal role in cellular regulation', *Science*, 207: 19-27.

Choe, H. W., Y. J. Kim, J. H. Park, T. Morizumi, E. F. Pai, N. Krauss, K. P. Hofmann, P. Scheerer, and O. P. Ernst. 2011. 'Crystal structure of metarhodopsin II', *Nature*, 471: 651-5.

Christensen, Mikkel, and Filip K %J Current diabetes reports Knop. 2010. 'Once-weekly GLP-1 agonists: How do they differ from exenatide and liraglutide?', 10: 124-32.

Christopoulos, A., G. Christopoulos, M. Morfis, M. Udawela, M. Laburthe, A. Couvineau, K. Kuwasako, N. Tilakaratne, and P. M. Sexton. 2003. 'Novel receptor partners and function of receptor activity-modifying proteins', *J Biol Chem*, 278: 3293-7.

Clairfeuille, T., Pavlos, N., Collins, B.M. 2015. "Crystal structure of the SNX27 PDZ domain bound to the C-terminal PTHR PDZ binding motif." In *4Z8J*. PDB bank: PDB bank.

Colditz, G. A., W. C. Willett, A. Rotnitzky, and J. E. Manson. 1995. 'Weight gain as a risk factor for clinical diabetes mellitus in women', *Ann Intern Med*, 122: 481-6.

Cong, Zhaotong, Li-Nan Chen, Honglei Ma, Qingtong Zhou, Xinyu Zou, Chenyu Ye, Antao Dai, Qing Liu, Wei Huang, Xianqiang Sun, Xi Wang, Peiyu Xu, Lihua Zhao, Tian Xia, Wenge Zhong, Dehua Yang, H. Eric Xu, Yan Zhang, and

Ming-Wei Wang. 2021. 'Molecular insights into ago-allosteric modulation of the human glucagon-like peptide-1 receptor', *Nat Commun*, 12: 3763.

Conklin, Bruce R, Zvi Farfel, Kevin D Lustig, David Julius, and Henry R Bourne. 1993. 'Substitution of three amino acids switches receptor specificity of Gqα to that of Giα', *Nature*, 363: 274-76.

Conner, A. C., D. L. Hay, J. Simms, S. G. Howitt, M. Schindler, D. M. Smith, M. Wheatley, and D. R. Poyner. 2005. 'A key role for transmembrane prolines in calcitonin receptor-like receptor agonist binding and signalling: implications for family B G-protein-coupled receptors', *Mol Pharmacol*, 67: 20-31.

Conner, Alex C, John Simms, Matthew T Conner, Denise L Wootten, Mark Wheatley, and David R Poyner. 2006. 'Diverse functional motifs within the three intracellular loops of the CGRP1 receptor', *Biochemistry*, 45: 12976-85.

Coopman, K., R. Wallis, G. Robb, A. J. Brown, G. F. Wilkinson, D. Timms, and G. B. Willars. 2011. 'Residues within the transmembrane domain of the glucagon-like peptide-1 receptor involved in ligand binding and receptor activation: modelling the ligand-bound receptor', *Mol Endocrinol*, 25: 1804-18.

Couvineau, A., C. Rouyer-Fessard, J. J. Maoret, P. Gaudin, P. Nicole, and M. Laburthe. 1996. 'Vasoactive intestinal peptide (VIP)1 receptor. Three nonadjacent amino acids are responsible for species selectivity with respect to recognition of peptide histidine isoleucineamide', *J Biol Chem*, 271: 12795-800.

Crane, Jonathan M., and Lukas K. Tamm. 2004. 'Role of cholesterol in the formation and nature of lipid rafts in planar and spherical model membranes', *Biophys J*, 86: 2965-79.

Creutzfeldt, W. O., N. Kleine, B. Willms, C. Orskov, J. J. Holst, and M. A. Nauck. 1996. 'Glucagonostatic actions and reduction of fasting hyperglycemia by exogenous glucagon-like peptide I(7-36) amide in type I diabetic patients', *Diabetes Care*, 19: 580-6.

Culhane, K. J., M. E. Belina, J. N. Sims, Y. Cai, Y. Liu, P. S. P. Wang, and E. C. Y. Yan. 2018. 'Parathyroid Hormone Senses Extracellular Calcium To Modulate Endocrine Signaling upon Binding to the Family B GPCR Parathyroid Hormone 1 Receptor', *ACS Chem Biol*, 13: 2347-58.

Cummings, Bethany P., Kimber L. Stanhope, James L. Graham, Denis G. Baskin, Steven C. Griffen, Cecilia Nilsson, Anette Sams, Lotte B. Knudsen, Kirsten Raun, and Peter J. Havel. 2010. 'Chronic administration of the glucagon-like peptide-1 analog, liraglutide, delays the onset of diabetes and lowers triglycerides in UCD-T2DM rats', *Diabetes*, 59: 2653-61.

Cypess, Aaron M, Cecilia G Unson, Cui-Rong Wu, and Thomas P Sakmar. 1999. 'Two cytoplasmic loops of the glucagon receptor are required to elevate cAMP or intracellular calcium', *Journal of Biological Chemistry*, 274: 19455-64.

Danev, Radostin, Haruaki Yanagisawa, and Masahide Kikkawa. 2019. 'Cryo-Electron Microscopy Methodology: Current Aspects and Future Directions', *Trends Biochem Sci*, 44: 837-48.

de Graaf, C., G. Song, C. Cao, Q. Zhao, M. W. Wang, B. Wu, and R. C. Stevens. 2017. 'Extending the Structural View of Class B GPCRs', *Trends Biochem Sci*, 42: 946-60.

de Graaf, Chris, Dan Donnelly, Denise Wootten, Jesper Lau, Patrick M. Sexton, Laurence J. Miller, Jung-Mo Ahn, Jiayu Liao, Madeleine M. Fletcher, Dehua Yang, Alastair J. H. Brown, Caihong Zhou, Jiejie Deng, and Ming-Wei Wang. 2016. 'Glucagon-Like Peptide-1 and Its Class B G Protein–Coupled

Receptors: A Long March to Therapeutic Successes', *Pharmacological Reviews*, 68: 954-1013.

De Lean, A., J. M. Stadel, and R. J. Lefkowitz. 1980. 'A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled β-adrenergic receptor', *Journal of Biological Chemistry*, 255: 7108-17.

Deacon, Carolyn F, Michael A Nauck, Maibritt Toft-Nielsen, Lone Pridal, Berend Willms, and Jens J %J Diabetes Holst. 1995. 'Both subcutaneously and intravenously administered glucagon-like peptide I are rapidly degraded from the NH2-terminus in type II diabetic patients and in healthy subjects', 44: 1126-31.

Del Castillo, J., and B. Katz. 1954. 'Quantal components of the end-plate potential', *J Physiol*, 124: 560-73.

Deupi, X., and B. Kobilka. 2007. 'Activation of G protein-coupled receptors', *Adv Protein Chem*, 74: 137-66.

Deupi, Xavier, and Brian K Kobilka. 2010. 'Energy landscapes as a tool to integrate GPCR structure, dynamics, and function', *Physiology*, 25: 293-303.

Devaux, Philippe F., and Roger Morris. 2004. 'Transmembrane Asymmetry and Lateral Domains in Biological Membranes', 5: 241-46.

Díaz, Graciela B., Ana Maria Cortizo, María Elisa Garciía, and Juan José Gagliardino. 1988. 'Lipid composition of normal male rat islets', 23: 1125-28.

Diaz, J. F., B. Wroblowski, J. Schlitter, and Y. Engelborghs. 1997. 'Calculation of pathways for the conformational transition between the GTP- and GDP-bound states of the Ha-ras-p21 protein: calculations with explicit solvent simulations and comparison with calculations in vacuum', *Proteins*, 28: 434-51.

Dibonaventura, M. D., J. S. Wagner, C. J. Girman, K. Brodovicz, Q. Zhang, Y. Qiu, S. R. Pentakota, and L. Radican. 2010. 'Multinational Internet-based survey of patient preference for newer oral or injectable Type 2 diabetes medication', *Patient Prefer Adherence*, 4: 397-406.

Dinner, Aaron R., Andrej Šali, Lorna J. Smith, Christopher M. Dobson, and Martin Karplus. 2000. 'Understanding protein folding via free-energy surfaces from theory and experiment', *Trends Biochem Sci*, 25: 331-39.

Dobolyi, Arpad, Eugene Dimitrov, Miklos Palkovits, and Ted Usdin. 2012. 'The Neuroendocrine Functions of the Parathyroid Hormone 2 Receptor', *Frontiers in Endocrinology*, 3.

Dobson, Christopher M. 1990. 'Hinge-bending and folding', *Nature*, 348: 198-99.

Dods, R. L., and D. Donnelly. 2015. 'The peptide agonist-binding site of the glucagon-like peptide-1 (GLP-1) receptor based on site-directed mutagenesis and knowledge-based modelling', *Biosci Rep*, 36: e00285.

Dong, Maoqing, Giuseppe Deganutti, Sarah J. Piper, Yi-Lynn Liang, Maryam Khoshouei, Matthew J. Belousoff, Kaleeckal G. Harikumar, Christopher A. Reynolds, Alisa Glukhova, Sebastian G. B. Furness, Arthur Christopoulos, Radostin Danev, Denise Wootten, Patrick M. Sexton, and Laurence J. Miller. 2020. 'Structure and dynamics of the active Gs-coupled human secretin receptor', *Nat Commun*, 11: 4137.

Donnelly, D. 2012. 'The structure and function of the glucagon-like peptide-1 receptor and its ligands', *Br J Pharmacol*, 166: 27-41.

Dror, Ron O., Daniel H. Arlow, Paul Maragakis, Thomas J. Mildorf, Albert C. Pan, Huafeng Xu, David W. Borhani, and David E. Shaw. 2011. 'Activation mechanism of the β2-adrenergic receptor', *Proc Natl Acad Sci U S A*, 108: 18684-89.

Drucker, D. J. 2016. 'The Cardiovascular Biology of Glucagon-like Peptide-1', *Cell Metab*, 24: 15-30.

Edward Zhou, X, Karsten Melcher, and H Eric Xu. 2019. 'Structural biology of G protein-coupled receptor signaling complexes', *Protein Science*, 28: 487-501.

Egan, Josephine M, Graydon S Meneilly, Joel F Habener, and Dariush Elahi. 2002. 'Glucagon-like peptide-1 augments insulin-mediated glucose uptake in the obese state', *The Journal of Clinical Endocrinology & Metabolism*, 87: 3768-73.

Ehrenmann, Janosch, Jendrik Schöppe, Christoph Klenk, Mathieu Rappas, Lutz Kummer, Andrew S. Doré, and Andreas Plückthun. 2018. 'High-resolution crystal structure of parathyroid hormone 1 receptor in complex with a peptide agonist', *Nature Structural & Molecular Biology*, 25: 1086-92.

Elber, R., and M. Karplus. 1987. 'Multiple conformational states of proteins: a molecular dynamics analysis of myoglobin', *Science*, 235: 318-21.

Essmann, Ulrich, Lalith Perera, Max L. Berkowitz, Tom Darden, Hsing Lee, and Lee G. Pedersen. 1995. 'A smooth particle mesh Ewald method', 103: 8577-93.

Fagerberg, Linn, Kalle Jonasson, Gunnar von Heijne, Mathias Uhlén, and Lisa Berglund. 2010. 'Prediction of the human membrane proteome', *Proteomics*, 10: 1141-49.

Fan, Guiyong, Qun Zhao, Pei Lu, Hao Chen, Wei Tan, Weixiao Guo, Chaoqun Liu, and Jinlian Liu. 2020. 'Comparison between teriparatide and bisphosphonates for improving bone mineral density in postmenopausal osteoporosis patients: A meta-analysis', *Medicine*, 99: e18964-e64.

Fandiño, J., L. Toba, L. C. González-Matías, Y. Diz-Chaves, and F. Mallo. 2020. 'GLP-1 receptor agonist ameliorates experimental lung fibrosis', *Sci Rep*, 10: 18091.

Farkas, Erzsébet, Anett Szilvásy-Szabó, Yvette Ruska, Richárd Sinkó, Morten Grønbech Rasch, Thomas Egebjerg, Charles Pyke, Balázs Gereben, Lotte Bjerre Knudsen, and Csaba Fekete. 2021. 'Distribution and ultrastructural localization of the glucagon-like peptide-1 receptor (GLP-1R) in the rat brain', *Brain Structure and Function*, 226: 225-45.

Farrens, David L, Christian Altenbach, Ke Yang, Wayne L Hubbell, and H Gobind Khorana. 1996. 'Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin', *Science*, 274: 768-70.

Fehse, F., M. Trautmann, J. J. Holst, A. E. Halseth, N. Nanayakkara, L. L. Nielsen, M. S. Fineman, D. D. Kim, and M. A. Nauck. 2005. 'Exenatide augments first- and second-phase insulin secretion in response to intravenous glucose in subjects with type 2 diabetes', *J Clin Endocrinol Metab*, 90: 5991-7.

Ferguson, S. S. 2001. 'Evolving concepts in G protein-coupled receptor endocytosis: the role in receptor desensitization and signaling', *Pharmacological Reviews*, 53: 1-24.

Ferrannini, Giulia, Thomas Hach, Susanne Crowe, Arjun Sanghvi, Kevin D. Hall, and Ele Ferrannini. 2015. 'Energy Balance After Sodium–Glucose Cotransporter 2 Inhibition', 38: 1730-35.

Ferrara, Philippe, Joannis Apostolakis, and Amedeo Caflisch. 2000. 'Targeted Molecular Dynamics Simulations of Protein Unfolding', *The Journal of Physical Chemistry B*, 104: 4511-18.

Fleetwood O., Ma§tricon P., Carlsson J., Delemotte L. 2020. ‘Enrgy landscapes reveal agonist control of GPCR activation via microswitches’, *Biochemistry,* 59,7 880-891.

Frauenfelder, H, SG Sligar, and PG Wolynes. 1991. 'The energy landscapes and motions of proteins', *Science*, 254: 1598-603.

Fredriksson, R., M. C. Lagerstrom, L. G. Lundin, and H. B. Schioth. 2003. 'The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints', *Mol Pharmacol*, 63: 1256-72.

Frías J. P., Davies M.J., Rosenstock J., Pérez F.C., Fernández L., Bergman B.K., Liu B., Cui X., Brown K., 2021 'Tirzepatide Versus Semaglutide Once Weekly in Patients With Type 2 Diabetes', *N Engl J Med,* 385:503-15.

Gabilondo, A. M., C. Krasel, and M. J. Lohse. 1996. 'Mutations of Tyr326 in the beta 2-adrenoceptor disrupt multiple receptor functions', *Eur J Pharmacol*, 307: 243-50.

Gallwitz, B, WE Schmidt, JM Conlon, and W Creutzfeldt. 1990. 'Glucagon-like peptide-1 (7–36) amide: characterization of the domain responsible for binding to its receptor on rat insulinoma RINm5F cells', *Journal of molecular endocrinology*, 5: 33-39.

Gallwitz, Baptist, Maike Witt, Gabriele Paetzold, Corinna Morys-Wortmann, Bodo Zimmermann, Klaus Eckart, Ulrich R Fölsch, and Wolfgang E Schmidt. 1994. 'Structure/activity characterization of glucagon-like peptide-1', *Eur J Biochem*, 225: 1151-56.

Garant, Montrose-Rafizadeh C Avdonin P, MJ Rodgers BD Kole S Yang, and H Levine MA Schwindinger W Bernier. 'M 1999 Pancreatic glucagon-like peptide-1 receptor couples to multiple G proteins and activates mitogen-activated protein kinase pathways in Chinese hamster ovary cells', *Endocrinology*, 140: 1132-40.

García-Nafría, J., R. Nehmé, P. C. Edwards, and C. G. Tate. 2018. 'Cryo-EM structure of the serotonin 5-HT(1B) receptor coupled to heterotrimeric G(o)', *Nature*, 558: 620-23.

García-Nafría, J., and C. G. Tate. 2019. 'Cryo-EM structures of GPCRs coupled to G(s), G(i) and G(o)', *Molecular and cellular endocrinology*, 488: 1-13.

Gardella, THOMAS J, H Jüppner, ANDREW K Wilson, HENRY T Keutmann, AB Abou-Samra, GINO V Segre, F RICHARD Bringhurst, JOHN T Potts Jr, SAMUEL R Nussbaum, and HENRY M Kronenberg. 1994. 'Determinants of [Arg2] PTH-(1-34) binding and signaling in the transmembrane region of the parathyroid hormone receptor', *Endocrinology*, 135: 1186-94.

Gardella, Thomas J, Michael D Luck, Ming-Hui Fan, and ChenWei Lee. 1996. 'Transmembrane residues of the parathyroid hormone (PTH)/PTH-related peptide receptor that specifically affect binding and signaling by agonist ligands', *Journal of Biological Chemistry*, 271: 12820-25.

Gether, Ulrik. 2000. 'Uncovering Molecular Mechanisms Involved in Activation of G Protein-Coupled Receptors', *Endocrine Reviews*, 21: 90-113.

GBD 2013 Mortality and Causes of Death Collaborators. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2015;385:117–171.

Gogala, M., T. Becker, B. Beatrix, J. P. Armache, C. Barrio-Garcia, O. Berninghausen, and R. Beckmann. 2014. 'Structures of the Sec61 complex engaged in nascent peptide translocation or membrane insertion', *Nature*, 506: 107-10.

Göke, R., H. C. Fehmann, T. Linn, H. Schmidt, M. Krause, J. Eng, and B. Göke. 1993. 'Exendin-4 is a high potency agonist and truncated exendin-(9-39)-amide an

antagonist at the glucagon-like peptide 1-(7-36)-amide receptor of insulin-secreting beta-cells', *J Biol Chem*, 268: 19650-5.

Gómez Santiago, Carla, Emanuele Paci, and Dan Donnelly. 2018. 'A mechanism for agonist activation of the glucagon-like peptide-1 (GLP-1) receptor through modelling & molecular dynamics', *Biochem Biophys Res Commun*, 498: 359-65. DOI: 10.1016/j.bbrc.2018.01.110.

Górriz, José Luis, María José Soler, Juan F. Navarro-González, Clara García-Carro, María Jesús Puchades, Luis D'Marco, Alberto Martínez Castelao, Beatriz Fernández-Fernández, Alberto Ortiz, Carmen Górriz-Zambrano, Jorge Navarro-Pérez, and Juan José Gorgojo-Martinez. 2020. 'GLP-1 Receptor Agonists and Diabetic Kidney Disease: A Call of Attention to Nephrologists', *Journal of clinical medicine*, 9: 947.

Grace, Christy Rani R., Marilyn H. Perrin, Jozsef Gulyas, Jean E. Rivier, Wylie W. Vale, and Roland Riek. 2010. 'NMR Structure of the First Extracellular Domain of Corticotropin-releasing Factor Receptor 1 (ECD1-CRF-R1) Complexed with a High Affinity Agonist', *J Biol Chem*, 285: 38580-89.

Grant, B. J., A. P. Rodrigues, K. M. ElSawy, J. A. McCammon, and L. S. Caves. 2006. 'Bio3d: an R package for the comparative analysis of protein structures', *Bioinformatics*, 22: 2695-6.

Green, Brian D, Peter R %J Best Practice Flatt, Research Clinical Endocrinology, and Metabolism. 2007. 'Incretin hormone mimetics and analogues in diabetes therapeutics', 21: 497-516.

Gregorio, G. G., M. Masureel, D. Hilger, D. S. Terry, M. Juette, H. Zhao, Z. Zhou, J. M. Perez-Aguilar, M. Hauge, S. Mathiasen, J. A. Javitch, H. Weinstein, B. K. Kobilka, and S. C. Blanchard. 2017. 'Single-molecule analysis of ligand efficacy in beta2AR-G-protein activation', *Nature*, 547: 68-73.

Griffith, David A., David J. Edmonds, Jean-Phillipe Fortin, Amit S. Kalgutkar, J. Brent Kuzmiski, Paula M. Loria, Aditi R. Saxena, Scott W. Bagley, Clare Buckeridge, John M. Curto, David R. Derksen, João M. Dias, Matthew C. Griffor, Seungil Han, V. Margaret Jackson, Margaret S. Landis, Daniel J. Lettiere, Chris Limberakis, Yuhang Liu, Alan M. Mathiowetz, David W. Piotrowski, David A. Price, Roger B. Ruggeri, and David A. Tess. 2020. 'A small-molecule oral agonist of the human glucagon-like peptide-1 receptor', *bioRxiv*: 2020.09.29.319483.

Theoretical and Computational Biophysics Group. 2016 'Data Analysis in VMD', Beckman Institute for Advanced Science and Technology
University of Illinois at Urbana-Champaign
405 N. Mathews
Urbana, IL 61801 Accessed 20.03.2019. https://www.ks.uiuc.edu/Training/Tutorials/vmd/tutorial-html/node7.html.

Grundmann, Manuel, Nicole Merten, Davide Malfacini, Asuka Inoue, Philip Preis, Katharina Simon, Nelly Rüttiger, Nicole Ziegler, Tobias Benkel, and Nina Katharina Schmitt. 2018. 'Lack of beta-arrestin signaling in the absence of active G proteins', *Nat Commun*, 9: 1-16.

Guex, Nicolas, Manuel C. Peitsch, and Torsten Schwede. 2009. 'Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective', 30: S162-S73.

"Guidance. Health matters: obesity and the food environment." In. 2017. edited by Public Health England.

Guixà-González, Ramon, José L Albasanz, Ismael Rodriguez-Espigares, Manuel Pastor, Ferran Sanz, Maria Martí-Solano, Moutusi Manna, Hector Martinez-

Seara, Peter W Hildebrand, and Mairena %J Nature communications Martín. 2017. 'Membrane cholesterol access into a G-protein-coupled receptor', 8: 14505.

Gurevich, Vsevolod V, and Eugenia V Gurevich. 2004. 'The molecular acrobatics of arrestin activation', *Trends Pharmacol Sci*, 25: 105-11.

Gurevich, Vsevolod V., and Eugenia V. Gurevich. 2019. 'GPCR Signaling Regulation: The Role of GRKs and Arrestins', 10.

Guvench, Olgun, and Alexander D. MacKerell. 2008. 'Comparison of Protein Force Fields for Molecular Dynamics Simulations.' in Andreas Kukol (ed.), *Molecular Modeling of Proteins* (Humana Press: Totowa, NJ).

Hackenbrock C.R., Schneider H., Lemasters J.J., Höchli M. 1980. *Relationships Between Bilayer Lipid, Motional Freedom of Oxidoreductase Components, and Electron Transfer in the Mitochondrial Inner Membrane.* (Springer, Boston, MA).

Hagren, Olof Idevall, and Anders %J Journal of Biological Chemistry Tengholm. 2006. 'Glucose and insulin synergistically activate phosphatidylinositol 3-kinase to trigger oscillations of phosphatidylinositol 3, 4, 5-trisphosphate in β-cells', 281: 39121-27.

Hallberg, A. 1984a. 'Effects of starvation and different culture conditions on the phospholipid content of isolated pancreatic islets', *Biochimica et Biophysica Acta (BBA)/Lipids and Lipid Metabolism*, 796: 328-35.

Hallberg, Anders. 1984b. 'Effects of starvation and different culture conditions on the phospholipid content of isolated pancreatic islets', *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism*, 796: 328-35.

Hallbrink, M., T. Holmqvist, M. Olsson, C. G. Ostenson, S. Efendic, and U. Langel. 2001. 'Different domains in the third intracellular loop of the GLP-1 receptor are responsible for Galpha(s) and Galpha(i)/Galpha(o) activation', *Biochim Biophys Acta*, 1546: 79-86.

Hanson, M. A., V. Cherezov, M. T. Griffith, C. B. Roth, V. P. Jaakola, E. Y. Chien, J. Velasquez, P. Kuhn, and R. C. Stevens. 2008. 'A specific cholesterol binding site is established by the 2.8 A structure of the human beta2-adrenergic receptor', *Structure*, 16: 897-905.

Hare, K. J., T. Vilsboll, M. Asmar, C. F. Deacon, F. K. Knop, and J. J. Holst. 2010. 'The glucagonostatic and insulinotropic effects of glucagon-like peptide 1 contribute equally to its glucose-lowering action', *Diabetes*, 59: 1765-70.

Hattersley, G., T. Dean, B. A. Corbin, H. Bahar, and T. J. Gardella. 2016. 'Binding Selectivity of Abaloparatide for PTH-Type-1-Receptor Conformations and Effects on Downstream Signaling', *Endocrinology*, 157: 141-9.

Hauser, Alexander S., Sreenivas Chavali, Ikuo Masuho, Leonie J. Jahn, Kirill A. Martemyanov, David E. Gloriam, and M. Madan Babu. 2018. 'Pharmacogenomics of GPCR Drug Targets', *Cell*, 172: 41-54.e19.

Hauser A.S., Attwood M.M., Rask-Andersen M., Schioth H.B., Gloriam D.E. Trends in GPCR drug discovery: New agents, targets and indications. *Nat. Rev. Drug Discov.* 2017;16:829–842. doi: 10.1038/nrd.2017.178.

Hay, D. L., and A. A. Pioszak. 2016. 'Receptor Activity-Modifying Proteins (RAMPs): New Insights and Roles', *Annu Rev Pharmacol Toxicol*, 56: 469-87.

He, Ziwei, Fabian Paul, and Benoît Roux. 2021. 'A critical perspective on Markov state model treatments of protein–protein association using coarse-grained simulations', *The Journal of Chemical Physics*, 154: 084101.

Hedger, George, and Mark SP %J Biochimica et Biophysica Acta -Biomembranes Sansom. 2016. 'Lipid interaction sites on channels, transporters and receptors: recent insights from molecular dynamics simulations', 1858: 2390-400.

Helfrich, W. 1973. 'Elastic properties of lipid bilayers: theory and possible experiments', *Z Naturforsch C*, 28: 693-703.

Henzler-Wildman, K., and D. Kern. 2007. 'Dynamic personalities of proteins', *Nature*, 450: 964-72.

Hess, B., C. Kutzner, D. van der Spoel, and E. Lindahl. 2008. 'GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation', *J Chem Theory Comput*, 4: 435-47.

Hesse, Eric, Saskia Schröder, Diana Brandt, Jenny Pamperin, Hiroaki Saito, and Hanna Taipaleenmäki. 2019. 'Sclerostin inhibition alleviates breast cancer-induced bone metastases and muscle weakness', *JCI insight*, 5: e125543.

Hilger, Daniel, Kaavya Krishna Kumar, Hongli Hu, Mie Fabricius Pedersen, Lise Giehm, Jesper Mosolff Mathiesen, Georgios Skiniotis, and Brian K. Kobilka. 2019. 'Structural insights into ligand efficacy and activation of the glucagon receptor', *bioRxiv*: 660837.

Hilger, Daniel, Kaavya Krishna Kumar, Hongli Hu, Mie Fabricius Pedersen, Evan S O'Brien, Lise Giehm, Christine Jennings, Gözde Eskici, Asuka Inoue, and Michael Lerch. 2020. 'Structural insights into differences in G protein activation by family A and family B GPCRs', *Science*, 369.

Hilger, Daniel, Kaavya Krishna Kumar, Hongli Hu, Mie Fabricius Pedersen, Evan S. O'Brien, Lise Giehm, Christine Jennings, Gözde Eskici, Asuka Inoue, Michael Lerch, Jesper Mosolff Mathiesen, Georgios Skiniotis, and Brian K. Kobilka. 2020. 'Structural insights into differences in G protein activation by family A and family B GPCRs', *Science*, 369: eaba3373.

Hoang Do, Oanh, and Peter Thorn. 2015. 'Insulin secretion from beta cells within intact islets: location matters', *Clinical and experimental pharmacology & physiology*, 42: 406-14.

Hoare, S. R. 2005. 'Mechanisms of peptide and nonpeptide ligand binding to Class B G-protein-coupled receptors', *Drug Discov Today*, 10: 417-27.

Hoare, Sam R. J., Thomas J. Gardella, and Ted B. Usdin. 2001. 'Evaluating the Signal Transduction Mechanism of the Parathyroid Hormone 1 Receptor: EFFECT OF RECEPTOR-G-PROTEIN INTERACTION ON THE LIGAND BINDING MECHANISM AND RECEPTOR CONFORMATION', *Journal of Biological Chemistry*, 276: 7741-53.

Hollenstein, K., C. de Graaf, A. Bortolato, M. W. Wang, F. H. Marshall, and R. C. Stevens. 2014. 'Insights into the structure of class B GPCRs', *Trends Pharmacol Sci*, 35: 12-22.

Hollenstein, K., J. Kean, A. Bortolato, R. K. Cheng, A. S. Dore, A. Jazayeri, R. M. Cooke, M. Weir, and F. H. Marshall. 2013. 'Structure of class B GPCR corticotropin-releasing factor receptor 1', *Nature*, 499: 438-43.

Holst, J. J., C. Orskov, O. V. Nielsen, and T. W. Schwartz. 1987. 'Truncated glucagon-like peptide I, an insulin-releasing hormone from the distal gut', *FEBS Lett*, 211: 169-74.

Holst, Jens Juul. 2019. 'From the Incretin Concept and the Discovery of GLP-1 to Today's Diabetes Therapy', 10.

Holz, G. G., and O. G. Chepurny. 2005. 'Diabetes outfoxed by GLP-1?', *Sci STKE*, 2005: pe2.

Huang, J., and A. D. MacKerell, Jr. 2013. 'CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data', *J Comput Chem*, 34: 2135-45.

Humphrey, William, Andrew Dalke, and Klaus Schulten. 1996. 'VMD: Visual molecular dynamics', *J Mol Graph*, 14: 33-38.

Hupe-Sodmann, Karin, Gerard Patrick McGregor, Robert Bridenbaugh, Rüdiger Göke, Burkhard Göke, Hubert Thole, Bodo Zimmermann, and Karlheinz %J Regulatory peptides Voigt. 1995. 'Characterisation of the processing by human neutral endopeptidase 24.11 of GLP-1 (7–36) amide and comparison of the substrate specificity of the enzyme for other glucagon-like peptides', 58: 149-56.

IDF. 2017. "IDF Diabetes Atlas Eight Edition." In, 147. International Diabetes Federation.

Iida-Klein, Akiko, Jun Guo, Masahiko Takemura, Matthew T Drake, John T Potts, Abdul Abou-Samra, F Richard Bringhurst, and Gino V Segre. 1997. 'Mutations in the second cytoplasmic loop of the rat parathyroid hormone (PTH)/PTH-related protein receptor result in selective loss of PTH-stimulated phospholipase C activity', *Journal of Biological Chemistry*, 272: 6882-89.

Inooka, H., T. Ohtaki, O. Kitahara, T. Ikegami, S. Endo, C. Kitada, K. Ogi, H. Onda, M. Fujino, and M. Shirakawa. 2001. 'Conformation of a peptide ligand bound to its G-protein coupled receptor', *Nat Struct Biol*, 8: 161-5.

Institute, UniProt European Bioinformatics. 2019. " UniProtKB/TrEMBL PROTEIN DATABASE RELEASE 2019_08 STATISTICS." In.

Iorga, Roua Anamaria, Nicolae Bacalbasa, Mara Carsote, Ovidiu Gabriel Bratu, Ana Maria Alexandra Stanescu, Simona Bungau, Carmen Pantis, and Camelia Cristina Diaconu. 2020. 'Metabolic and cardiovascular benefits of GLP-1 agonists, besides the hypoglycemic effect (Review)', *Experimental and therapeutic medicine*, 20: 2396-400.

Jabeen, Amara, Abidali Mohamedali, and Shoba Ranganathan. 2019. 'Protocol for Protein Structure Modelling.' in Shoba Ranganathan, Michael Gribskov, Kenta Nakai and Christian Schönbach (eds.), *Encyclopedia of Bioinformatics and Computational Biology* (Academic Press: Oxford).

Jackowski, S. 1994. 'Coordination of membrane phospholipid synthesis with the cell cycle', *J Biol Chem*, 269: 3858-67.

———. 1996. 'Cell cycle regulation of membrane phospholipid metabolism', *J Biol Chem*, 271: 20219-22.

Jackson, Sylvia H., Tonya S. Martin, Jocelyn D. Jones, David Seal, and Frank Emanuel. 2010. 'Liraglutide (victoza): the first once-daily incretin mimetic injection for type-2 diabetes', *P & T : a peer-reviewed journal for formulary management*, 35: 498-529.

Jarpe, Matthew B., Cindy Knall, Fiona M. Mitchell, Anne Mette Buhl, Emir Duzic, and Gary L. Johnson. 1998. '[d-Arg1,d-Phe5,d-Trp7,9,Leu11]Substance P Acts as a Biased Agonist toward Neuropeptide and Chemokine Receptors', *Journal of Biological Chemistry*, 273: 3097-104.

Jazayeri, Ali, Andrew S. Doré, Daniel Lamb, Harini Krishnamurthy, Stacey M. Southall, Asma H. Baig, Andrea Bortolato, Markus Koglin, Nathan J. Robertson, James C. Errey, Stephen P. Andrews, Iryna Teobald, Alastair J. H. Brown, Robert M. Cooke, Malcolm Weir, and Fiona H. Marshall. 2016. 'Extra-helical binding site of a glucagon receptor antagonist', *Nature*, 533: 274.

Jazayeri, Ali, Mathieu Rappas, Alastair J. H. Brown, James Kean, James C. Errey, Nathan J. Robertson, Cédric Fiez-Vandal, Stephen P. Andrews, Miles Congreve, Andrea Bortolato, Jonathan S. Mason, Asma H. Baig, Iryna Teobald, Andrew S. Doré, Malcolm Weir, Robert M. Cooke, and Fiona H. Marshall. 2017. 'Crystal structure of the GLP-1 receptor bound to a peptide agonist', *Nature*, 546: 254.

Ji, Tae H, William J Murdoch, and Inhae Ji. 1995. 'Activation of membrane receptors', *Endocrine*, 3: 187-94.

Ji, Tae H., Mathis Grossmann, and Inhae Ji. 1998. 'G Protein-coupled Receptors: I. DIVERSITY OF RECEPTOR-LIGAND INTERACTIONS', *Journal of Biological Chemistry*, 273: 17299-302.

Jo, Sunhwan, Taehoon Kim, and Wonpil Im. 2007. 'Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations', *PLOS ONE*, 2: e880.

Jo, Sunhwan, Taehoon Kim, Vidyashankara G. Iyer, and Wonpil Im. 2008. 'CHARMM-GUI: A web-based graphical user interface for CHARMM', 29: 1859-65.

Jo, Sunhwan, Joseph B. Lim, Jeffery B. Klauda, and Wonpil Im. 2009. 'CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes', *Biophys J*, 97: 50-58.

Jolley, C. C., S. A. Wells, P. Fromme, and M. F. Thorpe. 2008. 'Fitting low-resolution cryo-EM maps of proteins using constrained geometric simulations', *Biophys J*, 94: 1613-21.

Jones, Gareth, Peter Willett, Robert C. Glen, Andrew R. Leach, and Robin Taylor. 1997. 'Development and validation of a genetic algorithm for flexible docking11Edited by F. E. Cohen', *J Mol Biol*, 267: 727-48.

Jorgensen, R., V. Kubale, M. Vrecl, T. W. Schwartz, and C. E. Elling. 2007. 'Oxyntomodulin differentially affects glucagon-like peptide-1 receptor beta-arrestin recruitment and signaling through Galpha(s)', *J Pharmacol Exp Ther*, 322: 148-54.

Juppner, Harald, Abdul-Badi Abou-Samra, Mason Freeman, Xiang F Kong, Ernestina Schipani, Jennifer Richards, Lee F Kolakowski, Janet Hock, John T Potts, and Henry M Kronenberg. 1991. 'AG protein-linked receptor for parathyroid hormone and parathyroid hormone-related peptide', *Science*, 254: 1024-26.

Kahn, Steven E., Mark E. Cooper, and Stefano Del Prato. 2014. 'Pathophysiology and treatment of type 2 diabetes: perspectives on the past, present, and future', *Lancet*, 383: 1068-83.

Kang, Yanyong, Oleg Kuybeda, Parker W. de Waal, Somnath Mukherjee, Ned Van Eps, Przemyslaw Dutka, X. Edward Zhou, Alberto Bartesaghi, Satchal Erramilli, Takefumi Morizumi, Xin Gu, Yanting Yin, Ping Liu, Yi Jiang, Xing Meng, Gongpu Zhao, Karsten Melcher, Oliver P. Ernst, Anthony A. Kossiakoff, Sriram Subramaniam, and H. Eric Xu. 2018. 'Cryo-EM structure of human rhodopsin bound to an inhibitory G protein', *Nature*, 558: 553-58.

Kang, Yanyong, X. Edward Zhou, Xiang Gao, Yuanzheng He, Wei Liu, Andrii Ishchenko, Anton Barty, Thomas A. White, Oleksandr Yefanov, Gye Won Han, Qingping Xu, Parker W. de Waal, Jiyuan Ke, M. H. Eileen Tan, Chenghai Zhang, Arne Moeller, Graham M. West, Bruce D. Pascal, Ned Van Eps, Lydia N. Caro, Sergey A. Vishnivetskiy, Regina J. Lee, Kelly M. Suino-Powell, Xin Gu, Kuntal Pal, Jinming Ma, Xiaoyong Zhi, Sébastien Boutet, Garth J. Williams, Marc Messerschmidt, Cornelius Gati, Nadia A. Zatsepin, Dingjie Wang, Daniel James, Shibom Basu, Shatabdi Roy-Chowdhury,

Chelsie E. Conrad, Jesse Coe, Haiguang Liu, Stella Lisova, Christopher Kupitz, Ingo Grotjohann, Raimund Fromme, Yi Jiang, Minjia Tan, Huaiyu Yang, Jun Li, Meitian Wang, Zhong Zheng, Dianfan Li, Nicole Howe, Yingming Zhao, Jörg Standfuss, Kay Diederichs, Yuhui Dong, Clinton S. Potter, Bridget Carragher, Martin Caffrey, Hualiang Jiang, Henry N. Chapman, John C. H. Spence, Petra Fromme, Uwe Weierstall, Oliver P. Ernst, Vsevolod Katritch, Vsevolod V. Gurevich, Patrick R. Griffin, Wayne L. Hubbell, Raymond C. Stevens, Vadim Cherezov, Karsten Melcher, and H. Eric Xu. 2015. 'Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser', *Nature*, 523: 561.

Kaouane, Nadia, Yves Porte, Monique Vallée, Laurent Brayda-Bruno, Nicole Mons, Ludovic Calandreau, Aline Marighetto, Pier Vincenzo Piazza, and Aline Desmedt. 2012. 'Glucocorticoids can induce PTSD-like memory impairments in mice', *Science*, 335: 1510-13.

Karplus, Martin, and John Kuriyan. 2005. 'Molecular dynamics and protein function', *Proceedings of the National Academy of Sciences*, 102: 6679-85.

Kavishwar, Amol, and Anna Moore. 2013. 'Sphingomyelin patches on pancreatic beta-cells are indicative of insulin secretory capacity', *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society*, 61: 910-19.

Kawai T, Sun B, Yoshino H, Feng D, Suzuki Y, Fukazawa M, Nagao S, Wainscott DB, Showalter AD, Droz BA, Kobilka TS, Coghlan MP, Willard FS, Kawabe Y, Kobilka BK, Sloop KW. 2020 'Structural basis for GLP-1 receptor activation by LY3502970, an orally active nonpeptide agonist', Proc Natl Acad Sci USA. 24;117(47):29959-29967.

Kemp, Thomas M, Elizabeth LM Barr, Paul Z Zimmet, Adrian J Cameron, Timothy A Welborn, Stephen Colagiuri, Patrick Phillips, and Jonathan E %J Diabetes care Shaw. 2005. 'Glucose, lipid, and blood pressure control in Australian adults with type 2 diabetes: the 1999–2000 AusDiab', 28: 1490-92.

Kenakin, T. 2004. 'Principles: receptor theory in pharmacology', *Trends Pharmacol Sci*, 25: 186-92.

Kenakin T., Morgan P., Lutz M., Weiss J. 2000. 'The Evolution of Drug-Receptor Models: The Cubic Ternary Complex Model for G Protein-Coupled Receptors.' in, *The Pharmacology of Functional, Biochemical, and Recombinant Receptor Systems. Handbook of Experimental Pharmacology* (Springer, Berlin, Heidelberg: Berlin).

Kendler, David L., Fernando Marin, Cristiano A. F. Zerbini, Luis A. Russo, Susan L. Greenspan, Vit Zikan, Alicia Bagur, Jorge Malouf-Sierra, Péter Lakatos, Astrid Fahrleitner-Pammer, Eric Lespessailles, Salvatore Minisola, Jean Jacques Body, Piet Geusens, Rüdiger Möricke, and Pedro López-Romero. 2018. 'Effects of teriparatide and risedronate on new fractures in post-menopausal women with severe osteoporosis (VERO): a multicentre, double-blind, double-dummy, randomised controlled trial', *The Lancet*, 391: 230-40.

Key, T Alexander, Terry D Foutz, Vsevolod V Gurevich, Larry A Sklar, and Eric R Prossnitz. 2003. 'N-formyl peptide receptor phosphorylation domains differentially regulate arrestin and agonist affinity', *Journal of Biological Chemistry*, 278: 4041-47.

Kim, Tae Hun, Ka Young Chung, Aashish Manglik, Ron O Dror, David E Shaw, Brian K Kobilka, and R Scott Prosser. 2013. "THE ROLE OF LIGANDS ON THE THERMAL EQUILIBRIA BETWEEN FUNCTIONAL STATES OF AG PROTEIN-COUPLED RECEPTOR." In *JOURNAL OF RECEPTORS*

AND SIGNAL TRANSDUCTION, 193-93. INFORMA HEALTHCARE TELEPHONE HOUSE, 69-77 PAUL STREET, LONDON EC2A 4LQ, ENGLAND.

Kiviranta, Päivi H., Heikki S. Salo, Jukka Leppänen, Valtteri M. Rinne, Sergiy Kyrylenko, Erkki Kuusisto, Tiina Suuronen, Antero Salminen, Antti Poso, Maija Lahtela-Kakkonen, and Erik A. A. Wallén. 2008. 'Characterization of the binding properties of SIRT2 inhibitors with a N-(3-phenylpropenoyl)-glycine tryptamide backbone', *Bioorganic & medicinal chemistry*, 16: 8054-62.

Klenk, C., S. Schulz, D. Calebiro, and M. J. Lohse. 2010. 'Agonist-regulated cleavage of the extracellular domain of parathyroid hormone receptor type 1', *J Biol Chem*, 285: 8665-74.

Kleuss, C., A. S. Raw, E. Lee, S. R. Sprang, and A. G. Gilman. 1994. 'Mechanism of GTP hydrolysis by G-protein alpha subunits', *Proc Natl Acad Sci U S A*, 91: 9828-31.

Knudsen, L. B., D. Kiel, M. Teng, C. Behrens, D. Bhumralkar, J. T. Kodra, J. J. Holst, C. B. Jeppesen, M. D. Johnson, J. C. de Jong, A. S. Jorgensen, T. Kercher, J. Kostrowicki, P. Madsen, P. H. Olesen, J. S. Petersen, F. Poulsen, U. G. Sidelmann, J. Sturis, L. Truesdale, J. May, and J. Lau. 2007. 'Small-molecule agonists for the glucagon-like peptide 1 receptor', *Proc Natl Acad Sci U S A*, 104: 937-42.

Kobilka, Brian K. 2007. 'G protein coupled receptor structure and activation', *Biochim Biophys Acta*, 1768: 794-807.

Koehl, Antoine, Hongli Hu, Shoji Maeda, Yan Zhang, Qianhui Qu, Joseph M. Paggi, Naomi R. Latorraca, Daniel Hilger, Roger Dawson, Hugues Matile, Gebhard F. X. Schertler, Sebastien Granier, William I. Weis, Ron O. Dror, Aashish Manglik, Georgios Skiniotis, and Brian K. Kobilka. 2018. 'Structure of the μ-opioid receptor–Gi protein complex', *Nature*, 558: 547-52.

Kolakowski, L. F., Jr. 1994. 'GCRDb: a G-protein-coupled receptor database', *Receptors Channels*, 2: 1-7.

Koldso, H., D. Shorthouse, J. Helie, and M. S. Sansom. 2014. 'Lipid clustering correlates with membrane curvature as revealed by molecular simulations of complex lipid bilayers', *PLoS Comput Biol*, 10: e1003911.

Koole, Cassandra, Denise Wootten, John Simms, Celine Valant, Rohan Sridhar, Owen L Woodman, Laurence J Miller, Roger J Summers, Arthur Christopoulos, and Patrick M Sexton. 2010. 'Allosteric ligands of the glucagon-like peptide 1 receptor (GLP-1R) differentially modulate endogenous and exogenous peptide responses in a pathway-selective manner: implications for drug screening', *Mol Pharmacol*, 78: 456-65.

Kreymann, B., G. Williams, M. A. Ghatei, and S. R. Bloom. 1987. 'Glucagon-like peptide-1 7-36: a physiological incretin in man', *Lancet*, 2: 1300-4.

Krishna Kumar, Kaavya, Moran Shalev-Benami, Michael J. Robertson, Hongli Hu, Samuel D. Banister, Scott A. Hollingsworth, Naomi R. Latorraca, Hideaki E. Kato, Daniel Hilger, Shoji Maeda, William I. Weis, David L. Farrens, Ron O. Dror, Sanjay V. Malhotra, Brian K. Kobilka, and Georgios Skiniotis. 2019. 'Structure of a Signaling Cannabinoid Receptor 1-G Protein Complex', *Cell*, 176: 448-58.e12.

Krupnick, J. G., and J. L. Benovic. 1998. 'The role of receptor kinases and arrestins in G protein-coupled receptor regulation', *Annu Rev Pharmacol Toxicol*, 38: 289-319.

Kumar, Shiva, Augen Pioszak, Chenghai Zhang, Kunchithapadam Swaminathan, and H. Eric Xu. 2011. 'Crystal Structure of the PAC1R Extracellular Domain Unifies a Consensus Fold for Hormone Recognition by Class B G-Protein Coupled Receptors', *PLOS ONE*, 6: e19682.

Kumari, Punita, Ashish Srivastava, Eshan Ghosh, Ravi Ranjan, Shalini Dogra, Prem N Yadav, and Arun K Shukla. 2017. 'Core engagement with β-arrestin is dispensable for agonist-induced vasopressin receptor endocytosis and ERK activation', *Molecular biology of the cell*, 28: 1003-10.

Kunji, E. R., M. Harding, P. J. Butler, and P. Akamine. 2008. 'Determination of the molecular mass and dimensions of membrane proteins by size exclusion chromatography', *Methods*, 46: 62-72.

Laing, Susan P, AJ Swerdlow, SD Slater, AC Burden, A Morris, Norman Robert Waugh, W Gatling, PJ Bingley, and CC %J Diabetologia Patterson. 2003. 'Mortality from heart disease in a cohort of 23,000 patients with insulin-treated diabetes', 46: 760-65.

Lanske, B., A. C. Karaplis, K. Lee, A. Luz, A. Vortkamp, A. Pirro, M. Karperien, L. H. Defize, C. Ho, R. C. Mulligan, A. B. Abou-Samra, H. Jüppner, G. V. Segre, and H. M. Kronenberg. 1996. 'PTH/PTHrP receptor in early development and Indian hedgehog-regulated bone growth', *Science*, 273: 663-6.

Laporte, S. A., R. H. Oakley, J. Zhang, J. A. Holt, S. S. G. Ferguson, M. G. Caron, and L. S. Barak. 1999. 'The β2-adrenergic receptor/βarrestin complex recruits the clathrin adaptor AP-2 during endocytosis', *Proc Natl Acad Sci U S A*, 96: 3712-17.

Lebon, Guillaume. 2020. 'Fine-tuning receptor–G protein activation and signaling', *Science*, 369: 507-08.

Lee, C., T. J. Gardella, A. B. Abou-Samra, S. R. Nussbaum, G. V. Segre, Jr J. T. Potts, H. M. Kronenberg, and H. Jüppner. 1994. 'Role of the extracellular regions of the parathyroid hormone (PTH)/PTH-related peptide receptor in hormone binding', *Endocrinology*, 135: 1488-95.

Lee, ChenWei, Michael D Luck, Harald Jüppner, John T Potts Jr, Henry M Kronenberg, and Thomas J Gardella. 1995. 'Homolog-scanning mutagenesis of the parathyroid hormone (PTH) receptor reveals PTH-(1-34) binding determinants in the third extracellular loop', *Molecular Endocrinology*, 9: 1269-78.

Lee, Jumin, Xi Cheng, Jason M. Swails, Min Sun Yeom, Peter K. Eastman, Justin A. Lemkul, Shuai Wei, Joshua Buckner, Jong Cheol Jeong, Yifei Qi, Sunhwan Jo, Vijay S. Pande, David A. Case, Charles L. Brooks, Alexander D. MacKerell, Jeffery B. Klauda, and Wonpil Im. 2016. 'CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field', *J Chem Theory Comput*, 12: 405-13.

Lee, Jumin, Dhilon S. Patel, Jonas Ståhle, Sang-Jun Park, Nathan R. Kern, Seonghoon Kim, Joonseong Lee, Xi Cheng, Miguel A. Valvano, Otto Holst, Yuriy A. Knirel, Yifei Qi, Sunhwan Jo, Jeffery B. Klauda, Göran Widmalm, and Wonpil Im. 2019. 'CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans', *J Chem Theory Comput*, 15: 775-86.

Lee, Ray X., Greg J. Stephens, and Bernd Kuhn. 2018. 'Affective bonding explains post-traumatic behavioral development in adult mice', *bioRxiv*: 249870.

Leff, P. 1995. 'The two-state model of receptor activation', *Trends Pharmacol Sci*, 16: 89-97.

Lefkowitz, Robert J. 1998. 'G Protein-coupled Receptors: III. NEW ROLES FOR RECEPTOR KINASES AND β-ARRESTINS IN RECEPTOR SIGNALING AND DESENSITIZATION', *Journal of Biological Chemistry*, 273: 18677-80.

Lei, Saifei, Lachlan Clydesdale, Antao Dai, Xiaoqing Cai, Yang Feng, Dehua Yang, Yi-Lynn Liang, Cassandra Koole, Peishen Zhao, Thomas Coudrat, Arthur Christopoulos, Ming-Wei Wang, Denise Wootten, and Patrick M. Sexton. 2018. 'Two distinct domains of the glucagon-like peptide-1 receptor control peptide-mediated biased agonism', 293: 9370-87.

Lenstra, J. A., J. Hofsteenge, and J. J. Beintema. 1977. 'Invariant features of the structure of pancreatic ribonuclease. A test of different predictive models', *J Mol Biol*, 109: 185-93.

Leon, Nicholas, Richard LaCoursiere, Deirdre Yarosh, and Roshni S. Patel. 2017. 'Lixisenatide (Adlyxin): A Once-Daily Incretin Mimetic Injection for Type-2 Diabetes', *P & T : a peer-reviewed journal for formulary management*, 42: 676-711.

Lewiecki, E. M. 2006. 'RANK ligand inhibition with denosumab for the management of osteoporosis', *Expert Opin Biol Ther*, 6: 1041-50.

Li, W., and J. Frank. 2007. 'Transfer RNA in the hybrid P/E state: correlating molecular dynamics simulations with cryo-EM data', *Proc Natl Acad Sci U S A*, 104: 16540-5.

Li, Xiaoting, Tian Hua, Kiran Vemuri, Jo-Hao Ho, Yiran Wu, Lijie Wu, Petr Popov, Othman Benchama, Nikolai Zvonok, K'ara Locke, Lu Qu, Gye Won Han, Malliga R. Iyer, Resat Cinar, Nathan J. Coffey, Jingjing Wang, Meng Wu, Vsevolod Katritch, Suwen Zhao, George Kunos, Laura M. Bohn, Alexandros Makriyannis, Raymond C. Stevens, and Zhi-Jie Liu. 2019. 'Crystal Structure of the Human Cannabinoid Receptor CB2', *Cell*, 176: 459-67.e13.

Liang, Qian-Yun, Chun-Li Pang, Jun-Wei Li, Su-Hua Zhang, Hui Liu, Yong Zhan, and Hai-Long An. 2017. 'Allosteric Mechanism of Calmodulin Revealed by Targeted Molecular Dynamics Simulation', *Chinese Physics Letters*, 34: 068701.

Liang, Y. L., M. Khoshouei, A. Glukhova, S. G. B. Furness, P. Zhao, L. Clydesdale, C. Koole, T. T. Truong, D. M. Thal, S. Lei, M. Radjainia, R. Danev, W. Baumeister, M. W. Wang, L. J. Miller, A. Christopoulos, P. M. Sexton, and D. Wootten. 2018. 'Phase-plate cryo-EM structure of a biased agonist-bound human GLP-1 receptor-Gs complex', *Nature*, 555: 121-25.

Liang, Yi-Lynn, Matthew J. Belousoff, Peishen Zhao, Cassandra Koole, Madeleine M. Fletcher, Tin T. Truong, Villy Julita, George Christopoulos, H. Eric Xu, Yan Zhang, Maryam Khoshouei, Arthur Christopoulos, Radostin Danev, Patrick M. Sexton, and Denise Wootten. 2020. 'Toward a Structural Understanding of Class B GPCR Peptide Binding and Activation', *Molecular Cell*, 77: 656-68.e5.

Liang, Yi-Lynn, Maryam Khoshouei, Giuseppe Deganutti, Alisa Glukhova, Cassandra Koole, Thomas S Peat, Mazdak Radjainia, Jürgen M Plitzko, Wolfgang Baumeister, and Laurence J Miller. 2018. 'Cryo-EM structure of the active, G s-protein complexed, human CGRP receptor', *Nature*, 561: 492-97.

Liang, Yi-Lynn, Maryam Khoshouei, Mazdak Radjainia, Yan Zhang, Alisa Glukhova, Jeffrey Tarrasch, David M. Thal, Sebastian G. B. Furness, George Christopoulos, Thomas Coudrat, Radostin Danev, Wolfgang Baumeister, Laurence J. Miller, Arthur Christopoulos, Brian K. Kobilka, Denise Wootten,

Georgios Skiniotis, and Patrick M. Sexton. 2017. 'Phase-plate cryo-EM structure of a class B GPCR–G-protein complex', *Nature*, 546: 118.

Liao, Chenyi, Jacob M. Remington, Victor May, and Jianing Li. 2021. 'Molecular Basis of Class B GPCR Selectivity for the Neuropeptides PACAP and VIP', *Frontiers in Molecular Biosciences*, 8.

Lim, GB. 2019. 'GLP1R agonists: primary cardiovascular prevention and oral administration. ', *Nat Rev Cardiol* 16,.

Limited, Novo Nordisk. 2020. 'Rybelsus', Accessed September 2020. https://www.medicines.org.uk/emc/product/11507/smpc.

Lin, Chia-Wei, Feifei Yan, Satoko Shimamura, Sebastian Barg, and Show-Ling %J Diabetes Shyng. 2005. 'Membrane phosphoinositides control insulin secretion through their effects on ATP-sensitive K+ channel activity', 54: 2852-58.

Lin, S. C., C. R. Lin, I. Gukovsky, A. J. Lusis, P. E. Sawchenko, and M. G. Rosenfeld. 1993. 'Molecular basis of the little mouse phenotype and implications for cell type-specific growth', *Nature*, 364: 208-13.

Lindorff-Larsen K, Piana S, Dror RO, Shaw DE. How fast-folding proteins fold. Science. 2011;334(6055):517–20. Epub 2011/10/29. pmid:22034434.

Liu, J. J., R. Horst, V. Katritch, R. C. Stevens, and K. Wuthrich. 2012. 'Biased signaling pathways in beta2-adrenergic receptor characterized by 19F-NMR', *Science*, 335: 1106-10.

Liu, Wei, Eugene Chun, Aaron A Thompson, Pavel Chubukov, Fei Xu, Vsevolod Katritch, Gye Won Han, Christopher B Roth, Laura H Heitman, and Adriaan P %J Science IJzerman. 2012. 'Structural basis for allosteric regulation of GPCRs by sodium ions', 337: 232-36.

Liu, Y., Q. Tian, J. Yang, H. Wang, and T. Hong. 2018. 'No pancreatic safety concern following glucagon-like peptide-1 receptor agonist therapies: A pooled analysis of cardiovascular outcome trials', *Diabetes Metab Res Rev*, 34: e3061.

Lohse, Martin J, Jeffrey L Benovic, Juan Codina, Marc G Caron, and Robert J Lefkowitz. 1990. 'beta-Arrestin: a protein that regulates beta-adrenergic receptor function', *Science*, 248: 1547-50.

Lomize, Mikhail A., Irina D. Pogozheva, Hyeon Joo, Henry I. Mosberg, and Andrei L. Lomize. 2012. 'OPM database and PPM web server: resources for positioning of proteins in membranes', *Nucleic Acids Research*, 40: D370-D76.

Lopez de Maturana, R., and D. Donnelly. 2002. 'The glucagon-like peptide-1 receptor binding site for the N-terminus of GLP-1 requires polarity at Asp198 rather than negative charge', *FEBS Lett*, 530: 244-8.

Lorenz, M., and K. C. Holmes. 2010. 'The actin-myosin interface', *Proc Natl Acad Sci U S A*, 107: 12529-34.

Luck, Michael D, Percy H Carter, and Thomas J Gardella. 1999. 'The (1–14) fragment of parathyroid hormone (PTH) activates intact and amino-terminally truncated PTH-1 receptors', *Molecular Endocrinology*, 13: 670-80.

Luo, James Gumbart; Dong. 2007. "HBonds Plugin, Version 1.2." In, The Hbonds plugin counts the number of hydrogen bonds formed throughout a trajectory. The search can be restricted to a single selection or between two distinct selections, as well as a frame range given by the user. https://www.ks.uiuc.edu/Research/vmd/plugins/hbonds/.

Luttrell, L. M., S. S. Ferguson, Y. Daaka, W. E. Miller, S. Maudsley, G. J. Della Rocca, F. Lin, H. Kawakatsu, K. Owada, D. K. Luttrell, M. G. Caron, and R. J. Lefkowitz. 1999. 'Beta-arrestin-dependent formation of beta2 adrenergic receptor-Src protein kinase complexes', *Science*, 283: 655-61.

Luttrell, Louis M, Francine L Roudabush, Eric W Choy, William E Miller, Michael E Field, Kristen L Pierce, and Robert J Lefkowitz. 2001. 'Activation and targeting of extracellular signal-regulated kinases by β-arrestin scaffolds', *Proceedings of the National Academy of Sciences*, 98: 2449-54.

Luttrell, Louis M., Jialu Wang, Bianca Plouffe, Jeffrey S. Smith, Lama Yamani, Suneet Kaur, Pierre-Yves Jean-Charles, Christophe Gauthier, Mi-Hye Lee, Biswaranjan Pani, Jihee Kim, Seungkirl Ahn, Sudarshan Rajagopal, Eric Reiter, Michel Bouvier, Sudha K. Shenoy, Stéphane A. Laporte, Howard A. Rockman, and Robert J. Lefkowitz. 2018. 'Manifold roles of β-arrestins in GPCR signaling elucidated with siRNA and CRISPR/Cas9', 11: eaat7650.

Lyman, Edward, Chris Higgs, Byungchan Kim, Dmitry Lupyan, John C. Shelley, Ramy Farid, and Gregory A. Voth. 2009. 'A Role for a Specific Cholesterol Interaction in Stabilizing the Apo Configuration of the Human A2A Adenosine Receptor', *Structure*, 17: 1660-68.

M Bajaj, and, and T Blundell. 1984. 'Evolution and the Tertiary Structure of Proteins', *Annual Review of Biophysics and Bioengineering*, 13: 453-92.

Ma, Honglei, Wei Huang, Xiaoxi Wang, Lihua Zhao, Yi Jiang, Feng Liu, Wei Guo, Xianqiang Sun, Wenge Zhong, Daopeng Yuan, and H. Eric Xu. 2020. 'Structural insights into the activation of GLP-1R by a small molecule agonist', *Cell Research*, 30: 1140-42.

Ma, Shanshan, Qingya Shen, Li-Hua Zhao, Chunyou Mao, X. Edward Zhou, Dan-Dan Shen, Parker W. de Waal, Peng Bi, Chuntao Li, Yi Jiang, Ming-Wei Wang, Patrick M. Sexton, Denise Wootten, Karsten Melcher, Yan Zhang, and H. Eric Xu. 2020. 'Molecular Basis for Hormone Recognition and Activation of Corticotropin-Releasing Factor Receptors', *Molecular Cell*, 77: 669-80.e4.

MacDonald, Michael J., Lacmbouh Ade, James M. Ntambi, Israr-Ul H. Ansari, and Scott W. Stoker. 2015. 'Characterization of phospholipids in insulin secretory granules and mitochondria in pancreatic beta cells and their changes with glucose stimulation', *J Biol Chem*, 290: 11075-92.

MacDonald, Michael J., Agnieszka Dobrzyn, James Ntambi, and Scott W. Stoker. 2008. 'The role of rapid lipogenesis in insulin secretion: Insulin secretagogues acutely alter lipid composition of INS-1 832/13 cells', *Archives of Biochemistry and Biophysics*, 470: 153-62.

MacKerell, Alexander D. 2000. "Overview and parameter optimization of CHARMM Force FIeld." In *Especially Sanibel Conference 2003*. JCC v21, 86,105 (2000).

Maeda, Akira, Makoto Okazaki, David M. Baron, Thomas Dean, Ashok Khatri, Mathew Mahon, Hiroko Segawa, Abdul B. Abou-Samra, Harald Jüppner, Kenneth D. Bloch, John T. Potts, and Thomas J. Gardella. 2013. 'Critical role of parathyroid hormone (PTH) receptor-1 phosphorylation in regulating acute responses to PTH', *Proceedings of the National Academy of Sciences*, 110: 5864-69.

Maeda, Shoji, Qianhui Qu, Michael J Robertson, Georgios Skiniotis, and Brian K Kobilka. 2019. 'Structures of the M1 and M2 muscarinic acetylcholine receptor/G-protein complexes', *Science*, 364: 552-57.

Mahoney, Jacob P, and Roger K Sunahara. 2016. 'Mechanistic insights into GPCR–G protein interactions', *Curr Opin Struct Biol*, 41: 247-54.

Manglik, Aashish, Tae Hun Kim, Matthieu Masureel, Christian Altenbach, Zhongyu Yang, Daniel Hilger, Michael T Lerch, Tong Sun Kobilka, Foon Sun Thian, and Wayne L %J Cell Hubbell. 2015. 'Structural insights into the dynamic process of β2-adrenergic receptor signaling', 161: 1101-11.

Mann, Rosalind, Mark J. Wigglesworth, and Dan Donnelly. 2008. 'LIGAND-RECEPTOR INTERACTIONS AT THE PARATHYROID HORMONE (PTH) RECEPTORS: SUBTYPE BINDING SELECTIVITY IS MEDIATED VIA AN INTERACTION BETWEEN RESIDUE 23 ON THE LIGAND AND RESIDUE 41 ON THE RECEPTOR', *Mol Pharmacol*.

Manna, Moutusi, Miia Niemelä, Joona Tynkkynen, Matti Javanainen, Waldemar Kulig, Daniel J Müller, Tomasz Rog, and Ilpo %J Elife Vattulainen. 2016. 'Mechanism of allosteric regulation of β2-adrenergic receptor by cholesterol', 5: e18432.

Mapelli, Claudio, Sesha I Natarajan, Jean-Philippe Meyer, Margarita M Bastos, Michael S Bernatowicz, Ving G Lee, Jelka Pluscec, Douglas J Riexinger, Ellen S Sieber-McMaster, and Keith L Constantine. 2009. 'Eleven amino acid glucagon-like peptide-1 receptor agonists with antidiabetic activity', *Journal of Medicinal Chemistry*, 52: 7788-99.

Matheus, A. S., L. R. Tannus, R. A. Cobas, C. C. Palma, C. A. Negrato, and M. B. Gomes. 2013. 'Impact of diabetes on cardiovascular disease: an update', *Int J Hypertens*, 2013: 653789.

Mathi, S. K., Y. Chan, X. Li, and M. B. Wheeler. 1997. 'Scanning of the glucagon-like peptide-1 receptor localizes G protein-activating determinants primarily to the N terminus of the third intracellular loop', *Mol Endocrinol*, 11: 424-32.

Mattedi, G., S. Acosta-Gutiérrez, T. Clark, and F. L. Gervasio. 2020. 'A combined activation mechanism for the glucagon receptor', *Proc Natl Acad Sci U S A*, 117: 15414-22.

Mayer, G.P. 1979. 'Parathyroid hormone secretion', *Endocrinology*, 2: 607-11.

McCammon, J. Andrew, Bruce R. Gelin, and Martin Karplus. 1977. 'Dynamics of folded proteins', *Nature*, 267: 585-90.

McGreevy, Ryan, Ivan Teo, Abhishek Singharoy, and Klaus Schulten. 2016. 'Advances in the molecular dynamics flexible fitting method for cryo-EM modeling', *Methods (San Diego, Calif.)*, 100: 50-60.

McMullan, G, AR Faruqi, and R Henderson. 2016. 'Direct electron detectors.' in, *Methods in enzymology* (Elsevier).

McNicholas, S., E. Potterton, K. S. Wilson, and M. E. M. Noble. 2011. 'Presenting your structures: the CCP4mg molecular-graphics software', *Acta Crystallographica Section D*, 67: 386-94.

Meldolesi, J., J. D. Jamieson, and G. E. Palade. 1971. 'Composition of cellular membranes in the pancreas of the guinea pig. II. Lipids', *The Journal of cell biology*, 49: 130-49.

Michael J. Perley, David M. Kipnis. 1967. 'Plasma Insulin Responses to ORal and Intravenous Glucose: Studies in Normal and Diabetic Subjects', *The Journal of Clinical Investigation*, 46.

Miller, Paul D, Gary Hattersley, Bente Juel Riis, Gregory C Williams, Edith Lau, Luis Augusto Russo, Peter Alexandersen, Cristiano AF Zerbini, Ming-yi Hu, and Alan G Harris. 2016. 'Effect of abaloparatide vs placebo on new vertebral fractures in postmenopausal women with osteoporosis: a randomized clinical trial', *Jama*, 316: 722-33.

Miller-Gallacher, Jennifer L., Rony Nehmé, Tony Warne, Patricia C. Edwards, Gebhard F. X. Schertler, Andrew G. W. Leslie, and Christopher G. Tate. 2014. 'The 2.1 Å Resolution Structure of Cyanopindolol-Bound β1-Adrenoceptor Identifies an Intramembrane Na+ Ion that Stabilises the Ligand-Free Receptor', *PLOS ONE*, 9: e92727.

Milligan, G., and E. Kostenis. 2006. 'Heterotrimeric G-proteins: a short history', *Br J Pharmacol*, 147 Suppl 1: S46-55.

Mojsov, S., G. C. Weir, and J. F. Habener. 1987. 'Insulinotropin: glucagon-like peptide I (7-37) co-encoded in the glucagon gene is a potent stimulator of insulin release in the perfused rat pancreas', *J Clin Invest*, 79: 616-19.

Mojsov, Svetlana. 1992. 'Structural requirements for biological activity of glucagon-like peptide-I', *International journal of peptide and protein research*, 40: 333-43.

Montague, W., and E. N. Parkin. 1980. 'Changes in membrane lipids of the beta-cell during insulin secretion', *Horm Metab Res Suppl*, Suppl 10: 153-7.

Montrose-Rafizadeh, C., J. M. Egan, and J. Roth. 1994. 'Incretin hormones regulate glucose-dependent insulin secretion in RIN 1046-38 cells: mechanisms of action', *Endocrinology*, 135: 589-94.

Montrose-Rafizadeh, C., H. Yang, B. D. Rodgers, A. Beday, L. A. Pritchette, and J. Eng. 1997. 'High potency antagonists of the pancreatic glucagon-like peptide-1 receptor', *J Biol Chem*, 272: 21201-6.

Moon, Mi Jin, Yoo-Na Lee, Sumi Park, Arfaxad Reyes-Alcaraz, Jong-Ik Hwang, Robert Peter Millar, Han Choe, and Jae Young Seong. 2015. 'Ligand binding pocket formed by evolutionarily conserved residues in the glucagon-like peptide-1 (GLP-1) receptor core domain', *Journal of Biological Chemistry*, 290: 5696-706.

Muehleisen, B., D. D. Bikle, C. Aguilera, D. W. Burton, G. L. Sen, L. J. Deftos, and R. L. Gallo. 2012. 'PTH/PTHrP and vitamin D control antimicrobial peptide expression and susceptibility to bacterial skin infection', *Sci Transl Med*, 4: 135ra66.

Muirhead, H., J. M. Cox, L. Mazzarella, and M. F. Perutz. 1967. 'Structure and function of haemoglobin. 3. A three-dimensional fourier synthesis of human deoxyhaemoglobin at 5.5 Angstrom resolution', *J Mol Biol*, 28: 117-56.

Murakami, Akira, Toshihiro Yajima, Hitoshi Sakuma, Margaret J McLaren, and George Inana. 1993. 'X-Arrestin: a new retinal arrestin mapping to the X chromosome', *FEBS Lett*, 334: 203-09.

Myher, J. J., A. Kuksis, and S. Pind. 1989. 'Molecular species of glycerophospholipids and sphingomyelins of human erythrocytes: Improved method of analysis', *Lipids*, 24: 396-407.

Nauck, M., F. Stöckmann, R. Ebert, and W. Creutzfeldt. 1986. 'Reduced incretin effect in type 2 (non-insulin-dependent) diabetes', *Diabetologia*, 29: 46-52.

Nauck, Michael A., Daniel R. Quast, Jakob Wefers, and Juris J. Meier. 2021. 'GLP-1 receptor agonists in the treatment of type 2 diabetes – state-of-the-art', *Molecular Metabolism*, 46: 101102.

Neubig, RR. 1994. 'Membrane organization in G-protein mechanisms', *The FASEB Journal*, 8: 939-46.

NICE. 2020. "Type 2 diabetes in adults: management." In, 35. National Insittue for Health and Care Excellence.

Noble, Alex J, Venkata P Dandey, Hui Wei, Julia Brasch, Jillian Chase, Priyamvada Acharya, Yong Zi Tan, Zhening Zhang, Laura Y Kim, and Giovanna Scapin. 2018. 'Routine single particle CryoEM sample and grid characterization by tomography', *eLife*, 7: e34257.

Nobles, Kelly N., Kunhong Xiao, Seungkirl Ahn, Arun K. Shukla, Christopher M. Lam, Sudarshan Rajagopal, Ryan T. Strachan, Teng-Yi Huang, Erin A. Bressler, Makoto R. Hara, Sudha K. Shenoy, Steven P. Gygi, and Robert J. Lefkowitz. 2011. 'Distinct Phosphorylation Sites on the β₂-

Adrenergic Receptor Establish a Barcode That Encodes Differential Functions of β-Arrestin', *Science Signaling*, 4: ra51-ra51.

Noda, K., Y. Saad, R. M. Graham, and S. S. Karnik. 1994. 'The high affinity state of the beta 2-adrenergic receptor requires unique interaction between conserved and non-conserved extracellular loop cysteines', *J Biol Chem*, 269: 6743-52.

Nygaard, R., T. M. Frimurer, B. Holst, M. M. Rosenkilde, and T. W. Schwartz. 2009. 'Ligand binding and micro-switches in 7TM receptor structures', *Trends Pharmacol Sci*, 30: 249-59.

O'Harte, F. P., Y. H. Abdel-Wahab, J. M. Conlon, and P. R. Flatt. 1998. 'Glycation of glucagon-like peptide-1(7-36)amide: characterization and impaired action on rat insulin secreting cells', *Diabetologia*, 41: 1187-93.

Oakley, Robert H, Stéphane A Laporte, Jason A Holt, Larry S Barak, and Marc G Caron. 2001. 'Molecular determinants underlying the formation of stable intracellular G protein-coupled receptor-β-arrestin complexes after receptor endocytosis', *Journal of Biological Chemistry*, 276: 19452-60.

Oakley, Robert H, Stéphane A Laporte, Jason A Holt, Marc G Caron, and Larry S Barak. 2000. 'Differential affinities of visual arrestin, βarrestin1, and βarrestin2 for G protein-coupled receptors delineate two major classes of receptors', *Journal of Biological Chemistry*, 275: 17201-10.

Okashah, N., Q. Wan, S. Ghosh, M. Sandhu, A. Inoue, N. Vaidehi, and N. A. Lambert. 2019a. 'Variable G protein determinants of GPCR coupling selectivity', *Proc Natl Acad Sci U S A*, 116: 12054-59.

Okashah, Najeah, Qingwen Wan, Soumadwip Ghosh, Manbir Sandhu, Asuka Inoue, Nagarajan Vaidehi, and Nevin A. Lambert. 2019b. 'Variable G protein determinants of GPCR coupling selectivity', 116: 12054-59.

Oldham, W. M., and H. E. Hamm. 2008. 'Heterotrimeric G protein activation by G-protein-coupled receptors', *Nat Rev Mol Cell Biol*, 9: 60-71.

Orasanu, Gabriela, and Jorge %J Journal of the American College of Cardiology Plutzky. 2009. 'The pathologic continuum of diabetic vascular disease', 53: S35-S42.

Orskov, C., J. J. Holst, S. Knuhtsen, F. G. Baldissera, S. S. Poulsen, and O. V. Nielsen. 1986. 'Glucagon-like peptides GLP-1 and GLP-2, predicted products of the glucagon gene, are secreted separately from pig small intestine but not pancreas', *Endocrinology*, 119: 1467-75.

Orskov, C., L. Rabenhoj, A. Wettergren, H. Kofod, and J. J. Holst. 1994. 'Tissue and plasma concentrations of amidated and glycine-extended glucagon-like peptide I in humans', *Diabetes*, 43: 535-9.

Orskov, C., A. Wettergren, and J. J. Holst. 1993. 'Biological effects and metabolic rates of glucagonlike peptide-1 7-36 amide and glucagonlike peptide-1 7-37 in healthy subjects are indistinguishable', *Diabetes*, 42: 658-61.

Pace, C. N., and J. M. Scholtz. 1998. 'A helix propensity scale based on experimental studies of peptides and proteins', *Biophys J*, 75: 422-27.

Palczewski, K., T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. Le Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto, and M. Miyano. 2000. 'Crystal structure of rhodopsin: A G protein-coupled receptor', *Science*, 289: 739-45.

Pan, Ling, Eugenia V Gurevich, and Vsevolod V Gurevich. 2003. 'The nature of the arrestin· receptor complex determines the ultimate fate of the internalized receptor', *Journal of Biological Chemistry*, 278: 11623-32.

Parameswaran, N., and W. S. Spielman. 2006. 'RAMPs: The past, present and future', *Trends Biochem Sci*, 31: 631-8.

Park, J. H., P. Scheerer, K. P. Hofmann, H. W. Choe, and O. P. Ernst. 2008. 'Crystal structure of the ligand-free G-protein-coupled receptor opsin', *Nature*, 454: 183-7.

Parker, JC, KM Andrews, DM Rescek, W Massefski Jr, GC Andrews, LG Contillo, RW Stevenson, DH Singleton, and RT Suleske. 1998. 'Structure-function analysis of a series of glucagon-like peptide-1 analogs', *The Journal of peptide research*, 52: 398-409.

Parthier, C., S. Reedtz-Runge, R. Rudolph, and M. T. Stubbs. 2009. 'Passing the baton in class B GPCRs: peptide hormone activation via helix induction?', *Trends Biochem Sci*, 34: 303-10.

Patel, Dhaval, and Stephan N. Witt. 2017. 'Ethanolamine and Phosphatidylethanolamine: Partners in Health and Disease', *Oxidative medicine and cellular longevity*, 2017: 4829180-80.

Pedersen, J. S., D. Dikov, J. L. Flink, H. A. Hjuler, G. Christiansen, and D. E. Otzen. 2006. 'The changing face of glucagon fibrillation: structural polymorphism and conformational imprinting', *J Mol Biol*, 355: 501-23.

Pellegrini, Maria, Alessandro Bisello, Michael Rosenblatt, Michael Chorev, and Dale F. Mierke. 1998. 'Binding Domain of Human Parathyroid Hormone Receptor: From Conformation to Function', *Biochemistry*, 37: 12737-43.

Periole, X. 2017. 'Interplay of G Protein-Coupled Receptors with the Membrane: Insights from Supra-Atomic Coarse Grain Molecular Dynamics Simulations', *Chem Rev*, 117: 156-85.

Perrin, Marilyn H, Christy RR Grace, Michael R DiGruccio, Wolfgang H Fischer, Samir K Maji, Jeffrey P Cantle, Sean Smith, Gerard Manning, Wylie W Vale, and Roland Riek. 2007. 'Distinct structural and functional roles of conserved residues in the first extracellular domain of receptors for corticotropin-releasing factor and related G-protein-coupled receptors', *Journal of Biological Chemistry*, 282: 37529-36.

Perry, Stephen J, and Robert J Lefkowitz. 2002. 'Arresting developments in heptahelical receptor signaling and regulation', *Trends in cell biology*, 12: 130-38.

Pfister, C, M Chabre, J Plouet, VV Tuyen, Y De Kozak, JP Faure, and H Kuhn. 1985. 'Retinal S antigen identified as the 48K protein regulating light-dependent phosphodiesterase in rods', *Science*, 228: 891-93.

Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. 2005. 'Scalable molecular dynamics with NAMD', *J Comput Chem*, 26: 1781-802.

Pierce, Kristen L, Richard T Premont, and Robert J Lefkowitz. 2002. 'Seven-transmembrane receptors', *Nature Reviews Molecular Cell Biology*, 3: 639.

Pioszak, A. A., and H. E. Xu. 2008. 'Molecular recognition of parathyroid hormone by its G protein-coupled receptor', *Proc Natl Acad Sci U S A*, 105: 5034-9.

Pioszak, Augen A., Kaleeckal G. Harikumar, Naomi R. Parker, Laurence J. Miller, and H. Eric Xu. 2010. 'Dimeric Arrangement of the Parathyroid Hormone Receptor and a Structural Mechanism for Ligand-induced Dissociation *', *Journal of Biological Chemistry*, 285: 12435-44.

Pioszak, Augen A., Naomi R. Parker, Thomas J. Gardella, and H. Eric Xu. 2009. 'Structural Basis for Parathyroid Hormone-related Protein Binding to the Parathyroid Hormone Receptor and Design of Conformation-selective Peptides *', *Journal of Biological Chemistry*, 284: 28382-91.

Pitcher, J. A., N. J. Freedman, and R. J. Lefkowitz. 1998. 'G protein-coupled receptor kinases', *Annu Rev Biochem*, 67: 653-92.

Plamboeck, A., J. J. Holst, R. D. Carr, and C. F. Deacon. 2005. 'Neutral endopeptidase 24.11 and dipeptidyl peptidase IV are both mediators of the degradation of glucagon-like peptide 1 in the anaesthetised pig', *Diabetologia*, 48: 1882-90.

Poole-Wilson, Philip A., Karl Swedberg, John G. F. Cleland, Andrea Di Lenarda, Peter Hanrath, Michel Komajda, Jacobus Lubsen, Beatrix Lutiger, Marco Metra, Willem J. Remme, Christian Torp-Pedersen, Armin Scherhag, and Allan Skene. 2003. 'Comparison of carvedilol and metoprolol on clinical outcomes in patients with chronic heart failure in the Carvedilol Or Metoprolol European Trial (COMET): randomised controlled trial', *The Lancet*, 362: 7-13.

Potterton, Elizabeth, Peter Briggs, Maria Turkenburg, and Eleanor Dodson. 2003. *A graphical user interface to the CCP4 program suite. Acta Crystallogr D Biol Crystallogr D59:1131-1137*.

Pratley, Richard E., Matthew J. Crowley, Mette Gislum, Christin L. Hertz, Thomas B. Jensen, Kamlesh Khunti, Ofri Mosenzon, and John B. Buse. 2021. 'Oral Semaglutide Reduces HbA(1c) and Body Weight in Patients with Type 2 Diabetes Regardless of Background Glucose-Lowering Medication: PIONEER Subgroup Analyses', *Diabetes therapy : research, treatment and education of diabetes and related disorders*, 12: 1099-116.

Probst, William C., Lenore A. Snyder, David I. Schuster, JÜRgen Brosius, and Stuart C. Sealfon. 1992. 'Sequence Alignment of the G-Protein Coupled Receptor Superfamily', *DNA and Cell Biology*, 11: 1-20.

Qi, Yifei, Jumin Lee, Abhishek Singharoy, Ryan McGreevy, Klaus Schulten, and Wonpil Im. 2017. 'CHARMM-GUI MDFF/xMDFF Utilizer for Molecular Dynamics Flexible Fitting Simulations in Various Environments', *The Journal of Physical Chemistry B*, 121: 3718-23.

Qiao, Anna, Shuo Han, Xinmei Li, Zhixin Li, Peishen Zhao, Antao Dai, Rulve Chang, Linhua Tai, Qiuxiang Tan, Xiaojing Chu, Limin Ma, Thor Seneca Thorsen, Steffen Reedtz-Runge, Dehua Yang, Ming-Wei Wang, Patrick M. Sexton, Denise Wootten, Fei Sun, Qiang Zhao, and Beili Wu. 2020. 'Structural basis of G_s and G_i recognition by the human glucagon receptor', *Science*, 367: 1346-52.

Bernardi R., Bhandarkar M., Bhatele A., Bohm E., Brunner R., Buelens F., Chipot C., Dalke A., Dixit S., Fiorin G., Freddolino F., Fu H., Grayson P., Gullingsrud J., Gursoy A., Hardy D., Harrison C.,. Hénin J., Humphrey W., Hurwitz D., Hynninen A., Jain N., Krawetz N., Kumar S., Kunzman D., Lai J., Lee C., McGreevy R., Mei C., Melo M., Nelson M., Phillips J., Radak B., Rudack T., Sarood O., Shinozaki A., Tanner D., Wells D., Zheng G., Zhu F., 2016. "NAMD User's Guide Version 2.12." In. Urbana, IL 61801 Theoretical Biophysics Group University of Illinois and Beckman Institute 405 N. Mathews

Raebel, Marsha A., Jennifer L. Ellis, Emily B. Schroeder, Stanley Xu, Patrick J. O'Connor, Jodi B. Segal, Melissa G. Butler, Julie A. Schmittdiel, H. Lester Kirchner, Glenn K. Goodrich, Jean M. Lawrence, Gregory A. Nichols, Katherine M. Newton, Ram D. Pathak, and John F. Steiner. 2014. 'Intensification of antihyperglycemic therapy among patients with incident diabetes: a Surveillance Prevention and Management of Diabetes Mellitus (SUPREME-DM) study', 23: 699-710.

Ragnarsson, Lotten, Åsa Andersson, Walter G. Thomas, and Richard J. Lewis. 2019. 'Mutations in the NPxxY motif stabilize pharmacologically distinct conformational states of the α_{1B}- and β₂-adrenoceptors', *Science Signaling*, 12: eaas9485.

Rahuel-Clermont, Sophie, Raphaël Bchini, Sophie Barbe, Séverine Boutserin, Isabelle André, and François Talfournier. 2019. 'Enzyme Active Site Loop Revealed as a Gatekeeper for Cofactor Flip by Targeted Molecular Dynamics Simulations and FRET-Based Kinetics', *ACS Catalysis*, 9: 1337-46.

Rakesh, Kriti, ByungSu Yoo, Il-Man Kim, Natasha Salazar, Ki-Seok Kim, and Howard A. Rockman. 2010. 'β-Arrestin–Biased Agonism of the Angiotensin Receptor Induced by Mechanical Stress', *Science Signaling*, 3: ra46-ra46.

Rana, Rajendra S., Anjaneyulu Kowluru, and Michael J. MacDonald. 1986. 'Secretagogue-responsive and -unresponsive pools of phosphatidylinositol in pancreatic islets', *Archives of Biochemistry and Biophysics*, 245: 411-16.

Ranjan, Ravi, Pragya Gupta, and Arun K Shukla. 2016. 'GPCR Signaling: β-arrestins Kiss and Remember', *Current Biology*, 26: R285-R88.

Rankin, Wayne, Vivian Grill, and T John Martin. 1997. 'Parathyroid hormone-related protein and hypercalcemia', *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 80: 1564-71.

Rankovic, Zoran, Tarsis F. Brust, and Laura M. Bohn. 2016. 'Biased agonism: An emerging paradigm in GPCR drug discovery', *Bioorganic & Medicinal Chemistry Letters*, 26: 241-50.

Rasmussen, S. G., H. J. Choi, J. J. Fung, E. Pardon, P. Casarosa, P. S. Chae, B. T. Devree, D. M. Rosenbaum, F. S. Thian, T. S. Kobilka, A. Schnapp, I. Konetzki, R. K. Sunahara, S. H. Gellman, A. Pautsch, J. Steyaert, W. I. Weis, and B. K. Kobilka. 2011a. 'Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor', *Nature*, 469: 175-80.

Rasmussen, S. G., H. J. Choi, D. M. Rosenbaum, T. S. Kobilka, F. S. Thian, P. C. Edwards, M. Burghammer, V. R. Ratnala, R. Sanishvili, R. F. Fischetti, G. F. Schertler, W. I. Weis, and B. K. Kobilka. 2007. 'Crystal structure of the human beta2 adrenergic G-protein-coupled receptor', *Nature*, 450: 383-7.

Rasmussen, Søren G. F., Hee-Jung Choi, Juan Jose Fung, Els Pardon, Paola Casarosa, Pil Seok Chae, Brian T. DeVree, Daniel M. Rosenbaum, Foon Sun Thian, Tong Sun Kobilka, Andreas Schnapp, Ingo Konetzki, Roger K. Sunahara, Samuel H. Gellman, Alexander Pautsch, Jan Steyaert, William I. Weis, and Brian K. Kobilka. 2011b. 'Structure of a nanobody-stabilized active state of the β2 adrenoceptor', *Nature*, 469: 175-80.

Rasmussen, Søren G. F., Brian T. DeVree, Yaozhong Zou, Andrew C. Kruse, Ka Young Chung, Tong Sun Kobilka, Foon Sun Thian, Pil Seok Chae, Els Pardon, Diane Calinski, Jesper M. Mathiesen, Syed T. A. Shah, Joseph A. Lyons, Martin Caffrey, Samuel H. Gellman, Jan Steyaert, Georgios Skiniotis, William I. Weis, Roger K. Sunahara, and Brian K. Kobilka. 2011. 'Crystal structure of the β2 adrenergic receptor–Gs protein complex', *Nature*, 477: 549.

Raufman, Jean-Pierre. 1996. 'Bioactive peptides from lizard venoms', *Regul Pept*, 61: 1-18.

Reginster, J. Y., F. Bianic, R. Campbell, M. Martin, S. A. Williams, and L. A. Fitzpatrick. 2019. 'Abaloparatide for risk reduction of nonvertebral and vertebral fractures in postmenopausal women with osteoporosis: a network meta-analysis', *Osteoporosis International*, 30: 1465-73.

Ring, Aaron M., Aashish Manglik, Andrew C. Kruse, Michael D. Enos, William I. Weis, K. Christopher Garcia, and Brian K. Kobilka. 2013. 'Adrenaline-

activated structure of β2-adrenoceptor stabilized by an engineered nanobody', *Nature*, 502: 575.

Robinson, Emma, Mitchel Tate, Samuel Lockhart, Claire McPeake, Karla M. O'Neill, Kevin S. Edgar, Danielle Calderwood, Brian D. Green, Barbara J. McDermott, and David J. Grieve. 2016. 'Metabolically-inactive glucagon-like peptide-1(9–36)amide confers selective protective actions against post-myocardial infarction remodelling', *Cardiovascular Diabetology*, 15: 65.

Robinson, Scott W., Avid M. Afzal, and David P. Leader. 2014. 'Chapter 13 - Bioinformatics: Concepts, Methods, and Data.' in Sandosh Padmanabhan (ed.), *Handbook of Pharmacogenomics and Stratified Medicine* (Academic Press: San Diego).

Rose, A. S., M. Elgeti, U. Zachariae, H. Grubmüller, K. P. Hofmann, P. Scheerer, and P. W. Hildebrand. 2014. 'Position of transmembrane helix 6 determines receptor G protein coupling specificity', *J Am Chem Soc*, 136: 11244-7.

Rosenbaum, D. M., V. Cherezov, M. A. Hanson, S. G. Rasmussen, F. S. Thian, T. S. Kobilka, H. J. Choi, X. J. Yao, W. I. Weis, R. C. Stevens, and B. K. Kobilka. 2007. 'GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function', *Science*, 318: 1266-73.

Rovšnik, Urška, Yuxuan Zhuang, Björn O. Forsberg, Marta Carroni, Linnea Yvonnesdotter, Rebecca J. Howard, and Erik Lindahl. 2021. 'Dynamic closed states of a ligand-gated ion channel captured by cryo-EM and simulations', *Life Science Alliance*, 4: e202101011.

Rowlands, Jordan, Julian Heng, Philip Newsholme, and Rodrigo Carlessi. 2018. 'Pleiotropic Effects of GLP-1 and Analogs on Cell Signaling, Metabolism, and Function', *Frontiers in Endocrinology*, 9.

Rumpf, Tobias, Matthias Schiedel, Berin Karaman, Claudia Roessler, Brian J. North, Attila Lehotzky, Judit Oláh, Kathrin I. Ladwein, Karin Schmidtkunz, Markus Gajer, Martin Pannek, Clemens Steegborn, David A. Sinclair, Stefan Gerhardt, Judit Ovádi, Mike Schutkowski, Wolfgang Sippl, Oliver Einsle, and Manfred Jung. 2015. 'Selective Sirt2 inhibition by ligand-induced rearrangement of the active site', *Nat Commun*, 6: 6263.

Runge, S., C. Gram, H. Brauner-Osborne, K. Madsen, L. B. Knudsen, and B. S. Wulff. 2003. 'Three distinct epitopes on the extracellular face of the glucagon receptor determine specificity for the glucagon amino terminus', *J Biol Chem*, 278: 28005-10.

Runge, S., H. Thogersen, K. Madsen, J. Lau, and R. Rudolph. 2008. 'Crystal structure of the ligand-bound glucagon-like peptide-1 receptor extracellular domain', *J Biol Chem*, 283: 11340-7.

Ryba, D. M., J. Li, C. L. Cowan, B. Russell, B. M. Wolska, and R. J. Solaro. 2017. 'Long-Term Biased beta-Arrestin Signaling Improves Cardiac Structure and Function in Dilated Cardiomyopathy', *Circulation*, 135: 1056-70.

Sabbadin D., Moro S. Supervised molecular dynamics (SuMD) as a helpful tool to depict GPCR-ligand recognition pathway in a nanosecond time scale. *J. Chem. Inf. Model.* 2014;54:372–376. doi: 10.1021/ci400766b

Sackmann, E., J. Engelhardt, K. Fricke, and H. Gaub. 1984. 'On dynamic molecular and elastic properties of lipid bilayers and biological membranes', *Colloids and Surfaces*, 10: 321-35.

Samama, P., S. Cotecchia, T. Costa, and R. J. Lefkowitz. 1993a. 'A mutation-induced activated state of the beta 2-adrenergic receptor. Extending the ternary complex model', *J Biol Chem*, 268: 4625-36.

Samama, Pa, S Cotecchia, T Costa, and RJ Lefkowitz. 1993b. 'A mutation-induced activated state of the beta 2-adrenergic receptor. Extending the ternary complex model', *Journal of Biological Chemistry*, 268: 4625-36.

Sato, Emi, Jun Muto, Ling-Juan Zhang, Christopher A. Adase, James A. Sanford, Toshiya Takahashi, Teruaki Nakatsuji, Ted B. Usdin, and Richard L. Gallo. 2016a. 'The Parathyroid Hormone Second Receptor PTH2R and its Ligand Tuberoinfundibular Peptide of 39 Residues TIP39 Regulate Intracellular Calcium and Influence Keratinocyte Differentiation', *Journal of Investigative Dermatology*, 136: 1449-59.

———. 2016b. 'The Parathyroid Hormone Second Receptor PTH2R and its Ligand Tuberoinfundibular Peptide of 39 Residues TIP39 Regulate Intracellular Calcium and Influence Keratinocyte Differentiation', *The Journal of investigative dermatology*, 136: 1449-59.

Saxena, Roopali, and Amitabha %J Biochimica et Biophysica Acta -Biomembranes Chattopadhyay. 2012. 'Membrane cholesterol stabilizes the human serotonin1A receptor', 1818: 2936-42.

Schann, Stephan, S Mayer, M Frauli, C Franchet, and P Neuville. 2009. "Sounds of silence: innovative approach for identification of novel GPCR-modulator chemical entities." In *ABSTRACTS OF PAPERS OF THE AMERICAN CHEMICAL SOCIETY*. AMER CHEMICAL SOC 1155 16TH ST, NW, WASHINGTON, DC 20036 USA.

Scheerer, P., J. H. Park, P. W. Hildebrand, Y. J. Kim, N. Krauss, H. W. Choe, K. P. Hofmann, and O. P. Ernst. 2008. 'Crystal structure of opsin in its G-protein-interacting conformation', *Nature*, 455: 497-502.

Schiöth, Helgi B., and Robert Fredriksson. 2005. 'The GRAFS classification system of G-protein coupled receptors in comparative perspective', *General and Comparative Endocrinology*, 142: 94-101.

Schipani, E, Craig Langman, J Hunzelman, M Le Merrer, KY Loke, MJ Dillon, C Silve, and H Juppner. 1999. 'A novel parathyroid hormone (PTH)/PTH-related peptide receptor mutation in Jansen's metaphyseal chondrodysplasia', *The Journal of Clinical Endocrinology & Metabolism*, 84: 3052-57.

Schipani, Ernestina, Klaus Kruse, and Harald Juppner. 1995. 'A constitutively active mutant PTH-PTHrP receptor in Jansen-type metaphyseal chondrodysplasia', *Science*, 268: 98-100.

Schlegel, H. Bernhard. 1982. 'Optimization of equilibrium geometries and transition structures', 3: 214-18.

Schlick, Tamar. 2010. *Molecular modeling and simulation: an interdisciplinary guide: an interdisciplinary guide* (Springer Science & Business Media).

Schlitter, J., M. Engels, and P. Kruger. 1994. 'Targeted molecular dynamics: a new approach for searching pathways of conformational transitions', *J Mol Graph*, 12: 84-9.

Schneider, E. H., D. Schnell, A. Strasser, S. Dove, and R. Seifert. 2010. 'Impact of the DRY motif and the missing "ionic lock" on constitutive activity and G-protein coupling of the human histamine H4 receptor', *J Pharmacol Exp Ther*, 333: 382-92.

Schroder, G. F., A. T. Brunger, and M. Levitt. 2007. 'Combining efficient conformational sampling with a deformable elastic network model facilitates structure refinement at low resolution', *Structure*, 15: 1630-41.

Schrödinger, LLC. 2015. "The PyMOL Molecular Graphics System, Version 1.8." In.

Segre, G. V., M. Rosenblatt, B. L. Reiner, J. E. Mahaffey, and J. T. Potts, Jr. 1979. 'Characterization of parathyroid hormone receptors in canine renal cortical

plasma membranes using a radioiodinated sulfur-free hormone analogue. Correlation of binding with adenylate cyclase activity', *J Biol Chem*, 254: 6980-6.

Segre, Gino V. 1996. 'Receptors for parathyroid hormone and parathyroid hormone-related protein', *Principles of bone Biology.*: 377-403.

Sela, Michael, Christian B. Anfinsen, and William F. Harrington. 1957. 'The correlation of ribonuclease activity with specific aspects of tertiary structure', *Biochim Biophys Acta*, 26: 502-12.

Sexton, P. M., A. Albiston, M. Morfis, and N. Tilakaratne. 2001. 'Receptor activity modifying proteins', *Cell Signal*, 13: 73-83.

Shao, Zhenhua, Jie Yin, Karen Chapman, Magdalena Grzemska, Lindsay Clark, Junmei Wang, and Daniel M. Rosenbaum. 2016. 'High-resolution crystal structure of the human CB1 cannabinoid receptor', *Nature*, 540: 602.

Shenoy, S. K., L. S. Barak, K. Xiao, S. Ahn, M. Berthouze, A. K. Shukla, L. M. Luttrell, and R. J. Lefkowitz. 2007. 'Ubiquitination of β-arrestin links seven-transmembrane receptor endocytosis and ERK activation', *Journal of Biological Chemistry*, 282: 29549-62.

Shenoy, S. K., and R. J. Lefkowitz. 2005. 'Seven-transmembrane receptor signaling through beta-arrestin', *Sci STKE*, 2005: cm10.

Shenoy, Sudha K, and Robert J Lefkowitz. 2003. 'Trafficking patterns of β-arrestin and G protein-coupled receptors determined by the kinetics of β-arrestin deubiquitination', *Journal of Biological Chemistry*, 278: 14498-506.

Shenoy, Sudha K, Patricia H McDonald, Trudy A Kohout, and Robert J Lefkowitz. 2001. 'Regulation of receptor fate by ubiquitination of activated β2-adrenergic receptor and β-arrestin', *Science*, 294: 1307-13.

Shi, L., G. Liapakis, R. Xu, F. Guarnieri, J. A. Ballesteros, and J. A. Javitch. 2002. 'Beta2 adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch', *J Biol Chem*, 277: 40989-96.

Shimizu, Masaru, Percy H. Carter, Ashok Khatri, John T. Potts, Jr., and Thomas J. Gardella. 2001. 'Enhanced Activity in Parathyroid Hormone-(1–14) and -(1–11): Novel Peptides for Probing Ligand-Receptor Interactions*', *Endocrinology*, 142: 3068-74.

Shimizu, Masaru, Eri Joyashiki, Hiroshi Noda, Tomoyuki Watanabe, Makoto Okazaki, Miho Nagayasu, Kenji Adachi, Tatsuya Tamura, John T Potts Jr, and Thomas J Gardella. 2016. 'Pharmacodynamic actions of a long-acting PTH analog (LA-PTH) in thyroparathyroidectomized (TPTX) rats and normal monkeys', *Journal of Bone and Mineral Research*, 31: 1405-12.

Shukla, Arun K., Gerwin H. Westfield, Kunhong Xiao, Rosana I. Reis, Li-Yin Huang, Prachi Tripathi-Shukla, Jiang Qian, Sheng Li, Adi Blanc, Austin N. Oleskie, Anne M. Dosey, Min Su, Cui-Rong Liang, Ling-Ling Gu, Jin-Ming Shan, Xin Chen, Rachel Hanna, Minjung Choi, Xiao Jie Yao, Bjoern U. Klink, Alem W. Kahsai, Sachdev S. Sidhu, Shohei Koide, Pawel A. Penczek, Anthony A. Kossiakoff, Virgil L. Woods Jr, Brian K. Kobilka, Georgios Skiniotis, and Robert J. Lefkowitz. 2014. 'Visualization of arrestin recruitment by a G-protein-coupled receptor', *Nature*, 512: 218-22.

Shulman, G. I., D. L. Rothman, T. Jue, P. Stein, R. A. DeFronzo, and R. G. Shulman. 1990. 'Quantitation of muscle glycogen synthesis in normal subjects and subjects with non-insulin-dependent diabetes by 13C nuclear magnetic resonance spectroscopy', *N Engl J Med*, 322: 223-8.

Shyangdan, D. S., P. Royle, C. Clar, P. Sharma, N. Waugh, and A. Snaith. 2011. 'Glucagon-like peptide analogues for type 2 diabetes mellitus', *Cochrane Database of Systematic Reviews*.

Siegel, E. G., B. Gallwitz, G. Scharf, R. Mentlein, C. Morys-Wortmann, U. R. Fölsch, J. Schrezenmeir, K. Drescher, and W. E. Schmidt. 1999. 'Biological activity of GLP-1-analogues with N-terminal modifications', *Regul Pept*, 79: 93-102.

Singer, S. J., and G. L. Nicolson. 1972. 'The fluid mosaic model of the structure of cell membranes', *Science*, 175: 720-31.

Siu, Fai Yiu, Min He, Chris de Graaf, Gye Won Han, Dehua Yang, Zhiyun Zhang, Caihong Zhou, Qingping Xu, Daniel Wacker, Jeremiah S. Joseph, Wei Liu, Jesper Lau, Vadim Cherezov, Vsevolod Katritch, Ming-Wei Wang, and Raymond C. Stevens. 2013. 'Structure of the human glucagon class B G-protein-coupled receptor', *Nature*, 499: 444.

Skjærven, L., X. Q. Yao, G. Scarabelli, and B. J. Grant. 2014. 'Integrating protein structural dynamics and evolutionary analysis with Bio3D', *BMC Bioinformatics*, 15: 399.

Skrivanek, Zachary, Scott Berry, Don Berry, Jenny Chien, Mary Jane Geiger, James H. Anderson, and Brenda Gaydos. 2012. 'Application of adaptive design methodology in development of a long-acting glucagon-like peptide-1 analog (dulaglutide): statistical design and simulations', *Journal of diabetes science and technology*, 6: 1305-18.

Skyler, Jay S., George L. Bakris, Ezio Bonifacio, Tamara Darsow, Robert H. Eckel, Leif Groop, Per-Henrik Groop, Yehuda Handelsman, Richard A. Insel, Chantal Mathieu, Allison T. McElvaine, Jerry P. Palmer, Alberto Pugliese, Desmond A. Schatz, Jay M. Sosenko, John P. H. Wilding, and Robert E. Ratner. 2017. 'Differentiation of Diabetes by Pathophysiology, Natural History, and Prognosis', *Diabetes*, 66: 241.

Smyth, M. S., and J. H. Martin. 2000. 'x ray crystallography', *Molecular pathology : MP*, 53: 8-14.

Song, Gaojie, Dehua Yang, Yuxia Wang, Chris de Graaf, Qingtong Zhou, Shanshan Jiang, Kaiwen Liu, Xiaoqing Cai, Antao Dai, Guangyao Lin, Dongsheng Liu, Fan Wu, Yiran Wu, Suwen Zhao, Li Ye, Gye Won Han, Jesper Lau, Beili Wu, Michael A. Hanson, Zhi-Jie Liu, Ming-Wei Wang, and Raymond C. Stevens. 2017. 'Human GLP-1 receptor transmembrane domain structure in complex with allosteric modulators', *Nature*, 546: 312.

Spector, A. A., and M. A. Yorek. 1985. 'Membrane lipid composition and cellular function', *J Lipid Res*, 26: 1015-35.

Sprang, S. R. 1997. 'G protein mechanisms: insights from structural analysis', *Annu Rev Biochem*, 66: 639-78.

Sriram, Krishna, and Paul A. Insel. 2018. 'GPCRs as targets for approved drugs: How many targets and how many drugs?': mol.117.111062.

Stephenson, R. P. 1956. 'A modification of receptor theory', *British journal of pharmacology and chemotherapy*, 11: 379-93.

Strader, C. D., M. R. Candelore, W. S. Hill, I. S. Sigal, and R. A. Dixon. 1989. 'Identification of two serine residues involved in agonist activation of the beta-adrenergic receptor', *J Biol Chem*, 264: 13572-8.

Strader, C. D., I. S. Sigal, and R. A. Dixon. 1989. 'Structural basis of beta-adrenergic receptor function', *Faseb j*, 3: 1825-32.

Suhre, K., J. Navaza, and Y. H. Sanejouand. 2006. 'NORMA: a tool for flexible fitting of high-resolution protein structures into low-resolution electron-microscopy-derived density maps', *Acta Crystallogr D Biol Crystallogr*, 62: 1098-100.

Sun, Chaohong, Danying Song, Rachel A Davis-Taber, Leo W Barrett, Victoria E Scott, Paul L Richardson, Ana Pereda-Lopez, Marie E Uchic, Larry R Solomon, and Marc R Lake. 2007. 'Solution structure and mutational analysis of pituitary adenylate cyclase-activating polypeptide binding to the extracellular domain of PAC1-RS', *Proceedings of the National Academy of Sciences*, 104: 7875-80.

Sutkeviciute I., Clark L.J., White A.D., Gardella T.J., Vilardaga J. 2019. 'PTH/PTHrP Receptor Signaling, Allostery, and Structures', Trends Endocrinol Metab, 30(11): 860-874.

Suvannasankha, Attaya, Colin Crean, Isaac Carrera Ochoa, G. David Roodman, and John Chirgwin. 2018. 'PTHrP May Contribute to Cachexia Accompanying Multiple Myeloma Bone Disease', *Blood*, 132: 4450-50.

Swope W.C., Andersen H. C., Berens P. H., and Wilson K. R. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: application to small water clusters. J. Chem. Phys., 76:637–649, 1982

Szczepek, Michal, Florent Beyrière, Klaus Peter Hofmann, Matthias Elgeti, Roman Kazmin, Alexander Rose, Franz J Bartl, David Von Stetten, Martin Heck, and Martha E Sommer. 2014. 'Crystal structure of a common GPCR-binding interface for G protein and arrestin', *Nat Commun*, 5: 1-8.

Tahrani, Abd A., Anthony H. Barnett, and Clifford J. Bailey. 2016. 'Pharmacology and therapeutic implications of current drugs for type 2 diabetes mellitus', *Nature Reviews Endocrinology*, 12: 566-92.

Takhar, S., S. Gyomorey, R. C. Su, S. K. Mathi, X. Li, and M. B. Wheeler. 1996. 'The third cytoplasmic domain of the GLP-1[7-36 amide] receptor is required for coupling to the adenylyl cyclase system', *Endocrinology*, 137: 2175-8.

Tama, F., O. Miyashita, and C. L. Brooks, 3rd. 2004. 'Normal mode based flexible fitting of high-resolution structure into low-resolution experimental data from cryo-EM', *J Struct Biol*, 147: 315-26.

ter Haar, Ernst, Christopher M Koth, Norzehan Abdul-Manan, Lora Swenson, Joyce T Coll, Judith A Lippke, Christopher A Lepre, Miguel Garcia-Guzman, and Jonathan M Moore. 2010. 'Crystal structure of the ectodomain complex of the CGRP receptor, a class-B GPCR, reveals the site of drug antagonism', *Structure*, 18: 1083-93.

Thore, Sophia, Oleg Dyachok, Erik Gylfe, and Anders %J Journal of cell science Tengholm. 2005. 'Feedback activation of phospholipase C via intracellular mobilization and store-operated influx of Ca2+ in insulin-secreting β-cells', 118: 4463-71.

Thore, Sophia, Oleg Dyachok, and Anders %J Journal of Biological Chemistry Tengholm. 2004. 'Oscillations of phospholipase C activity triggered by depolarization and Ca2+ influx in insulin-secreting cells', 279: 19396-400.

Thornton, Kevin, and David G Gorenstein. 1994. 'Structure of glucagon-like peptide (7-36) amide in a dodecylphosphocholine micelle as determined by 2D NMR', *Biochemistry*, 33: 3532-39.

Tomas, E., V. Stanojevic, and J. F. Habener. 2011. 'GLP-1-derived nonapeptide GLP-1(28-36)amide targets to mitochondria and suppresses glucose production and oxidative stress in isolated mouse hepatocytes', *Regul Pept*, 167: 177-84.

Topf, M., M. L. Baker, M. A. Marti-Renom, W. Chiu, and A. Sali. 2006. 'Refinement of protein structures by iterative comparative modeling and CryoEM density fitting', *J Mol Biol*, 357: 1655-68.

Topf, M., K. Lasker, B. Webb, H. Wolfson, W. Chiu, and A. Sali. 2008. 'Protein structure fitting and refinement guided by cryo-EM density', *Structure*, 16: 295-307.

Trabuco, Leonardo G., Elizabeth Villa, Kakoli Mitra, Joachim Frank, and Klaus Schulten. 2008. 'Flexible Fitting of Atomic Structures into Electron Microscopy Maps Using Molecular Dynamics', *Structure*, 16: 673-83.

Trabuco, Leonardo G., Elizabeth Villa, Eduard Schreiner, Christopher B. Harrison, and Klaus Schulten. 2009. 'Molecular Dynamics Flexible Fitting: A practical guide to combine cryo-electron microscopy and X-ray crystallography', *Methods (San Diego, Calif.)*, 49: 174-80.

Turner, J. T., S. B. Jones, and D. B. Bylund. 1986. 'A fragment of vasoactive intestinal peptide, VIP(10-28), is an antagonist of VIP in the colon carcinoma cell line, HT29', *Peptides*, 7: 849-54.

Ugleholdt, R., X. Zhu, C. F. Deacon, C. Orskov, D. F. Steiner, and J. J. Holst. 2004. 'Impaired intestinal proglucagon processing in mice lacking prohormone convertase 1', *Endocrinology*, 145: 1349-55.

Underwood, Christina Rye, Patrick Garibay, Lotte Bjerre Knudsen, Sven Hastrup, Günther H. Peters, Rainer Rudolph, and Steffen Reedtz-Runge. 2010. 'Crystal Structure of Glucagon-like Peptide-1 in Complex with the Extracellular Domain of the Glucagon-like Peptide-1 Receptor', 285: 723-30.

Unson, Cecilia G. 2002. 'Molecular determinants of glucagon receptor signaling', *Peptide Science*, 66: 218-35.

Usdin, Ted B, Catherine Gruber, and Tom I Bonner. 1995. 'Identification and functional expression of a receptor selectively recognizing parathyroid hormone, the PTH2 receptor', *Journal of Biological Chemistry*, 270: 15455-58.

Vahl, Torsten P, Breay W Paty, Bradley D Fuller, Ronald L Prigeon, and David A D'Alessio. 2003. 'Effects of GLP-1-(7–36) NH2, GLP-1-(7–37), and GLP-1-(9–36) NH2 on intravenous glucose tolerance and glucose-induced insulin secretion in healthy humans', *The Journal of Clinical Endocrinology & Metabolism*, 88: 1772-79.

Valabhji, Professor Jonathan. 2018. "Type 2 diabetes and the importance of prevention." In *NHS Engalnd*, edited by NHS England. England.

van der Zijl, N. J., G. H. Goossens, C. C. Moors, D. H. van Raalte, M. H. Muskiet, P. J. Pouwels, E. E. Blaak, and M. Diamant. 2011. 'Ectopic fat storage in the pancreas, liver, and abdominal fat depots: impact on beta-cell function in individuals with impaired glucose metabolism', *J Clin Endocrinol Metab*, 96: 459-67.

van Meer, Gerrit, and Anton I. P. M. de Kroon. 2011. 'Lipid map of the mammalian cell', 124: 5-8.

van Meer, Gerrit, Dennis R. Voelker, and Gerald W. Feigenson. 2008. 'Membrane lipids: where they are and how they behave', *Nature Reviews Molecular Cell Biology*, 9: 112.

Vance, Dennis E., and Jean E. Vance. 2008. 'CHAPTER 8 - Phospholipid biosynthesis in eukaryotes.' in Dennis E. Vance and Jean E. Vance (eds.), *Biochemistry of Lipids, Lipoproteins and Membranes (Fifth Edition)* (Elsevier: San Diego).

Velazquez-Muriel, J. A., M. Valle, A. Santamaria-Pang, I. A. Kakadiaris, and J. M. Carazo. 2006. 'Flexible fitting in 3D-EM guided by the structural variability of protein superfamilies', *Structure*, 14: 1115-26.

Venkatakrishnan, A. J., Xavier Deupi, Guillaume Lebon, Franziska M. Heydenreich, Tilman Flock, Tamara Miljus, Santhanam Balaji, Michel Bouvier, Dmitry B.

Veprintsev, Christopher G. Tate, Gebhard F. X. Schertler, and M. Madan Babu. 2016. 'Diverse activation pathways in class A GPCRs converge near the G-protein-coupling region', *Nature*, 536: 484.

Verkleij, A. J., R. F. Zwaal, B. Roelofsen, P. Comfurius, D. Kastelijn, and L. L. van Deenen. 1973. 'The asymmetric distribution of phospholipids in the human red cell membrane. A combined study using phospholipases and freeze-etch electron microscopy', *Biochim Biophys Acta*, 323: 178-93.

Verlet, L. Computer experiments on classical fluids. ii. equilibrium correlation functions. Phys. Rev., 165:201–214, 1968.

Vilardaga, Jean-Pierre, Guillermo Romero, Peter A Friedman, and Thomas J Gardella. 2011. 'Molecular basis of parathyroid hormone receptor signaling and trafficking: a family B GPCR paradigm', *Cellular and Molecular Life Sciences*, 68: 1-13.

Villa, Leonardo Trabuco and Elizabeth. 2006. "Salt Bridges Plugin." In, The Salt Bridges plugin searches for salt bridges formed in a protein throughout a trajectory. The search can be restricted to a selection and/or a frame range given by the user.

Volkmann, N., D. Hanein, G. Ouyang, K. M. Trybus, D. J. DeRosier, and S. Lowey. 2000. 'Evidence for cleft closure in actomyosin upon ADP release', *Nat Struct Biol*, 7: 1147-55.

Vries, Luc De, Bin Zheng, Thierry Fischer, Eric Elenko, and Marilyn G. Farquhar. 2000. 'The Regulator of G Protein Signaling Family', *Annual Review of Pharmacology and Toxicology*, 40: 235-71.

Wang, Jia, Xianqiang Song, Dandan Zhang, Xiaoqing Chen, Xun Li, Yaping Sun, Cui Li, Yunpeng Song, Yao Ding, Ruobing Ren, Essa Hu Harrington, Liaoyuan A. Hu, Wenge Zhong, Cen Xu, Xin Huang, Hong-Wei Wang, and Yingli Ma. 2020. 'Cryo-EM structures of PAC1 receptor reveal ligand binding mechanism', *Cell Research*, 30: 436-45.

Wang, Jingjing, Min Yu, Jian Xu, Yusheng Cheng, Xiang Li, Guihong Wei, Hong Wang, Hui Kong, and Weiping Xie. 2019. 'Glucagon-like peptide-1 (GLP-1) mediates the protective effects of dipeptidyl peptidase IV inhibition on pulmonary hypertension', *Journal of Biomedical Science*, 26: 6.

Wang, Shuo, Jiayi Yu, Wei Li, and Fei Li. 2011. 'Structural study of an active analog of EX-4 in solution and micelle associated states', *Peptide Science*, 96: 348-57.

Warne, Tony, Rouslan Moukhametzianov, Jillian G. Baker, Rony Nehmé, Patricia C. Edwards, Andrew G. W. Leslie, Gebhard F. X. Schertler, and Christopher G. Tate. 2011. 'The structural basis for agonist and partial agonist action on a β1-adrenergic receptor', *Nature*, 469: 241-44.

Watanabe, Y., K. Kawai, S. Ohashi, C. Yokota, S. Suzuki, and K. Yamashita. 1994. 'Structure-activity relationships of glucagon-like peptide-1(7-36)amide: insulinotropic activities in perfused rat pancreases, and receptor binding and cyclic AMP production in RINm5F cells', *J Endocrinol*, 140: 45-52.

Waterhouse, Andrew, Martino Bertoni, Stefan Bienert, Gabriel Studer, Gerardo Tauriello, Rafal Gumienny, Florian T. Heer, Tjaart A P de Beer, Christine Rempfer, Lorenza Bordoli, Rosalba Lepore, and Torsten Schwede. 2018. 'SWISS-MODEL: homology modelling of protein structures and complexes', *Nucleic Acids Research*, 46: W296-W303.

Waterhouse, Andrew, Christine Rempfer, Florian T Heer, Gabriel Studer, Gerardo Tauriello, Lorenza Bordoli, Martino Bertoni, Rafal Gumienny, Rosalba Lepore, Stefan Bienert, Tjaart A P de Beer, and Torsten Schwede. 2018.

'SWISS-MODEL: homology modelling of protein structures and complexes', *Nucleic Acids Research*, 46: W296-W303.

Weaver, Richard E., Juan C. Mobarec, Mark J. Wigglesworth, Christopher A. Reynolds, and Dan Donnelly. 2017. 'High affinity binding of the peptide agonist TIP-39 to the parathyroid hormone 2 (PTH2) receptor requires the hydroxyl group of Tyr-318 on transmembrane helix 5', *Biochem Pharmacol*, 127: 71-81.

Weaver, Richard E., Mark J. Wigglesworth, and Dan Donnelly. 2014. 'A salt bridge between Arg-20 on parathyroid hormone (PTH) and Asp-137 on the PTH1 receptor is essential for full affinity', *Peptides*, 61: 83-87.

Wess, J. 1997a. 'G-protein-coupled receptors: molecular mechanisms involved in receptor activation and selectivity of G-protein recognition', *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 11: 346-54.

Wess, Jürgen. 1997b. 'G-protein-coupled receptors: molecular mechanisms involved in receptor activation and selectivity of G-protein recognition', *The FASEB Journal*, 11: 346-54.

Weston, C., D. Poyner, V. Patel, S. Dowell, and G. Ladds. 2014. 'Investigating G protein signalling bias at the glucagon-like peptide-1 receptor in yeast', *Br J Pharmacol*, 171: 3651-65.

Wettergren, A., L. Pridal, M. Wojdemann, and J. J. Holst. 1998. 'Amidated and non-amidated glucagon-like peptide-1 (GLP-1): non-pancreatic effects (cephalic phase acid secretion) and stability in plasma in humans', *Regul Pept*, 77: 83-7.

Wollmann, P., S. Cui, R. Viswanathan, O. Berninghausen, M. N. Wells, M. Moldt, G. Witte, A. Butryn, P. Wendler, R. Beckmann, D. T. Auble, and K. P. Hopfner. 2011. 'Structure and mechanism of the Swi2/Snf2 remodeller Mot1 in complex with its substrate TBP', *Nature*, 475: 403-7.

Wootten, D., H. Lindmark, M. Kadmiel, H. Willcockson, K. M. Caron, J. Barwell, T. Drmota, and D. R. Poyner. 2013. 'Receptor activity modifying proteins (RAMPs) interact with the VPAC2 receptor and CRF1 receptors and modulate their function', *Br J Pharmacol*, 168: 822-34.

Wootten, D., C. A. Reynolds, C. Koole, K. J. Smith, J. C. Mobarec, J. Simms, T. Quon, T. Coudrat, S. G. Furness, L. J. Miller, A. Christopoulos, and P. M. Sexton. 2016. 'A Hydrogen-Bonded Polar Network in the Core of the Glucagon-Like Peptide-1 Receptor Is a Fulcrum for Biased Agonism: Lessons from Class B Crystal Structures', *Mol Pharmacol*, 89: 335-47.

Wootten, D., C. A. Reynolds, K. J. Smith, J. C. Mobarec, C. Koole, E. E. Savage, K. Pabreja, J. Simms, R. Sridhar, S. G. B. Furness, M. Liu, P. E. Thompson, L. J. Miller, A. Christopoulos, and P. M. Sexton. 2016a. 'The Extracellular Surface of the GLP-1 Receptor Is a Molecular Trigger for Biased Agonism', *Cell*, 165: 1632-43.

Wootten, Denise, and Laurence J. Miller. 2020. 'Structural Basis for Allosteric Modulation of Class B G Protein–Coupled Receptors', *Annual Review of Pharmacology and Toxicology*, 60: 89-107.

Wootten, Denise, Christopher A. Reynolds, Kevin J. Smith, Juan C. Mobarec, Sebastian G. B. Furness, Laurence J. Miller, Arthur Christopoulos, and Patrick M. Sexton. 2016b. 'Key interactions by conserved polar amino acids located at the transmembrane helical boundaries in Class B GPCRs modulate activation, effector specificity and biased signalling in the glucagon-like peptide-1 receptor', *Biochem Pharmacol*, 118: 68-87.

Wootten, Denise, John Simms, Laurence J. Miller, Arthur Christopoulos, and Patrick M. Sexton. 2013. 'Polar transmembrane interactions drive formation of ligand-specific and signal pathway-biased family B G protein-coupled receptor conformations', 110: 5211-16.

Wriggers, W., R. K. Agrawal, D. L. Drew, A. McCammon, and J. Frank. 2000. 'Domain motions of EF-G bound to the 70S ribosome: insights from a hand-shaking between multi-resolution structures', *Biophys J*, 79: 1670-8.

Wriggers, W., and S. Birmanns. 2001. 'Using situs for flexible and rigid-body fitting of multiresolution single-molecule data', *J Struct Biol*, 133: 193-202.

Wriggers, Willy. 2012. 'Conventions and workflows for using Situs', *Acta Crystallogr D Biol Crystallogr*, 68: 344-51.

Wriggers, Willy, and Pablo Chacón. 2001. 'Modeling Tricks and Fitting Techniques for Multiresolution Structures', *Structure*, 9: 779-88.

Wu, Emilia L., Xi Cheng, Sunhwan Jo, Huan Rui, Kevin C. Song, Eder M. Dávila-Contreras, Yifei Qi, Jumin Lee, Viviana Monje-Galvan, Richard M. Venable, Jeffery B. Klauda, and Wonpil Im. 2014. 'CHARMM-GUI Membrane Builder toward realistic biological membrane simulations', 35: 1997-2004.

Wu, Fan, Linlin Yang, Kaini Hang, Mette Laursen, Lijie Wu, Gye Won Han, Qiansheng Ren, Nikolaj Kulahin Roed, Guangyao Lin, Michael A. Hanson, Hualiang Jiang, Ming-Wei Wang, Steffen Reedtz-Runge, Gaojie Song, and Raymond C. Stevens. 2020. 'Full-length human GLP-1 receptor structure without orthosteric ligands', *Nat Commun*, 11: 1272.

Wu, Huixian, Chong Wang, Karen J Gregory, Gye Won Han, Hyekyung P Cho, Yan Xia, Colleen M Niswender, Vsevolod Katritch, Jens Meiler, and Vadim %J Science Cherezov. 2014. 'Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator', 344: 58-64.

Wuttke, Anne, Jenny Sågetorp, and Anders %J J Cell Sci Tengholm. 2010. 'Distinct plasma-membrane PtdIns (4) P and PtdIns (4, 5) P2 dynamics in secretagogue-stimulated β-cells', 123: 1492-502.

Xiao, Q, J Giguere, M Parisien, W Jeng, SA St-Pierre, PL Brubaker, and MB Wheeler. 2001. 'Biological activities of glucagon-like peptide-1 analogues in vitro and in vivo', *Biochemistry*, 40: 2860-69.

Xiao, Q., W. Jeng, and M. B. Wheeler. 2000. 'Characterization of glucagon-like peptide-1 receptor-binding determinants', *J Mol Endocrinol*, 25: 321-35.

Xie, Beichen, Phuoc My Nguyen, Alenka Guček, Antje Thonig, Sebastian Barg, and Olof Idevall-Hagren. 2016. 'Plasma Membrane Phosphatidylinositol 4,5-Bisphosphate Regulates Ca2+-Influx and Insulin Secretion from Pancreatic β Cells', *Cell Chemical Biology*, 23: 816-26.

Xu, Y., Y. Wang, Y. Wang, K. Liu, Y. Peng, D. Yao, H. Tao, H. Liu, and G. Song. 2019. 'Mutagenesis facilitated crystallization of GLP-1R', *IUCrJ*, 6: 996-1006.

Yang, Dehua, Chris de Graaf, Linlin Yang, Gaojie Song, Antao Dai, Xiaoqing Cai, Yang Feng, Steffen Reedtz-Runge, Michael A. Hanson, Huaiyu Yang, Hualiang Jiang, Raymond C. Stevens, and Ming-Wei Wang. 2016. 'Structural Determinants of Binding the Seven-transmembrane Domain of the Glucagon-like Peptide-1 Receptor (GLP-1R)', 291: 12991-3004.

Yang, L., D. Yang, C. de Graaf, A. Moeller, G. M. West, V. Dharmarajan, C. Wang, F. Y. Siu, G. Song, S. Reedtz-Runge, B. D. Pascal, B. Wu, C. S. Potter, H. Zhou, P. R. Griffin, B. Carragher, H. Yang, M. W. Wang, R. C. Stevens, and H. Jiang. 2015a. 'Conformational states of the full-length glucagon receptor', *Nat Commun*, 6: 7859.

Yang, Linlin, Dehua Yang, Chris de Graaf, Arne Moeller, Graham M. West, Venkatasubramanian Dharmarajan, Chong Wang, Fai Y. Siu, Gaojie Song, Steffen Reedtz-Runge, Bruce D. Pascal, Beili Wu, Clinton S. Potter, Hu Zhou, Patrick R. Griffin, Bridget Carragher, Huaiyu Yang, Ming-Wei Wang, Raymond C. Stevens, and Hualiang Jiang. 2015b. 'Conformational states of the full-length glucagon receptor', *Nat Commun*, 6: 7859.

Yano, M., K. Watanabe, T. Yamamoto, K. Ikeda, T. Senokuchi, M. Lu, T. Kadomatsu, H. Tsukano, M. Ikawa, M. Okabe, S. Yamaoka, T. Okazaki, H. Umehara, T. Gotoh, W. J. Song, K. Node, R. Taguchi, K. Yamagata, and Y. Oike. 2011. 'Mitochondrial dysfunction and increased reactive oxygen species impair insulin secretion in sphingomyelin synthase 1-null mice', *J Biol Chem*, 286: 3992-4002.

Yao, X. J., G. Velez Ruiz, M. R. Whorton, S. G. Rasmussen, B. T. DeVree, X. Deupi, R. K. Sunahara, and B. Kobilka. 2009. 'The effect of ligand efficacy on the formation and stability of a GPCR-G protein complex', *Proc Natl Acad Sci U S A*, 106: 9501-6.

Yao, Z., and B. Kobilka. 2005. 'Using synthetic lipids to stabilize purified beta2 adrenoceptor in detergent micelles', *Anal Biochem*, 343: 344-6.

Yin, J., J. C. Mobarec, P. Kolb, and D. M. Rosenbaum. 2015. 'Crystal structure of the human OX2 orexin receptor bound to the insomnia drug suvorexant', *Nature*, 519: 247-50.

Yin, Yanting, Parker W de Waal, Yuanzheng He, Li-Hua Zhao, Dehua Yang, Xiaoqing Cai, Yi Jiang, Karsten Melcher, Ming-Wei Wang, and H Eric Xu. 2017. 'Rearrangement of a polar core provides a conserved mechanism for constitutive activation of class BG protein-coupled receptors', *Journal of Biological Chemistry*, 292: 9865-81.

Yokomori, Hiroaki, and Wataru Ando. 2020. 'Spatial expression of glucagon-like peptide 1 receptor and caveolin-1 in hepatocytes with macrovesicular steatosis in non-alcoholic steatohepatitis', *BMJ Open Gastroenterology*, 7: e000370.

Zaretsky, Joseph Z., and Daniel H. Wreschner. 2008. 'Protein multifunctionality: principles and mechanisms', *Translational oncogenomics*, 3: 99-136.

Zhang, Haonan, Anna Qiao, Dehua Yang, Linlin Yang, Antao Dai, Chris de Graaf, Steffen Reedtz-Runge, Venkatasubramanian Dharmarajan, Hui Zhang, Gye Won Han, Thomas D. Grant, Raymond G. Sierra, Uwe Weierstall, Garrett Nelson, Wei Liu, Yanhong Wu, Limin Ma, Xiaoqing Cai, Guangyao Lin, Xiaoai Wu, Zhi Geng, Yuhui Dong, Gaojie Song, Patrick R. Griffin, Jesper Lau, Vadim Cherezov, Huaiyu Yang, Michael A. Hanson, Raymond C. Stevens, Qiang Zhao, Hualiang Jiang, Ming-Wei Wang, and Beili Wu. 2017. 'Structure of the full-length glucagon class B G-protein-coupled receptor', *Nature*, 546: 259.

Zhang, Haonan, Anna Qiao, Linlin Yang, Ned Van Eps, Klaus S. Frederiksen, Dehua Yang, Antao Dai, Xiaoqing Cai, Hui Zhang, Cuiying Yi, Can Cao, Lingli He, Huaiyu Yang, Jesper Lau, Oliver P. Ernst, Michael A. Hanson, Raymond C. Stevens, Ming-Wei Wang, Steffen Reedtz-Runge, Hualiang Jiang, Qiang Zhao, and Beili Wu. 2018. 'Structure of the glucagon receptor in complex with a glucagon analogue', *Nature*, 553: 106.

Zhang, Jintu, Qifeng Bai, Horacio Pérez-Sánchez, Shuxia Shang, Xiaoli An, and Xiaojun Yao. 2019. 'Investigation of ECD conformational transition mechanism of GLP-1R by molecular dynamics simulations and Markov state model', *Physical Chemistry Chemical Physics*, 21: 8470-81.

Zhang, Xin, Matthew J. Belousoff, Peishen Zhao, Albert J. Kooistra, Tin T. Truong, Sheng Yu Ang, Christina Rye Underwood, Thomas Egebjerg, Petr Šenel, Gregory D. Stewart, Yi-Lynn Liang, Alisa Glukhova, Hari Venugopal, Arthur Christopoulos, Sebastian G. B. Furness, Laurence J. Miller, Steffen Reedtz-Runge, Christopher J. Langmead, David E. Gloriam, Radostin Danev, Patrick M. Sexton, and Denise Wootten. 2020a. 'Differential GLP-1R binding and activation by peptide and non-peptide agonists', *bioRxiv*: 2020.08.16.252585.

———. 2020b. 'Differential GLP-1R Binding and Activation by Peptide and Non-peptide Agonists', *Molecular Cell*, 80: 485-500.e7.

Zhang, Xin, Rachel M. Johnson, Ieva Drulyte, Lingbo Yu, Abhay Kotecha, Radostin Danev, Denise Wootten, Patrick M. Sexton, and Matthew J. Belousoff. 'Evolving cryo-EM structural approaches for GPCR drug discovery', *Structure*.

Zhang, Yan, Bingfa Sun, Dan Feng, Hongli Hu, Matthew Chu, Qianhui Qu, Jeffrey T. Tarrasch, Shane Li, Tong Sun Kobilka, Brian K. Kobilka, and Georgios Skiniotis. 2017. 'Cryo-EM structure of the activated GLP-1 receptor in complex with a G protein', *Nature*, 546: 248.

Zhao, Li-Hua, Shanshan Ma, Ieva Sutkeviciute, Dan-Dan Shen, X. Edward Zhou, Parker W. de Waal, Chen-Yao Li, Yanyong Kang, Lisa J. Clark, Frederic G. Jean-Alphonse, Alex D. White, Dehua Yang, Antao Dai, Xiaoqing Cai, Jian Chen, Cong Li, Yi Jiang, Tomoyuki Watanabe, Thomas J. Gardella, Karsten Melcher, Ming-Wei Wang, Jean-Pierre Vilardaga, H. Eric Xu, and Yan Zhang. 2019. 'Structure and dynamics of the active human parathyroid hormone receptor-1', *Science*, 364: 148-53.

Zhao, Li-Hua, Yanting Yin, Dehua Yang, Bo Liu, Li Hou, Xiaoxi Wang, Kuntal Pal, Yi Jiang, Yang Feng, and Xiaoqing Cai. 2016. 'Differential requirement of the extracellular domain in activation of class BG protein-coupled receptors', *Journal of Biological Chemistry*, 291: 15119-30.

Zhao, Peishen, Yi-Lynn Liang, Matthew J. Belousoff, Giuseppe Deganutti, Madeleine M. Fletcher, Francis S. Willard, Michael G. Bell, Michael E. Christe, Kyle W. Sloop, Asuka Inoue, Tin T. Truong, Lachlan Clydesdale, Sebastian G. B. Furness, Arthur Christopoulos, Ming-Wei Wang, Laurence J. Miller, Christopher A. Reynolds, Radostin Danev, Patrick M. Sexton, and Denise Wootten. 2020. 'Activation of the GLP-1 receptor by a non-peptidic agonist', *Nature*.

Zhao Fenghui, Zhou Qingtong, Cong Zhaotong, Hang Kaini Hang, ZouXinyu Zou, Zhang Chao Zhang, Chen Yan, Dai Antao, Anyi Liang, Qianqian Ming, Mu Wang, Linan Chen, Peiyu Xu, Rulue Chang, Wenbo Feng, Tian Xia , Yan

Zhang , Beili Wu, Dehua Yang, Lihua Zhao, H. Eric Xu, Ming-Wei Wang. 2021. 'Structural basis for the therapeutic advantage of dual and triple agonists at the human GIP, GLP-1 or GCG receptors', preprint bioRxiv https://doi.org/10.1101/2021.07.29.454286

Zhou, Fulai, Huibing Zhang, Zhaotong Cong, Li-Hua Zhao, Qingtong Zhou, Chunyou Mao, Xi Cheng, Dan-Dan Shen, Xiaoqing Cai, Cheng Ma, Yuzhe Wang, Antao Dai, Yan Zhou, Wen Sun, Fenghui Zhao, Suwen Zhao, Hualiang Jiang, Yi Jiang, Dehua Yang, H. Eric Xu, Yan Zhang, and Ming-Wei Wang. 2020. 'Structural basis for activation of the growth hormone-releasing hormone receptor', *Nat Commun*, 11: 5205.

Zhu, X., A. Zhou, A. Dey, C. Norrbom, R. Carroll, C. Zhang, V. Laurent, I. Lindberg, R. Ugleholdt, J. J. Holst, and D. F. Steiner. 2002. 'Disruption of PC1/3

expression in mice causes dwarfism and multiple neuroendocrine peptide processing defects', *Proc Natl Acad Sci U S A*, 99: 10293-8.

Zhuang, Xiaohong, Judah R. Makover, Wonpil Im, and Jeffery B. Klauda. 2014. 'A systematic molecular dynamics simulation study of temperature dependent bilayer structural properties', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1838: 2520-29.

Zocher, M., C. Zhang, S. G. Rasmussen, B. K. Kobilka, and D. J. Muller. 2012. 'Cholesterol increases kinetic, energetic, and mechanical stability of the human beta2-adrenergic receptor', *Proc Natl Acad Sci U S A*, 109: E3463-72.