

# **Structural Mechanism of Activation of G Protein-Coupled Receptors**

Ph.D. Thesis

**Argha Mitra**

Szeged

2021

# **Structural Mechanism of Activation of G Protein-Coupled Receptors**

Ph.D. Thesis

**Argha Mitra**

*Laboratory of Chemical Biology, Institute of Biochemistry,  
Biological Research Centre, Szeged*

Supervisor

**Dr. Attila Borics**

*Doctoral School of Theoretical Medicine  
Albert Szent-Györgyi Medical School, University of Szeged*

Szeged

2021

## Publications Related to the Thesis

- I. **Mitra, A** and Sarkar, A, Szabó, MR and Borics, A.

*Correlated motions of conserved polar motifs lay out a plausible mechanism of G Protein-Coupled Receptor activation.*

BIOMOLECULES, 11 (5), 670 (2021).

doi: 10.3390/biom11050670.

IF: 4.879 [Q2]

- II. **Mitra, A** and Sarkar, and Borics, A.

*Universal properties and specificities of the  $\beta_2$ -adrenergic receptor- $G_s$  protein complex activation mechanism revealed by all-atom molecular dynamics simulations.*

INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES, 22(19), 10423 (2021).

doi: 10.3390/ijms221910423

IF: 5.923 [Q1]

## Acknowledgements

First and foremost, I want to extend my heartfelt gratitude towards my advisor, Dr. Attila Borics for his support, patient guidance, encouragement, constructive criticism throughout this journey.

I would like to thank Dr. György Pósfai, former Director the of Institute of Biochemistry, for accepting me to the Institute as a PhD student. Furthermore, I extend this gratitude to Dr. Péter Horváth, the present Director for his ongoing support and providing a fellowship to carry out my PhD research.

I would like to thank our former lab member Márton Richárd Szabó and present member Arijit Sarkar for their help and support.

I would like to thank Dr. Csaba Tömböly and other members from the Laboratory of Chemical Biology for their kind support and assistance.

I would like to acknowledge the collaborators Dr. Szabolcs Dvorácskó, Dr. Balázs István Tóth, Dr. Zoltán Lipinszki and Dr. Gabriel Fenteany for providing me with opportunity to work in different projects.

Thanks to Basu, Anushree, Mouboni, Manna, Eszter, Kewin for their suggestions and encouragements.

I would like to express my deepest gratitude to my mother Swarnalata Mitra, and my father Chanchal Mitra for their unparalleled support and cheering me up for countless ways and to my uncle Mrinal Ghosh for unwavering suggestions.



# Table of Contents

1. Introduction .....	1
1.1 Proteins.....	1
1.1.1. General Properties of Proteins.....	1
1.1.2. Physical Properties and Structure of Proteins.....	1
1.2. G Protein-Coupled Receptors (GPCRs).....	4
1.2.1. Biological role of GPCRs .....	4
1.2.2. Classification of GPCRs.....	4
1.2.3. GPCR Ligands and Therapeutic Applications.....	5
1.2.4. Interacting Partners and Downstream Signaling.....	6
1.3. The Structure of GPCRs.....	7
1.3.1. Domain Architecture of GPCRs.....	7
1.3.2. Experimental Investigations of GPCR Structure.....	9
1.3.3. Limitations of Experiments .....	12
1.3.4. Computational Studies of GPCRs .....	13
1.3.5. Limitations of Computational Techniques and Studies .....	16
2. Aims and Scope.....	18
3. Materials and Methods.....	19
3.1. System Building.....	19
3.1.1. Obtaining Atomic Resolution Starting Structures.....	19
3.1.2. Modeling and Attachment of Missing Domains and Residues.....	20
3.2.3. Post-Translational Modifications.....	20
3.1.4. Ligand Docking.....	21
3.2.5. Membrane Building and the Insertion of Ligand-Receptor Complexes.....	22
3.2. Energy Minimization and Equilibration .....	23

3.3. Production Simulations.....	23
3.4. MD Trajectory Analysis.....	25
3.5. Sequence Alignment and Conservation Analysis.....	26
4. Results and Discussion.....	27
4.1. Membrane Properties .....	27
4.2. Simulation System Integrity .....	28
4.2.1. N- and C-terminal Domain Dynamics.....	28
4.2.2. Dissociation of Molecular Parts .....	30
4.3. Allosteric Na <sup>+</sup> Binding .....	31
4.4. TM6 Dynamics.....	33
4.5. Loop Dynamics and Intermolecular Interactions .....	34
4.6. Intramolecular Interactions.....	39
4.7. Transmembrane Helix Properties.....	41
4.8. NPxxY Motif Dynamics.....	42
4.8. Correlated Side-chain Motions in the Transmembrane Domain.....	42
5. Conclusions.....	49
6. Summary of findings .....	49
References.....	51

## List of Abbreviations

<b>2-AG</b>	2-arachidonoylglycerol
<b>5-HT<sub>2B</sub></b>	5-hydroxytryptamine receptor-2B
<b>aMD</b>	Accelerated molecular dynamics
<b>cAMP/cGMP</b>	adenylyl or guanylyl cyclase
<b>CB1</b>	type 1 cannabinoid receptor
<b>CB2</b>	type 2 cannabinoid receptor
<b>CL</b>	C-loop
<b>Cryo-EM</b>	cryogenic electron microscopy
<b>DCCM</b>	dynamic cross-correlation matrix
<b>DOP</b>	$\delta$ -opioid receptor
<b>ECL</b>	extracellular loop
<b>EM2</b>	endomorphin 2
<b>FL</b>	finger loop
<b>GDP</b>	guanosine diphosphate
<b>GPCR</b>	G protein-coupled receptor
<b>GRK</b>	G protein-coupled receptor kinases
<b>GTP</b>	guanosine triphosphate
<b>H5</b>	$\alpha$ 5 helix or helix 5 of the G $\alpha$ protein
<b>H8</b>	cytosolic helix or helix 8
<b>ICL</b>	intracellular loop
<b>KOP</b>	kappa opioid receptor
<b>M2R</b>	type 2 muscarinic receptor
<b>MD</b>	molecular dynamics
<b>ML</b>	middle loop
<b>MOP</b>	$\mu$ -opioid receptor
<b>Nb</b>	nanobody

<b>NMR</b>	nuclear magnetic resonance
<b>NTS1</b>	type 1 neurotensin receptor
<b>PME</b>	particle-mesh Ewald summation
<b>PTM</b>	Post translational modification
<b>RMSD</b>	root mean square deviation
<b>suMD</b>	supervised molecular dynamics
<b>TM</b>	transmembrane
<b><math>\alpha_{1B}</math>AR</b>	$\alpha_{1B}$ adrenergic receptor
<b><math>\beta_2</math>AR</b>	$\beta_2$ adrenergic receptor
<b><math>\beta_1</math>AR</b>	$\beta_1$ adrenergic receptor

# 1. Introduction

## 1.1. Proteins

### 1.1.1. General Properties of Proteins

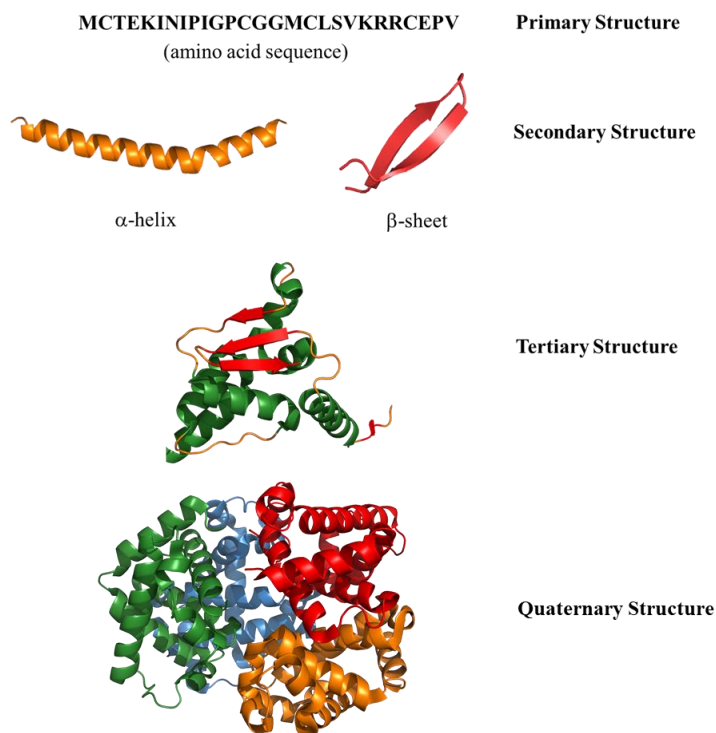
Proteins are versatile macromolecules and serve crucial functions in essentially all biological processes. They transport and store other molecules, catalyze chemical reactions, provide mechanical support and immune protection, generate movement, transmit nerve impulses and control growth and differentiation. Proteins are built from  $\alpha$ -amino acids through enzymatic polymerization. An  $\alpha$ -amino acid consists of a central carbon atom, referred to as the  $\alpha$ -carbon, which is linked to a primary amino group (except for proline), a carboxyl group, a hydrogen atom, and a variable side chain. Proteins are generally classified on the basis of their structure and function. Considering their shape and overall morphology, proteins can be either globular or fibrillar. Globular proteins are folded to form a compact, often spherical architecture, whereas fibrillar proteins form filamentous and sheet-like structures. Alternatively, proteins could be classified as structural or functional. The primary role of structural proteins is to provide mechanical support, while functional proteins execute and regulate cellular processes. Conjugated proteins contain non-protein components or prosthetic groups such as lipids, carbohydrates, nucleic acids, or small molecule co-factors. Simple proteins consist entirely of amino acids. Depending on the number of subunits proteins can be monomeric (single continuous chain) and multimeric (more than protein chains) form. After ribosomal biosynthesis, the polypeptide chains fold and adopt specific structures necessary for their biological function. Further processing of the folded protein may occur through post-translational modifications (PTMs). There are several known PTMs such as phosphorylation, acetylation, glycosylation, lipidation, methylation and ubiquitination. These modifications can alter the activity of proteins through the modulation of protein-protein interactions.

### 1.1.2. Physical Properties and Structure of Proteins

The three-dimensional structure and physico-chemical properties of proteins depend on those of the constituent amino acids. Amino acids are most often classified on the basis of the

chemical properties of their side chains. According to this classification, amino acids are either polar (see below) or nonpolar (glycine, Gly, G; alanine, Ala, A; valine, Val, V; leucine, Leu, L; isoleucine, Ile, I; proline, Pro, P; phenylalanine, Phe, F; tryptophan, Trp, W; and methionine, Met, M). The group of polar amino acids can be further divided into acidic (aspartic acid, Asp, D; and glutamic acid, Glu, E), basic (lysine, Lys, K; arginine, Arg, R; and histidine, His, H), or neutral amino acids (serine, Ser, S; threonine, Thr, T; asparagine, Asn, N; glutamine, Gln, Q; cysteine, Cys, C, and tyrosine, Tyr, Y). Amino acids can also be sorted based on the hydrophobicity of their side chains, differentiating between hydrophilic (R, N, Q, D, E, S, T, C, H), hydrophobic (A, V, L, I, F, P) and amphipathic (K, Y, M, W). In this latter classification G forms, its own class.

The amino acid side chains may vary in further physico-chemical parameters such as aromaticity, bulk, conformational flexibility, ability to cross-link, ability to form hydrogen bonds and chemical reactivity. Therefore, important protein attributes such as structural stability, hydrophilicity, pH dependence of function, to name a few, depend highly on amino acid compositions. Disulfide bridges, formed between two C residues and salt bridges, formed between ionizable side chains bearing opposite charge are the two most common primary bonds that stabilize the three-dimensional structure of proteins. Intramolecular hydrogen bonds play a crucial role in stabilizing extended, periodic protein structures. Polar interactions of permanent dipole moments of non-ionizable groups, dispersion and Van der Waals forces and weakly polar interactions of aromatic quadrupoles provide further stabilization to local structural elements. The folding of proteins is proposed to be driven by the so-called hydrophobic effect which is, by definition, the tendency of hydrophobic amino acid side chains to aggregate in aqueous solutions in order to decrease their surface area and consequently, the number of energetically unfavorable contacts with water molecules. As a result, globular protein structures form which contain hydrophilic amino acids on the surface and bury hydrophobic amino acids in the inner regions, inaccessible to water. The native structures of proteins maintain some degree of flexibility and they can be featured as an ensemble of interconverting structural states. This structural elasticity is most often associated with the function of the protein. The structures of proteins are marginally stable owing to the fact that the final, native structure is a thermodynamic trade-off with regard to the enthalpy and entropy change that accompany the folding process.



**Figure 1.** Hierarchic classification of protein structure.

The function of proteins determined by their three-dimensional structures, which is encoded by the corresponding amino acid sequences. Protein structures have been traditionally classified at four levels of organization. The primary structure refers to the amino acid sequence, whereas the secondary structure classifies ordered local structures of relatively short segments. The tertiary structure describes the overall three-dimensional structure of a single polypeptide chain, and the relative orientation of polypeptide subunits of a multi-chain protein is given by its quaternary structure (Figure 1). The three-dimensional structure of proteins is defined by the local conformations of their polypeptide backbones. Since peptide bonds are inherently planar and rigid, local structures are formed through the rotation around the single  $N-C_{\alpha}$  and  $C_{\alpha}-C(O)$  bonds of the polypeptide backbone and stabilized by the above-mentioned intramolecular interactions. In this regard, local secondary structural elements such as helices ( $\alpha$ -helix,  $3_{10}$ -helix,  $\pi$ -helix, etc.), parallel or antiparallel  $\beta$ -pleated sheets,  $\beta$ -turns and  $\gamma$ -turns are most frequently stabilized through H-bonds between H-bond donors and acceptors of the polypeptide backbone (amide N-H and C=O groups), whereas the role of interactions between the corresponding amino acid side chains is more

pronounced in stabilizing tertiary and quaternary structures. Certain combinations of secondary structural elements form different domains and motifs. Motifs are short, conserved sequence patterns and corresponding three dimensional structures which provides the protein with a complex and specific interaction surface and a distinctive function of that particular macromolecular part (Zn-finger, leucine-zipper, EF-hand, etc.). A domain, on the other hand, is traditionally defined as an independent, autonomous folding unit of the protein sequence that can also function independently of the rest of the protein.

## 1.2. G Protein-Coupled Receptors (GPCRs)

### 1.2.1. Biological role of GPCRs

Cell signaling is a vital process required for the growth and the development of cells and to sustain various environmental stress conditions. Membrane proteins are present on the surface of cells and serve as communication interfaces between extracellular and intracellular environments. A diverse array of intracellular processes initiated in response of primary stimuli by the external binding of ions, photons, hormones, neurotransmitters, small organic molecules or the external surface of another cell, to name a few. In the primary response to this stimulus, membrane proteins form a “signaling triad” consisting of a receptor, transducer and an effector. Such triads are involved in complex cytosolic signaling networks that control the activation of transcription factors that regulate gene expression and consequently induce cellular response. The largest and most diverse membrane protein family is G protein-coupled receptors (GPCRs). These proteins play essential roles in physiological cell homeostasis and also involved in a great number of pathophysiological conditions.

### 1.2.2. Classification of GPCRs

GPCRs are the largest protein superfamily of the human genome and are subdivided into several classes based on their amino acid sequences and functional similarities. According to one of the frequently used classifications, GPCRs are grouped into six classes, known as the A-F system (Brogi et al. 2014, Kolakowski 1994). Class A, also known as the rhodopsin-like family, is the largest group of GPCRs, which includes hormone, neurotransmitter, and light receptors and



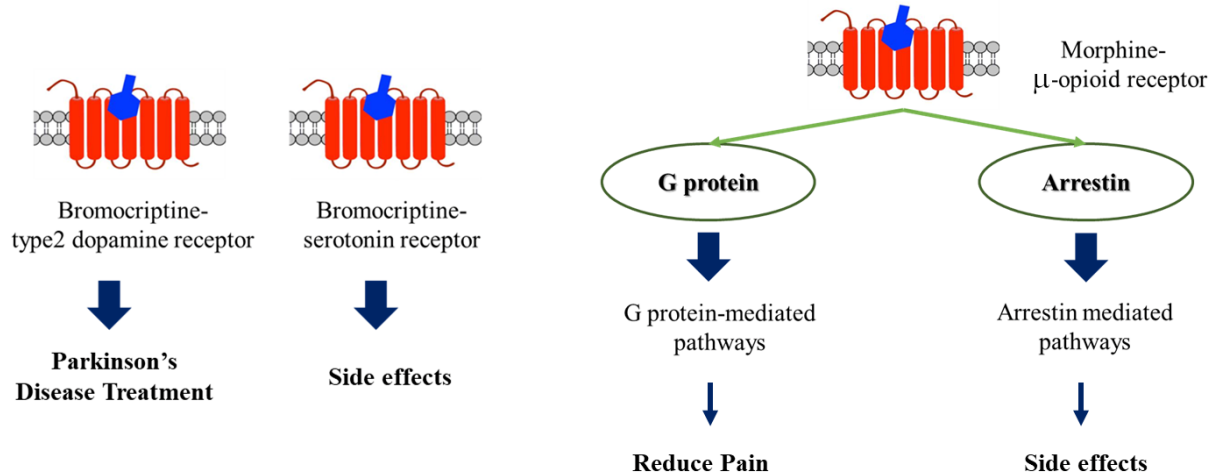
accounts for approximately 80% of GPCRs. Class B includes secretin receptors, Class C contains the metabotropic glutamate family, GABA receptors, calcium-sensing receptors, and taste receptors. Class D includes fungal mating pheromone receptors, whereas Class E includes cAMP receptors and Class F includes frizzled/smoothed receptors.

Another classification of GPCRs called the GRAFS system has also been proposed based on the phylogenetic tree of approximately 800 human GPCR sequences (Schiöth et al. 2004). This system contains five main families, namely the Glutamate (G), Rhodopsin (R), Adhesion (A), Frizzled/Taste2 (F), and Secretin (S) families.

### **1.2.3. GPCR Ligands and Therapeutic Applications**

An established pharmacological strategy of treating pathophysiological conditions is to generate alternative cellular responses through hijacking, blocking or inducing the function of GPCRs by suitable external agents, including ions, small organic molecules, hydrophobic odorants, amines, peptides, lipids, nucleotides and photons. GPCRs are implicated in a wide variety of human pathophysiological conditions such as cancer, pain, migraine, metabolic diseases and inflammation to name a few. Due to their accessibility at the cell surface, GPCRs are easily targeted by external therapeutic agents. Between 2011 and 2015, the proportion of FDA approved drugs that target GPCRs was approximately 27% of all marketed prescription pharmaceuticals. According to a recent report, this proportion have increased to 34% by 2017 (Hauser et al. 2017).

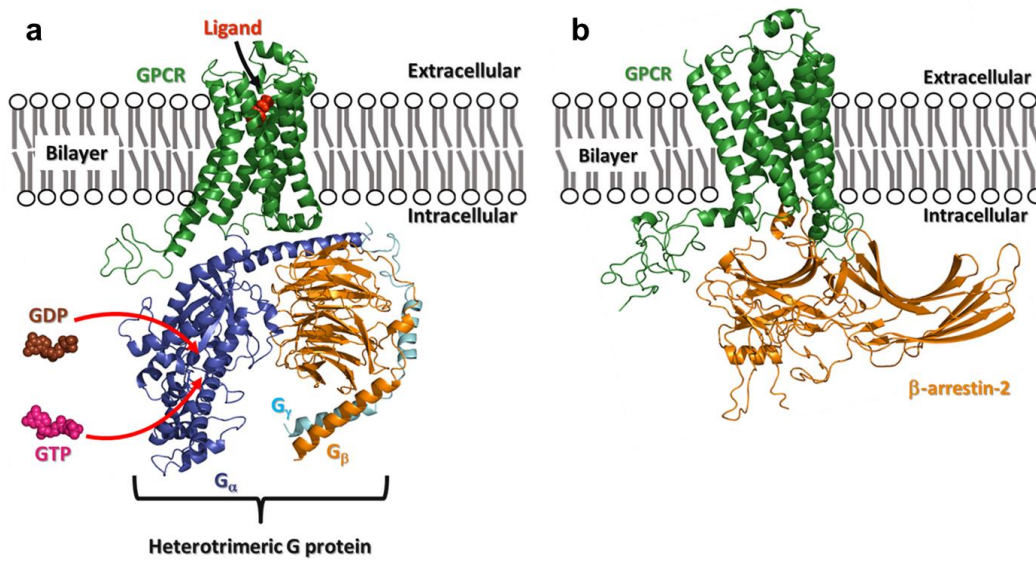
A major drawback of several important GPCR-targeting therapeutics is that their application is often limited by severe side effects. Such side effects are originating from non-selective, simultaneous activation of multiple signaling pathways (Figure 2). One example is the  $\mu$ -opioid receptor (MOP) of which external agonist, morphine, is still widely used in clinical practice to suppress chronic pain. The analgesic effect of morphine is exerted through the  $G_i$  protein-mediated pathway, but the arrestin-mediated pathway simultaneously activated upon morphine binding leads to severe, life-threatening side effects, such as respiratory depression. This necessitates the development of a new generation of GPCR drugs with improved receptor and pathway specificity. To do so, the structural mechanism of GPCR activation has to be elucidated in fine details.



**Figure 2.** The origin of side effects of GPCR drugs: non-selective activation (left panel) and activation of multiple pathways (right panel).

#### 1.2.4. Intracellular Interacting Partners and Downstream Signaling

GPCR signaling is mediated by heterotrimeric G proteins (Figure 3). Heterotrimeric G proteins are composed of three subunits  $G\alpha$ ,  $G\beta$  and  $G\gamma$ . The  $G\alpha$  subunit contains a guanosine triphosphate/ guanosine triphosphate (GDP/GTP) binding site and has GTPase activity. In the resting state this nucleotide binding site is occupied by GDP. Upon allosteric activation by the agonist bound active GPCR the GDP gets released and rapidly replaced by cytosolic GTP. The heterotrimeric complex dissociates from the GPCR followed by the dissociation of the  $G\alpha$  subunit from the  $G\beta\gamma$  dimer and consequent binding and activation of downstream effectors. G protein signaling is terminated by the hydrolysis of GTP to GDP by the  $G\alpha$  subunit and consequent re-association of the three subunits. It is debated if the G protein complex is recruited only upon GPCR activation or if it is pre-coupled with the inactive, ligand-free GPCR (de Oliveira et al. 2019). In contrast the large number of different GPCRs (~800), there are only four heterotrimeric G protein families ( $G_s$ ,  $G_{i/o}$ ,  $G_{q/11}$  and  $G_{12/13}$ ) with approximately 20 different  $G\alpha$  subunits (Strathmann et al. 1991). The diversity of  $G\beta$  and  $G\gamma$  subunits is even lower. This limited number of G protein families suggest that although they bind to very different GPCRs, the G protein-mediated signal transduction follows a conserved, general mechanism. In addition, elevated concentration of  $Na^+$  ions inhibit signaling of class A GPCRs through an allosteric mechanism (Pert et al. 1973). Arrestins are also a small family of proteins, of which primary role



**Figure 3.** A GPCR bound to heterotrimeric G protein (a) and arrestin (b).

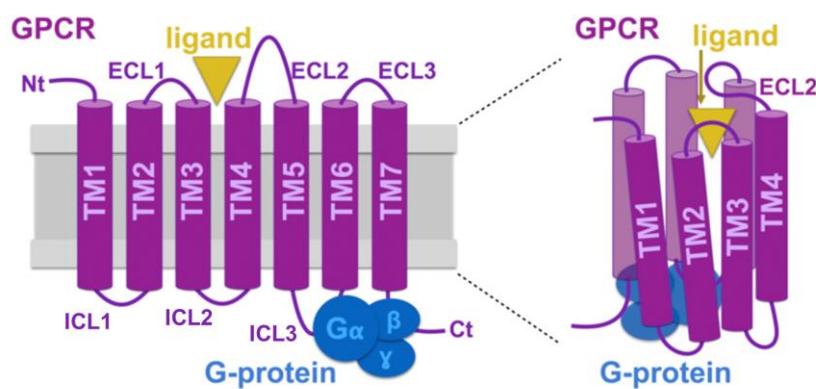
is the regulation of GPCR signaling.  $\beta$ -Arrestin-1 (arrestin 2) and  $\beta$ -arrestin-2 (arrestin 3) are the most frequent regulatory partners of GPCRs, whereas visual arrestin (arrestin 1) is specific to rhodopsin (Figure 3). These cytosolic proteins desensitize cognate GPCRs via a two-step mechanism. Active state GPCRs are flagged for arrestin binding by G protein-coupled receptor kinases (GRKs) through the phosphorylation of specific Ser and/or Thr residues on the cytosolic surface of the receptors. This is followed by arrestin binding which blocks G protein re-coupling and promotes internalization of GPCRs into clathrin-coated vesicles and subsequent trafficking, degradation or recycling of the receptors. In addition to their role in desensitization,  $\beta$ -arrestin-1 and  $\beta$ -arrestin-2 are involved in activating a rapidly expanding list of signaling pathways, separate from the G protein-mediated ones (Lohse et al. 1992, Lefkowitz et al. 2005, Moore et al. 2007).

## 1.3. The Structure of GPCRs

### 1.3.1. Domain Architecture of GPCRs

GPCRs consist of an N-terminal extracellular domain, a C-terminal intracellular domain and a transmembrane (TM) domain formed by a bundle of seven  $\alpha$ -helices (TM1-TM7), which

are connected by extracellular (ECL1-ECL3) and intracellular loops (ICL1-ICL3) of various lengths. In addition, the TM domain contains a membrane-parallel amphipathic helix/cytosolic helix, located between TM7 and the C-terminal domain, on the intracellular surface of the receptors (Figure 4). The N-terminal domain is the most diverse part of GPCRs regarding its size (sequence length) and structure. The adhesion family receptors have the largest N-terminal domains (up to 3000 residues), followed by the glutamate or glycoprotein hormone family (350-600 residues), while receptors binding neuropeptides or monoamine neurotransmitters have very short N-terminal domains (10-50 residues). This diversity controls the ability of GPCRs to interact with very different ligands. Furthermore, this domain often contains N-linked glycosylation at N residues, that influence cellular trafficking. In some cases, these PTMs are involved in the downregulation of receptor activity. The extracellular loops (ECLs) connecting the TM domains on the extracellular surface are variable in sequence and size too. These loops also contain conserved N residues that could be glycosylated to control the localization of the receptor. In addition, these loops contain conserved C residues which may form disulfide bridges, that further stabilize the structure of the receptor. The common structural signature observed in GPCRs, is the TM helical bundle formed by seven amphipathic  $\alpha$ -helices, which are evolutionary conserved among all classes. This domain is rich in hydrophobic residues and forms a tightly packed, tube-shaped central core. The hydrophobic residues are arranged to face the fatty acid tails of the surrounding lipid bilayer, whereas the hydrophilic amino acids of the TM helices face towards the core of the receptor. Helix-helix interactions also contribute to the stability of the tertiary structure



**Figure 4.** Domain architecture of GPCRs (Schneider et al. 2018).

of GPCRs. The orthosteric binding pocket is often located in this domain and the external signal is transmitted to the intracellular G protein complex through the conformational changes of the transmembrane helices. There is a short alpha helical segment, positioned perpendicular to the TM helices, connecting the last TM helix with the C-terminal domain. This segment is called helix 8 (H8) or cytosolic helix, as it is accessible from the cytosolic side. H8 has an important role in the binding of G proteins and arrestins and subsequent signal transmission.

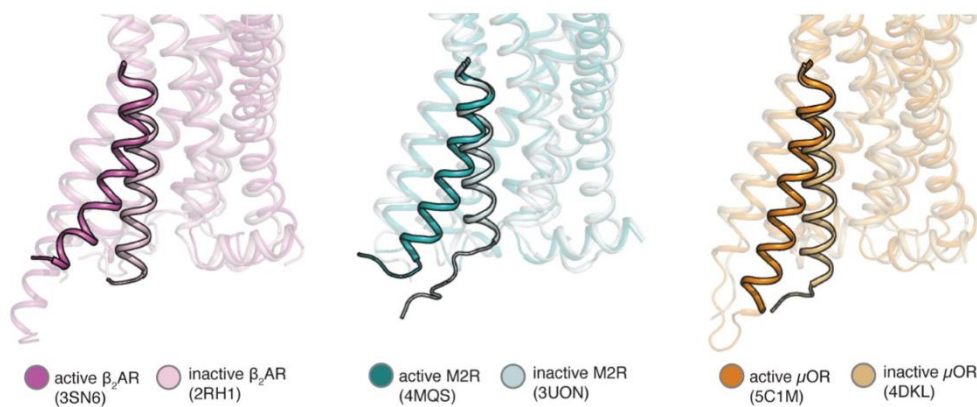
The intracellular loops (ICLs) connect the TM helices on the intracellular side. These loops are also involved in the binding of G proteins and arrestins, as well as kinases. The C-terminal domains are variable in sequence, but less variable in size. Several conserved S and T residues are contained in these domains which get phosphorylated by GRKs to promote arrestin binding and subsequent desensitization and internalization of the receptors. A conserved, lipidated C residue in this domain furnish receptor localization and provides further structural stability to these domains through the insertion of the attached lipid tail into the bilayer.

### **1.3.2. Experimental Investigations of GPCR Structure**

Advanced experimental techniques such as x-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and most recently cryogenic electron microscopy (cryo-EM) have provided invaluable structural information about GPCRs in their active and inactive states. Innovative developments, such as the application of fusion proteins and specific antibody fragments as crystallization chaperones as well as the introduction of the cryo-EM technique resulted in an exponential growth in the number of GPCRs with known atomic resolution structures (Cherezov et al. 2007, Rasmussen et al. 2011). A major milestone of structural studies of GPCRs was recognized by awarding the Nobel Prize in Chemistry to Brian K. Kobilka and Robert J. Lefkowitz “for studies of G-protein–coupled receptors” in 2012. In the past 20 years, 113 high resolution structures of GPCRs have been published, in complex with various small molecules, peptides, G-proteins, and arrestins. Currently available high-resolution structures of active, and inactive state GPCRs provide detailed information about the large structural rearrangements associated with receptor activation. Also, intermediate structures of a few GPCRs (pdb codes: 6DRZ, 5TZY, 6Z10, 5NX2, 7CA5) have been reported recently. Structurally diverse ligands have been shown to bind to GPCRs, but considerable evidence have been gathered in support of that they induce similar transmembrane movements during activation, in accordance

with their functional properties. A large displacement of TM6 is the most specific and pronounced structural change that was shown to take place during GPCR activation.

Rhodopsin was the first GPCR of which atomic resolution structural data was reported, which provided a detailed picture of the architecture of the transmembrane domain, but did not sufficiently describe conformational changes that accompany receptor activation. A hydrogen bond connecting the adjacent E134<sup>3.49</sup> and R135<sup>3.50</sup> residues of the highly conserved E/DRY motif in TM3 with E247<sup>6.30</sup> of TM6 (Ballesteros-Weinstein numbering is indicated as upper indices), forming a functional “ionic lock”, stabilize the inactive conformation (pdb code: 1F88; Palczewski et al. 2000). This was later supported by results obtained for the  $\beta_2$  adrenergic receptor ( $\beta_2$ AR) bound to a partial agonist, where this ionic lock was disrupted upon receptor Activation (Yao et al. 2006). In another study, agonist-induced conformational changes of  $\beta_2$ AR, specifically the rotation and/or tilting of TM6 was reported (Ghanouni et al. 2001). The rearrangement of TM6 was also suggested earlier for rhodopsin- $G_T$  photoactivation (Farrens et al. 1996). High resolution structures of the inactive and active  $\beta_2$ AR published in 2007 and 2011, respectively, revealed an approximately 14 Å outward movement of TM6, a smaller inward movement of TM5 and TM7, the formation of a two-turn  $\alpha$ -helix in ICL2 and the displacement of the  $\alpha$ -helical domain of the  $G_s\alpha$  protein upon receptor activation (pdb codes: 2RH1, 2R4R and 3SN6; Cherezov et al. 2007, Rasmussen et al. 2007, Rasmussen et al. 2011). A so-called conformational toggle switch (W246<sup>6.48</sup>) was proposed to play an important role in the transition between active and inactive states of the  $A_{2A}$ -adenosine receptor, and potentially all class A GPCRs. (pdb code: 3EML; Jaakola



**Figure 5.** Comparison of the structures of active and inactive state GPCRs (Latorraca et al. 2017).

et al. 2008). Another important finding indicated the position of the orthosteric and allosteric binding sites of the the M2 muscarinic acetylcholine receptor (M2R) and that they form a long aqueous channel extending approximately two-thirds of the vertical length of the TM domain (pdb codes: 4MQT, 3UON; Kruse et al. 2013, Haga et al. 2012).

The crystallographic structure of the inactive MOP indicated the formation of a salt bridge between D164<sup>3.49</sup> and R165<sup>3.50</sup> of the conserved E/DRY motif. Similar to  $\beta_2$ AR, no ionic lock interaction was observed between D164<sup>3.49</sup> and the cytoplasmic tip of TM6. Instead, polar interactions between D164<sup>3.49</sup> and R179 of ICL2 and R165<sup>3.50</sup> and T279<sup>6.34</sup> were indicated. This latter interaction was suggested to furnish stability to the inactive structure, similar to the ionic lock discussed above. (pdb code: 4DKL; Manglik et al. 2012). A detailed study on the inactive  $\delta$ -opioid receptor (DOP) revealed that a cluster of a Na<sup>+</sup> ion coordinated by D95<sup>2.50</sup>, N131<sup>3.35</sup>, S135<sup>3.39</sup>, W274<sup>6.48</sup> and water molecules could stabilize the inactive structural state (pdb code: 4N6H; Fenalti et al. 2014). NMR spectroscopic analysis of MOP bound by the small molecule agonist BU72 exhibited the role of ICL1 and H8 during receptor activation. It was indicated that the initial interactions between the receptor and the G<sub>i</sub> protein complex are furnished by ICL1 and H8 (Sounier et al. 2015). The simultaneously published x-ray crystallographic structure of the active MOP-BU72 complex indicated an approximately 10 Å displacement of TM6, accompanied by the moderate inward shifts of TM5 and TM7. No salt bridge between the adjacent D164<sup>3.49</sup> and R165<sup>3.50</sup> residues of the E/DRY motif or allosteric Na<sup>+</sup> binding was observed in this active state of the receptor. Compared to the previously published inactive state structure of MOP, residues F289<sup>6.44</sup>, P244<sup>5.50</sup> and I155<sup>3.40</sup> (the "core triad" or PIF motif) reoriented concurrently with the movements of TM5 and TM6. In the active state the orthosteric binding pocket is connected to the G protein-coupling interface through a water-mediated polar network. This network involves the polar amino acid side chains of the ortho- and allosteric binding pockets and the conserved CWxP and NPxxY motifs. An interaction between H54 of the N-terminal domain of the receptor and the secondary amine of the bound BU72 implied that the N-terminal can act as lid over the receptor core (pdb code: 5C1M; Huang et al. 2015). Similar trend was observed for the human type 1 cannabinoid receptor (CB1), although in the inverse agonist-bound inactive state in which the ECL2 and N-terminal domain formed a cap on the orthosteric binding pocket blocking solvent entry (pdb code: 5U09; Shao et al. 2016). The active structural state of the 5-hydroxytryptamine receptor-2B (5-HT<sub>2B</sub>) demonstrated a 6.7 Å and 2.9 Å displacement of TM6 and TM5,

respectively, measured at the intracellular tips of these TM helices with respect to the corresponding inactive structure and the breakdown of the intramolecular salt bridge involving the E/DRY motif (pdb code: 5TUD; Ishchenko et al. 2017). This was supported by the active structure of the  $\kappa$ -opioid receptor (KOP). The role of the W293<sup>6,48</sup> residue of the conserved CWxP motif as a central conformational switch of class A GPCR activation was further supported by results obtained for the active KOP. According to the proposed mechanism, the W293<sup>6,48</sup> side chain adjacent to the orthosteric binding site changes conformation and consequently the kink of TM6 introduced by P295<sup>6,50</sup> is reoriented upon agonist binding leading to the notable outward shift of the cytoplasmic end of TM6 (pdb code: 6B73; Che et al. 2018). Very recently, the cryo-EM structure of the MOP coupled with the G<sub>i</sub> protein complex was reported. In this native-like complex the interaction between the MOP and the G<sub>i</sub> protein involved ICL2, ICL3 and TM3, TM5, and TM6 of the receptor and the  $\alpha$ N helix, the  $\alpha$ N- $\beta$ 1 loop and the  $\alpha$ 5 helix of the G<sub>i</sub> $\alpha$  subunit (H5G $\alpha$ ). The conformation of ICL3 was different from that observed for the Nb39 nanobody-stabilized active state structure of the MOP, although the overall structures were highly similar (pdb code: 6DDE; Koehl et al. 2018). A recently published NMR spectroscopic study of the  $\beta_2$ AR indicated that the  $\alpha$ -helical conformation of ICL2 is preserved when the receptor is bound by the G<sub>s</sub> protein complex and partially unfolded upon interaction with the G<sub>i</sub> protein complex. This suggests that the structure of ICL2 determines G protein specificity (Ma et al. 2020). Interestingly, the ICL2 of the  $\beta_1$  adrenergic receptor ( $\beta_1$ AR) was found to be helical in both the G<sub>i</sub> and G<sub>s</sub> protein-bound state, although the specific interactions between this intracellular loop and the H5G $\alpha$  of G<sub>i</sub> and G<sub>s</sub> were entirely different (Alegre et al. 2021).

### 1.3.3. Limitations of Experiments

Atomic resolution experimental methods have been successfully applied to solve the three-dimensional structure of the transmembrane domains of GPCRs. However, in the majority of cases these structures represent a single low energy conformation of either the active or the inactive state of the receptor and lack information about the dynamic process of activation. Furthermore, the N- and C-terminal domains as well as the longer ILC3s, along with the corresponding PTMs are generally omitted from the x-ray crystallographic structures owing to the difficulty of crystallization of GPCRs when these flexible domains are attached. Point mutations that improve the thermal stability of GPCRs and hence aid their crystallization are also often applied, although



such mutations often lock the protein in the inactive state. A further tool of structural stabilization to facilitate protein crystallization is the application of so-called crystallization chaperones, such as specific antibody fragments (nanobodies), or fusion proteins. Such additives which may severely affect the relative mobility of the transmembrane helices and the overall dynamics of the transmembrane domains of GPCRs. Crystal lattice forces in the crystalline state could also impose unnatural restraints and distort the tertiary structure of proteins.

Recent reports indicated that negatively charged lipids are facilitating the interaction of GPCRs and G proteins, for example in the  $\beta_1$ AR -  $G_s$ , type 1 neurotensin receptor (NTS1) -  $G_q$ , type 2 cannabinoid receptor (CB2) -  $G_i$  and rhodopsin –  $G_T$  protein complexes (Yen et al. 2018, Inagaki et al. 2012, Vukoti et al. 2012, Kaya et al. 2011). While the local membrane composition was found to modulate signaling pathways of  $\beta_2$ AR bound to  $G_s$ , negatively charged phospholipids were specifically suggested to enhance the affinity of agonists and stabilize the active state of the  $\beta_2$ AR (Verkleij et al. 2000, Strohman et al. 2019). Unfortunately, the effects of the local membrane compositions and specific lipid-protein interactions are difficult to investigate using the currently available atomic resolution experimental techniques.

### 1.3.4 Computational Studies of GPCRs

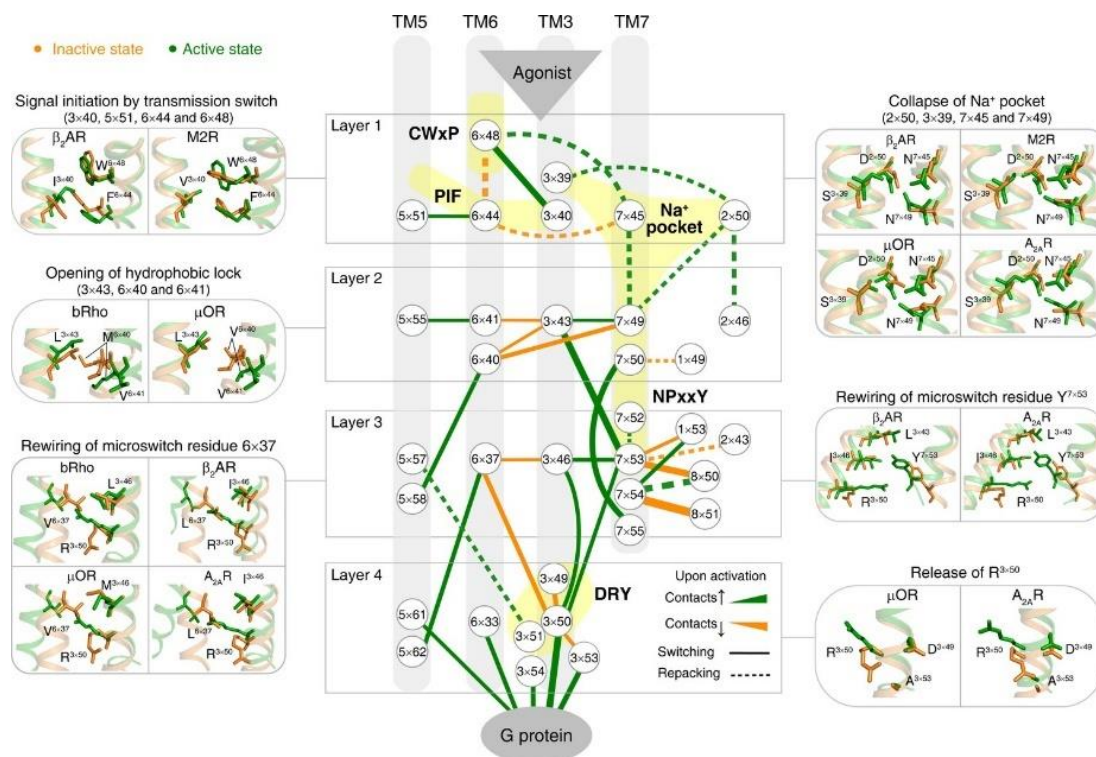
Comparative modeling studies and molecular dynamics (MD) simulations of the DOP, KOP and MOP receptors revealed that the receptor microenvironments were primary determinants of the subtype specificities in opiate ligand binding and the concerted movements of distant parts of ligand-binding regions suggested motion-sensitive components of ligand binding (Strahs et al. 1997). Homology modeling of the  $\alpha_{1B}$  adrenergic receptor ( $\alpha_{1B}$ AR) suggested that the equilibrium between active and inactive states depends on the protonation state of D142<sup>3,49</sup> of the E/DRY motif. It also revealed the fundamental role of R143<sup>3,50</sup> in receptor activation. Mutation of R143<sup>3,50</sup> of the conserved 'polar pocket' formed by N63<sup>1,50</sup>, D91<sup>2,50</sup>, R143<sup>3,50</sup>, N344<sup>7,49</sup> and Y348<sup>7,53</sup> have elevated the constitutive activity of the receptor (Scheer et al. 1996). The P side chain usually precludes normal helix geometry, due to its lack of an amide proton, thus causing a break/bend in the  $\alpha$ -helix. The role of P kinks, present in TM helices of GPCRs, has been investigated and it was suggested that these kinks may act as hinges which facilitate the reorientation of TM helices (Sansom et al. 2000). An earlier MD study of the homology model of the 5-hydroxytryptamine receptor-1A (5-HT<sub>1A</sub>) receptor revealed that the chemical information may be transferred from the

extracellular to the cytosolic domains through a cluster of aromatic amino acids along TM6 and the reorientation of the helix kink at P360<sup>6,50</sup> was implied in the breakage of the interhelical salt bridge between R134<sup>3,50</sup> of E/DRY motif and E340<sup>6,30</sup> (Seeber et al. 2003).

The first long time-scale all-atom simulation along with experimental methods was applied for  $\beta_2$ AR bound to an inverse agonist. It was concluded that the active state of the receptor spontaneously destabilizes and transitions into the inactive state in the absence of either G proteins or a stabilizing nanobody (Rosenbaum et al. 2011). Similar approach was applied in another study of  $\beta_2$ AR aimed at the mechanism of transition from the active to inactive state. In that study 92 simulations of total 656  $\mu$ s was performed for the  $\beta_2$ AR bound either by the  $G_s$  protein complex or by the Nb80 nanobody. An articulate population of stable intermediates was identified during the deactivation simulations with regard to the conformation and/or position of specific residues of the orthosteric binding pocket and the PIF and NPxxY motifs. Consequently, the structure of the receptor was described as a dynamic ensemble of multiple active, inactive and intermediate states. The populations of these states are shifted upon the binding of intracellular signaling molecules and ligands of different functional properties. The binding of a G protein to the intermediate state was proposed as a catalyst of the activation mechanism and that the presence of an agonist in the orthosteric binding site shifts the conformational equilibrium towards the active state (Dror et al. 2009 and 2011, Latorraca et al. 2017). Although the information provided by the above large-scale simulations is tremendous, the exceptional computing power such simulations demand was only accessible on dedicated supercomputers.

Recently, as an alternative, cloud computing was utilized to perform large scale simulations of the  $\beta_2$ AR. Markov state models were used to build a statistical model from independent simulations of 2.15 ms in total which confirmed the presence metastable intermediate states and explained how ligands of different functional properties alter the conformational landscape of the receptor (Kohlhoff et al. 2014). The use of biased MD for GPCRs was first reported for bovine rhodopsin in different activation states. The free-energy landscape constructed along the pre-determined pathways of structural transition revealed the existence two different pathways and four metastable intermediate states, characterized by the disposition of TM6. Rhodopsin is a unique GPCR in a sense that it is covalently bound to 11-*cis*-retinal in the resting state and receptor activation is initiated by the absorption of a photon and subsequent isomerization of 11-*cis*-retinal

to 11-*trans*-retinal. Covalently bound 11-*cis*-retinal was indicated to destabilize active-like intermediate states through the disruption of the ionic lock between TM3 and TM6 (between R135<sup>3,50</sup> and E247<sup>6,30</sup>) and the modulation of the rotamer toggle switch (W265<sup>6,48</sup>) (Provasi et al. 2010). The potential role of water molecules in the activation mechanism of GPCRs was analyzed by accelerated MD (aMD) simulations of the A<sub>2A</sub>-adenosine receptor, which confirmed the presence of water molecules in the receptor core, while demonstrated higher conformational sampling efficiency, compared to conventional MD simulations (Zia et al. 2016). In another attempt to reduce the computing time, supervised MD (suMD) simulations were carried out for the A<sub>3</sub>-adenosine receptor in order to study the role of ECL2 in the ligand binding event. Residues K152, T154, V169 and R173 of ECL2 were shown to establish contacts with the agonist which trigger conformational changes in the ECL2. This the initiated the formation of further ligand interactions involving residues of the orthosteric binding site. ECL2 was proposed to act as a cap of the orthosteric binding site closing the entrance to the site upon agonist binding (Deganutti et al. 2015).



**Figure 6.** Cascade mechanism of GPCR activation involving conserved conformational switches (Zhou et al. 2019).

The translocation of  $\text{Na}^+$  in the transmembrane core and its effect on the class A GPCRs have been extensively studied by both experimental and theoretical methods (Mahaut-Smith et al. 2008, Zhang et al. 2012, Fenalti et al. 2014, Yuan et al. 2013, Shang et al. 2014, Hu et al. 2019, Selent et al. 2010, Fleetwood et al. 2020). The allosteric effect of  $\text{Na}^+$  ions was explained on the basis of advanced MD simulations of the MOP receptor.  $\text{Na}^+$  was found to associate quickly with its allosteric binding site in ligand free states of the receptor. Apart from D114<sup>2.50</sup> of the allosteric site,  $\text{Na}^+$  frequently interacts with the nearby W293<sup>6.48</sup>, N328<sup>7.45</sup> and N332<sup>7.49</sup>, which form a secondary  $\text{Na}^+$  coordination site. The activation of receptor was shown to be accompanied by the collapse of the allosteric  $\text{Na}^+$  binding site, resulting in the ejection of the bound  $\text{Na}^+$  and its migration towards the cytosol. Energy landscape analysis indicated that an energy barrier is to be crossed for  $\text{Na}^+$  translocation (Yuan et al. 2013, Shang et al. 2014, Hu et al. 2019). A comparative study of several class A GPCRs on the water network in the receptor core indicated, that the water molecules form two types of polar networks: a network that is maintained across the active and the inactive states and a network that rearranges during activation (Venkatakrishnan et al. 2019).

Most recently, conformational changes in 234 structures of 45 different class A GPCRs were analyzed in fine details. Residue contact maps, allowed researchers to string together the well-known, conserved, but spatially scattered key functional motifs that link the orthosteric ligand binding pocket to the intracellular G protein-coupling interface (Figure 6). Four sequential stages of receptor activation were proposed which is initiated in “layer 1” where the ligand binds to the orthosteric binding pocket resulting in change in the residue contacts at the bottom of the binding site, leading to collapse of the allosteric  $\text{Na}^+$  binding site. This is then relayed to “layer 2” opening a hydrophobic lock, resulting in the rewiring of residues of NPxxY and E/DRY microswitches in layer 3 and layer 4, respectively, facilitating the outward movement of TM6, making the receptor competent for G protein coupling (Zhou et al. 2019).

### 1.3.5. Limitations of Computational Techniques and Studies

The majority of information about GPCR dynamics was provided by MD simulations. The previous section demonstrates how the theory of GPCR activation evolved from classic static two-state models to the current state-of-the-art, proposing that GPCR activity is linked to structurally distinct, dynamically interconverting functional states and the process of activation relies on an interconnecting network of conformational switches in the transmembrane domain. Conventional

MD simulations, however, have a certain limitation that they need advanced computing hardware and/or tremendous amount of computing time to acquire a statistically sound picture of large-scale conformational transitions, occurring on long time scales. The limited capability for the assessment of pairwise interactions that involve subtle, *in situ* changes in the electronic structure and polarization is a general limitation of classical force fields used in MD simulations. The application of coarse-grained models further limits the assessment of pairwise interactions. Enhanced sampling techniques require less computing power but only at the sacrifice of natural or realistic conditions.

Apart from the above limitations, there are some application-based approximations, specific to GPCR simulations. These simulations share a common theme by confirming the structural ensemble of active, inactive and intermediate states and providing insight into the molecular details and kinetics of transitions between them. However, there are several crude approximations commonly employed and reported in the vast literature of GPCR simulation studies. The ICL3, often missing from experimental structures of GPCRs, was not explicitly taken into account in several GPCR simulations. The highly variable N- and C-terminal domains are generally missing from the experimental structures of GPCRs, because these domains cannot be crystallized and resolved owing to their inherently flexible structures. The computational prediction of the structure of these domains is a lot more problematic, than that of the ICL3 (Dror et al. 2011, Manglik et al. 2012, Fenalti et al. 2014, Huang et al. 2015, Sounier et al. 2015, Koehl et al. 2018, and Marino et al. 2018). Consequently, PTMs of the N- and C-terminal domains, which could have a crucial role in the activation mechanism were omitted too. The application of single-component membrane bilayers or very simplistic membrane models are dominating the practice of MD simulation studies of GPCRs, whereas there are numerous reports advocating the role of the lipid environment in the function of transmembrane proteins, including a recent one highlighting the effect of charged interfacial lipids on the activation of the  $\beta_2$ AR (Strohman et al. 2019).

## 2. Aims and Scope

The explanation of the activation mechanism of GPCRs according to the current state-of-the-art, presented above in detail, gets complicated when multiple active states or structurally similar but functionally different ligands are considered. Models built exclusively on a structural basis have limited capabilities to quantitatively address differences between ligands with similar constitution, three-dimensional structure, physico-chemical properties and binding affinity, but different efficacy. In order to respond to the most recent challenge of rational drug design and to develop high-affinity, high-efficacy, and functionally selective GPCR ligands, a quantitative model of the activation mechanism is needed. In light of the above discussed shortcomings of the existing models, this necessitates the introduction of new perspectives. In this thesis a new model is proposed. According to this model the activation of GPCRs is accompanied by a shift of macroscopic polarization in a shielded central duct of the TM domain, initiated by ligand binding and propagated to the intracellular G protein-binding interface. This hypothesis was formulated on the basis of the abundance of highly conserved polar/ionizable species which were indicated in previous reports to be key participants of the activation mechanism.

The main objective of the study presented in this thesis is to test the feasibility of the above hypothesis and to reveal further key determinants of the mechanism of activation of GPCRs, by utilizing large scale MD simulations in which the physiological state and environment of the receptor is represented as accurately as possible. Evidently, MD simulations employing fixed point charge force fields cannot reveal exact or quantitative details of processes involving charge shift. Nevertheless, statistical analysis of the dynamic motions of key molecular parts could provide convincing support for the interplay of the above-mentioned polar species.

With regard to the above considerations the specific aims of this study were set as follows.

1. To build stable and realistic MD simulation systems, which represent the native state and environment of the MOP,  $\beta_2$ AR and CB1 receptors, representative class A GPCRs addressed in this study.

- 1.1 To model and include the flexible N- and C-terminal domains and ICL3 to study the dynamics of the full sequence receptors.

- 1.2 To obtain plausible ligand-receptor complex structures via molecular docking if experimental data of such complexes are unavailable.
  - 1.3 To apply native conditions for protein function through the explicit inclusion of all PTMs such as phosphorylation, glycosylation, and lipidation.
  - 1.4 To approximate native membrane environment through the application of explicit caveolar (raft-like) membrane composition during simulations.
2. To utilize large scale MD simulations in order to investigate specific structural transitions of GPCRs in the presence/absence of ligands and intracellular signaling proteins that correlate with the activation state of the receptor.
  3. To design and employ specific analysis methods or strategies in order to identify further potential parameters of GPCR signaling which could rationally extend the current state-of-the-art.

## 3. Materials and Methods

### 3.1. System Building

#### 3.1.1. Obtaining Atomic Resolution Starting Structures

All X-ray crystallographic structures used in this study were downloaded from the Brookhaven Protein Data Bank (<http://www.rcsb.org>). The active and inactive state crystallographic structures of the MOP (pdb code: 5C1M and 4DKL, respectively) bound to  $G_i$  and GTP (pdb code: 1GP2) and  $\beta$ -arrestin-2 (pdb code: 3P2D) and the endogenous ligand endomorphin 2 (EM2) were used as the starting structures of MD simulations of this receptor. For  $\beta_2$ AR the active and inactive state structures of the receptor (pdb code: 3SN6 and 2RH1, respectively) were retrieved along with the structure of the full  $G_s$  protein and GDP complex (pdb code: 1AZT) and  $\beta$ -arrestin-2 (pdb code: 3P2D). Starting structures of this receptor also included the endogenous agonist epinephrine (pdb code: 4LDO). The  $\alpha$  subunit of the  $G_s$  protein had missing parts in the crystallographic structure of the active  $\beta_2$ AR (pdb code: 3SN6), therefore this subunit was supplemented from an independent crystallographic structure of the  $G_s$  protein (pdb code: 1AZT). The active and inactive state structures of the CB1 receptor (pdb codes: 6N4B

and 5TGZ, respectively) complexed with the  $G_i$  protein and GDP (pdb code: 1GP2) or  $\beta$ -arrestin-2 (pdb code: 3P2D) and 2-arachidonoylglycerol (2-AG) were used to build the third model GPCR system. A large portion of the  $G\alpha$  subunit is missing from the active state CB1 crystallographic structure (pdb code: 6N4B), hence coordinates of this subunit from a separate crystallographic structure of the  $G_i$  heterotrimeric complex were included (pdb code: 1GP2). The crystallographic structures of the  $G_s$  protein-bound active  $\beta_2$ AR (pdb code: 3SN6) and the visual arrestin-bound rhodopsin (pdb code: 4ZWJ) served as templates for the alignment of G protein and arrestin-bound GPCR complexes, respectively.

### 3.1.2. Modeling and Attachment of Missing Domains and Residues

The full sequence of the murine MOP (UniProtKB-P42866-OPRM1), human  $\beta_2$ AR (UniProtKB-P07550-ADRB2) and human CB1 (UniProtKB-P21554-CNR1) was obtained from UniProt (<http://www.uniprot.org>). The ICL3s missing from the crystallographic structures were modeled using the Modeller 9.20 software and other missing or mutated residues were added or reverted using the Swiss-PdbViewer program (Webb et al. 2016, Guex et al. 1997). The missing N- and C-terminal domains were modeled by performing 10 ns folding simulations using the GROMACS 5.1.4 or 2018.3 program packages, the AMBER ff99SB-ILDN-NMR force field and the GB/SA implicit solvation model (Abraham et al. 2015). During MD simulations, the system temperature was set to 310 K and maintained by the v-rescale algorithm. Ten parallel simulations were run for both the N- and C-terminal domains from where the resultant, folded structures were evaluated and selected based on their compactness and the accessibility of PTM and TM region attachment sites. The selected N- and C-terminal domain structures were linked to the TM region using Pymol 2.1.0.

### 3.1.3. Post-Translational Modifications

The CHARMM-GUI web server was used to introduce PTMs in the receptor structures (Jo et al. 2008). Complex type glycans were added to the glycosylation prone sites of the N-terminal domain, containing a common core (Man $\alpha$ 1-3(Man $\alpha$ 1-6) Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc $\beta$ 1-N) and sialic acid (N-acetylneuraminic acid). Glycosylation sites were predicted using the NetNGlyc 1.0 online server (Gupta et al. 2004). Phosphorylation of S and T residues in the C-terminal domain



with regard to the G protein and  $\beta$ -arrestin bound states and palmitoylation of a C residue in the ICL3 were also added. PTMs applied for the different receptors are summarized in Table 1.

### 3.1.4. Ligand Docking

The coordinates of  $\beta_2$ AR-bound epinephrine was retrieved from the x-ray crystallographic structure of this complex (pdb code: 4LDO). Epinephrine was inserted in the binding pocket of the receptor in a protonated form. Due to the lack of high-resolution structures of EM2 bound to the MOP and 2-AG bound to the CB1, the binding sites of these ligands were predicted utilizing molecular docking. The structure of EM2, a peptide agonist of the MOP (Zadina et al. 1997), was built manually using Pymol 2.1.0. The orthosteric binding site was confirmed by flexible docking of this ligand to the active state MOP crystallographic structure (5C1M) using the Autodock 4.2 software and the Lamarckian genetic algorithm. All  $\phi$ ,  $\psi$ , and  $\chi^1$  ligand torsions, as well as receptor side chains in contact with the bound ligand were kept flexible (Huang et al. 2015). Docking of EM2 was performed in an 8.0 nm  $\times$  8.0 nm  $\times$  8.0 nm grid volume, large enough to cover the whole surface of the receptor accessible from the extracellular side, including the orthosteric binding pocket (D147<sup>3.32</sup>, Y148<sup>3.33</sup>, M151<sup>3.36</sup>, K233<sup>5.39</sup>, W293<sup>6.48</sup>, I296<sup>6.51</sup>, H297<sup>6.52</sup>, V300<sup>6.55</sup>, I322<sup>7.39</sup>, Y326<sup>7.43</sup>). The spacing of grid points was set at 0.0375 nm and 1000 dockings were done. The resultant ligand-receptor complexes were clustered and ranked according to the corresponding binding free energies.

**Table 1.** List of PTMs sites in MOP,  $\beta_2$ AR and CB1 receptors.

GPCR	Glycosylation	Phosphorylation	Lipidation
MOP	N9, N31 and N38 (Huang et al. 2008)	G <sub>i</sub> protein: S363, T370 $\beta$ -arrestin-2: S363, T370, S375, T376 and T379 (Mann et al. 2015)	C170 (Zheng et al. 2012)
$\beta_2$ AR	N6, N15 and N187 (Mialet-Perez et al. 2004)	G <sub>s</sub> protein: no phosphorylation $\beta$ -arrestin-2: S355, S356 and S364 (Zamah et al. 2002, Hausdorff et al. 1989)	C341 (O'Dowd et al. 1989)
CB1	N77 and N83 (Howlett et al. 1991, Song et al. 1995, Ruehle et al. 2017, Rapino et al. 2019)	G <sub>i</sub> protein: no phosphorylation $\beta$ -arrestin-2: T460, S462, S464, T465, T467 and S468 (Jin et al. 1999, Howlett et al. 2010)	C415 (Oddi et al. 2012)

The lowest energy bound state was selected for simulations, in which specific ligand-receptor interactions observed in the crystallographic structures were present. Cryo-EM structure of the MOP and the peptide agonist DAMGO, published later, have confirmed the correct localization and similar orientation of pharmacophores of EM2 (Koehl et al. 2018).

The above docking protocol was applied to find a plausible orientation of 2-AG in the CB1 receptor. The orthosteric binding site (F177<sup>2.64</sup>, F189<sup>3.25</sup>, L193<sup>3.29</sup>, V196<sup>3.33</sup>, F200<sup>3.36</sup>, F268<sup>ECL2</sup>, L359<sup>6.51</sup>, M363<sup>6.55</sup>, F379<sup>7.35</sup>, S383<sup>7.39</sup>, and L387<sup>7.43</sup>) was confirmed by flexible dockings of 2-AG to the crystallographic structure of the active state CB1 receptor (pdb code: 5XRA, Hua et al. 2017). The affinity of 2-AG to the CB1 was predicted to be 1.99  $\mu$ M ( $K_i$ ), which is in line with experimental data, appointing 2-AG to be a low/moderate affinity agonist of the CB1 receptor (Sugiura et al. 1995). The lowest energy bound state was selected for simulations, in which the orientation of 2-AG as well as specific ligand-receptor interactions were analogous to those observed between AM11542 and CB1 in the experimental structure (pdb code: 5XRA).

### 3.1.5. Membrane Building and the Insertion of Ligand-Receptor Complexes

Previous reports proposed the localization of several GPCRs, including the  $\beta_2$ AR, MOP and CB1 receptors, in caveolar lipid-rafts (Oh et al. 2001, Xiang 2002, Chini et al. 2004, Head et al. 2005, Calizo et al. 2012). Membrane bilayers with caveolar compositions were built in line with lipidomic data using the membrane builder tool of CHARMM-GUI and CHARMM36 parameters were included for the lipid components (Pike et al. 2002). The resultant multicomponent, raft-like membrane systems included cholesterol (CHL-32.8%), 1-palmitoyl-2-oleoyl-glycero-3-phosphatidylcholine (POPC-14.9%), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylethanolamine (POPE-27.8%), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidyl-L-serine (POPS-3.6%), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylinositol (POPI2-6%), palmitoyl-sphingomyelin (PSM-9.9%) and monosialodihexosylganglioside (GM3-5%). The asymmetric upper and the lower leaflet membrane compositions were specified in a most probable ratio (Ingólfsson et al. 2014) and the coordinates of the membrane orientation were obtained from the OPM server (<http://opm.phar.umich.edu>). The CHARMM-GUI membrane builder was used to embed the glycosylated, palmitoylated and phosphorylated full sequence receptors into the membrane slabs. Systems were then solvated explicitly with TIP3P water molecules in a hexagonal shaped periodic

box, and sodium and chloride ions (0.15 M) were added to neutralize the net charge and to attain physiological ionic strength. System coordinates and topologies were generated in GROMACS format.

### 3.2. Energy Minimization and Equilibration

After orienting and adding ligands and membrane insertion of ligand-receptor complexes, the resultant systems were energy minimized thoroughly by performing 5000 steps steepest descent, followed by 5000 steps conjugate gradient minimization, having the convergence criteria set to 1000 kJ/mol nm<sup>-1</sup> in both cases. After energy minimizations, systems were subjected to a six-step equilibration protocol, as recommended by CHARMM, using the GROMACS 5.1.4 and GROMACS 2018.3 program packages. According to this protocol, restrained MD simulations were executed in the NVT and then, after 2 steps, in the NPT ensemble at 303.15 K and 1 bar, having the positional restraints on the heavy atoms of the proteins and membrane constituents decreasing gradually. The first three equilibration MD runs were 25 ps long and were performed in 1 fs time steps. The following two were continued for 100 ps in 2 fs time steps. The last equilibration step was extended to 50 ns and was executed in 2 fs time steps. The LINCS algorithm was used to constrain all bonds to their correct length. System temperature was regulated by the v-rescale algorithm with a coupling constant of 1 ps and semi-isotropic pressure coupling was applied with a coupling constant of 5 ps and compressibility of  $4.5 \times 10^{-5}$  bar<sup>-1</sup>. The Particle Mesh Ewald (PME) method was used to calculate energy contributions from electrostatic interactions. Van der Waals interactions were calculated using a twin-range cutoff. All cut-off values were set to 1.2 nm.

### 3.3. Production Simulations

The production simulations were performed at 310 K in the NPT ensemble, using GROMACS 5.1.4 for the MOP and GROMACS 2018.3 for the  $\beta_2$ AR and CB1 receptors with other parameters similar to the above. Each production simulation was 1  $\mu$ s long. System coordinates were stored in every 5000th steps yielding trajectories of 100.000 snapshots.

In total, ten independent simulations were performed for the inactive and active MOP, complexed with heterotrimeric G<sub>i</sub> protein,  $\beta$ -arrestin-2, Nb39 nanobody and T4-lysozyme. These latter two macromolecular complexes served as references representing the effect of crystallization

conditions on the structure of the MOP (pdb codes: 4DKL and 5C1M). Two additional control simulation systems were built for the active state,  $G_i$  protein-bound receptor, either with the exclusion of bound EM2, or with the inclusion of EM2 in the orthosteric site together with a  $Na^+$  placed in the proximity of D<sup>2.50</sup> in the allosteric binding pocket. The ten independent simulations yielded a total of 10  $\mu s$  of trajectories.

Eleven independent production simulations were performed for the active and inactive state  $\beta_2AR$ , complexed either by the heterotrimeric  $G_s$  protein or  $\beta$ -arrestin-2 and in the presence or in

**Table 2.** Simulation systems

Receptor	State	Intracellular proteins/chaperones	Ligand	Total
MOP	active	$G_i$ protein	none	10 simulations
		$G_i$ protein	EM2	
		$G_i$ protein	EM2 + allosteric $Na^+$	
		$\beta$ -arrestin-2	EM2	
		Nb39	EM2	
		T4-lysozyme	EM2	
	inactive	$G_i$ protein	EM2	
		$\beta$ -arrestin-2	EM2	
		Nb39	EM2	
		T4-lysozyme	EM2	
$\beta_2AR$	active	$G_s$ protein	none	11 simulations
		$G_s$ protein-replica1	epinephrine	
		$G_s$ protein-replica2	epinephrine	
		$G_s$ protein-replica3	epinephrine	
		$G_s$ protein	epinephrine (restrained)	
		$\beta$ -arrestin-2	none	
	inactive	$\beta$ -arrestin-2	epinephrine	
		$G_s$ protein	none	
		$G_s$ protein	epinephrine	
		$\beta$ -arrestin-2	none	
CB1	active	$G_i$ protein	2-AG	4 simulations
		$\beta$ -arrestin-2	2-AG	
	inactive	$G_i$ protein	2-AG	
		$\beta$ -arrestin-2	2-AG	

the absence of orthosterically bound epinephrine. The simulation of the active  $\beta_2$ AR bound to the  $G_s$  protein and epinephrine was performed in three replicates. An additional reference simulation was performed for this system in which mild positional restraints ( $200 \text{ kJ mol}^{-1}$  on heavy atoms) were applied for epinephrine to prevent its spontaneous ejection from the binding pocket. None of the other production simulations have included any restraints.

Four independent simulations were performed for the active and inactive state CB1 receptor bound by 2-AG and complexed with either the heterotrimeric  $G_i$  protein and  $\beta$ -arrestin-2. The simulations performed for the study presented in this thesis are summarized in Table 2.

### 3.4. MD Trajectory Analysis

Analysis of MD trajectories was performed using the analysis suite of the GROMACS 5.1.4 and 2018.3 packages. Generally, the analysis of MD trajectories was performed to evaluate membrane properties and protein conformational changes, stability of the molecular complexes, as well as to investigate previously described interactions and their role in different activation states of the receptor.

Folding simulations were assessed through the analysis of root mean square deviation (RMSD) of backbone atom positions with respect to the starting structures. Furthermore, the radius of gyration of the N- and C-terminal domains were calculated and the number of intramolecular H-bonds were counted along the trajectories using the `gmx gyrate` and `gmx hbond` utilities, respectively. The evolution of secondary structure was monitored using the DSSP method. Membrane thickness and area per lipid head group values were calculated using the FATS LiM 0.2.1 program (Buchoux et al. 2017).

For each production MD trajectory RMSD calculations were carried out to assess the structural stability of the complex and to identify significant displacements of structural components as a function of time. RMSD values of protein backbone atoms were calculated in comparison with the active and inactive state starting structures. The dynamics of terminal domains during simulations was examined by monitoring the radii of gyration. Conformational fluctuations of specific amino acid side chains were analyzed by measuring side chain  $\chi^1$  angles and calculating the frequency of transitions between rotameric states using `gmx chi`. The `gmx helix`

utility was used to calculate helix properties (Hol 1985). Secondary structure assignment was, again, done using the DSSP method (Kabsch et al. 1983).

The occurrence and frequency of intra- and intermolecular H-bonds were calculated using the gmh utility. The donor-acceptor distance and donor-hydrogen-acceptor angle cut-offs for H-bond assignments were set to 0.35 nm and 60.0 degrees, respectively. The presence of salt bridges was monitored by measuring distance and angle between the corresponding acidic and basic side chain functional groups using gmh distance and gmh gangle, respectively. The distance threshold for salt bridge assignment was 0.4 nm and the angle threshold was 90.0 degrees. Where applicable, penetration of Na<sup>+</sup> ions into the allosteric Na<sup>+</sup> binding site, D<sup>2.50</sup>, was checked using the gmh mindist utility.

The extent of correlation of atomic displacements was examined by dynamic cross-correlation matrix analysis (DCCM) integrated into an earlier version of the GROMACS suite (g\_correlation, 3.3) (Lange et al. 2006). The Gimp 2.8 software was used for image analysis of the obtained DCCM maps, where the extent of correlation was demonstrated by color intensity. The threshold of assignment of correlation was red color intensity corresponding to > 0.65 value of the correlation coefficient. Amino acid side chains having at least 4 atoms participating in correlated motions were considered. The threshold of 4 atoms have been set in order to exclude irrelevant sidechain motions, such as torsional rotations of methyl groups.

Systems were visualized using Pymol 2.1.0 and VMD 1.9.3 (Humphrey et al. 1996) software and graphs were prepared using the Xmgrace 5.1.25 program.

### **3.5. Sequence Alignment and Conservation Analysis**

244 sequences of mouse and 267 human class A GPCRs (without orphan and olfactory receptors) were retrieved from the UniProt database in FASTA format. The Clustal Omega program (Sievers et al. 2011) was used to align those multiple sequences and the results were analyzed using Jalview 2.10.5 (Waterhouse et al. 2009). The OPRM2\_MOUSE (P42866), ADRB2\_HUMAN (P07550) and CNR1\_HUMAN (P21554) sequences were set as references for the three different receptors. The sequences were compared based on percentage of identity

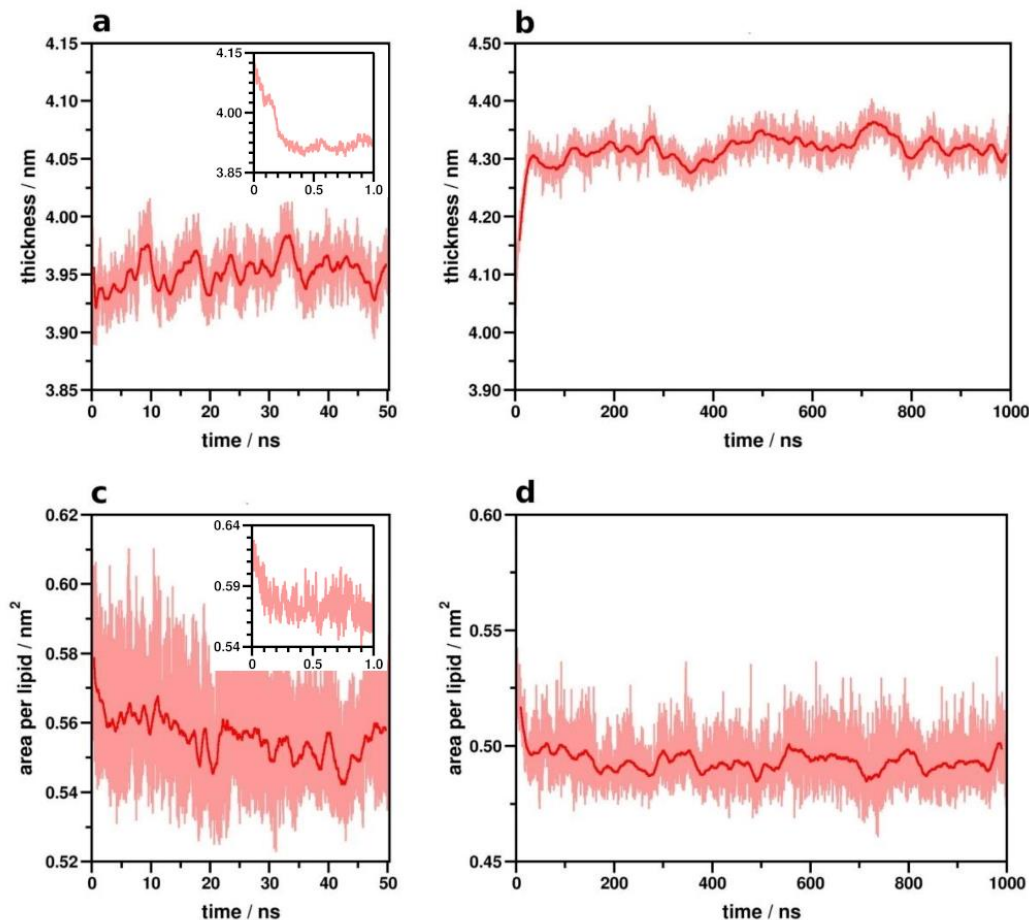
## 4. Results and Discussion

The page limit and format requirements specified by the Doctoral School of Theoretical Medicine do not allow for the complete demonstration of the large amount of data gathered during the research project presented in this thesis. Therefore, only representative graphs and tables are included along with the detailed description of observations and conclusions. Scientific journal articles published on the topic by the candidate, including all relevant data and analysis are attached to this thesis to facilitate objective assessment.

### 4.1. Membrane Properties

Raft-like lipid membrane slabs were prepared following lipidomic data and simulated with GPCR complexes (Pike et al. 2002). An insufficiently equilibrated membrane bilayer could include holes and gaps and could introduce serious artifacts. Thus, thorough minimization and sufficiently long equilibration of the caveolar membrane surrounding the protein was performed prior to production simulations. Specific properties of the membrane have been examined and compared to experimental data to establish the integrity of the membrane system applied for simulations.

Rapid vertical contraction of the membrane bilayer was observed during the equilibration phase, approaching the thickness of homogenous phospholipid bilayers while lateral contraction was gradual (Kučerka et al. 2011). After the removal of positional restraints, a slight vertical expansion took place in the first 20-30 ns of the production simulations, then the thickness of the membrane stabilized at approximately 4.30 nm, in agreement with previous reports about the thickness of lipid rafts (Figure 7) (Niemela et al. 2007). Lateral contraction of the membrane continued in the initial production phase until stabilization at approximately 0.49 nm<sup>2</sup> after 30 ns. It is important to note, that the initial imbalance of membrane parameters in the production phase emerge from that the positional restraints on the proteins were released only at the beginning of the production phase. Since simulations were intended to monitor the process of activation, of which timescale could not be estimated, the possibility of structural changes in the equilibration phase had to be kept minimal. Therefore, positional restraints on the heavy atoms of the protein were decreased step-by-step, but maintained throughout the equilibration. No artificial trends reflecting this initial membrane imbalance was perceived in the results. The average values for bilayer thickness and area per lipid



**Figure 7.** Analysis of membrane properties during equilibration and production simulation of the EM2-bound active MOP -  $G_i$  protein complex as a representative example. (a) Membrane bilayer thickness in the equilibration phase. (b) Membrane bilayer thickness in the production phase. (c) Area per lipid headgroup in the equilibration phase. (d) Area per lipid headgroup in the production phase. Inserts show the initial phase of the simulation.

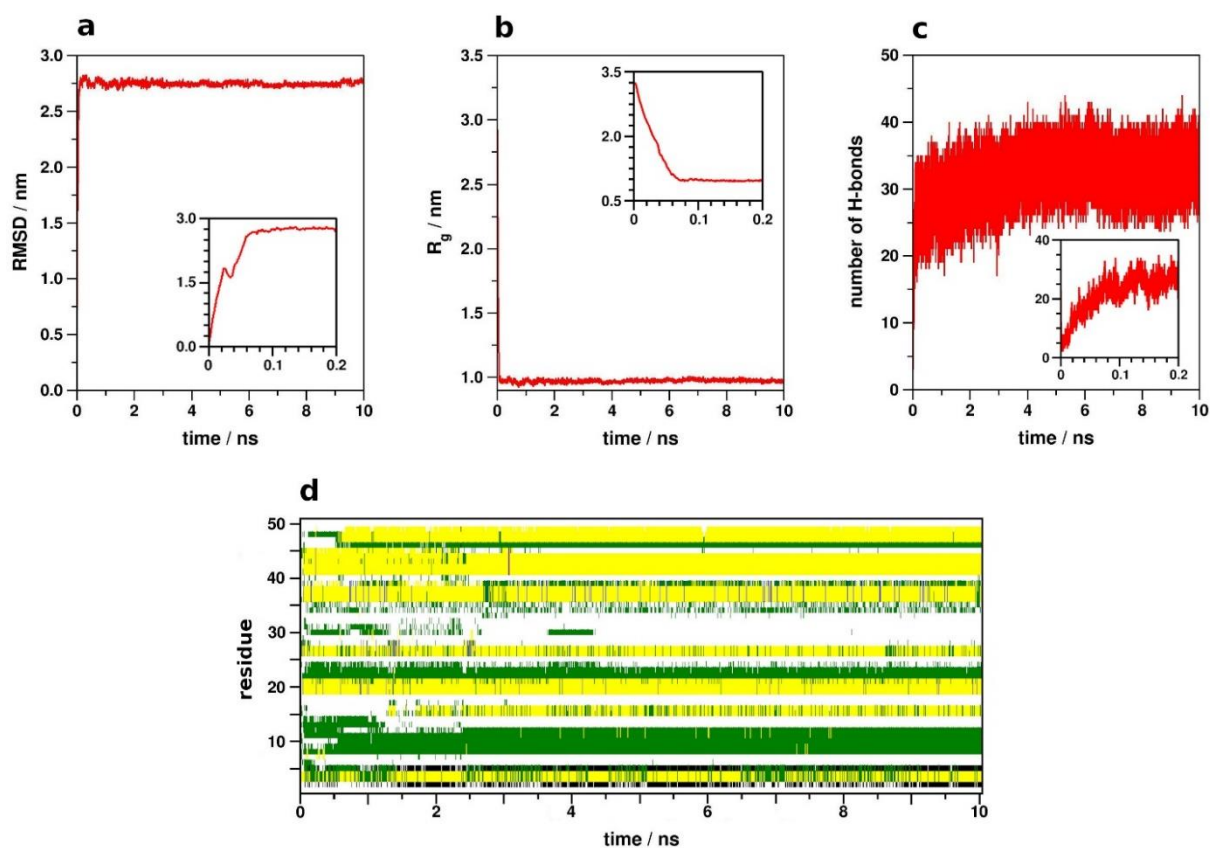
head group, were 4.328 nm and 0.493 nm<sup>2</sup>, respectively, having excluded the values of the first 40 ns.

## 4.2. Simulation System Integrity

### 4.2.1. N- and C-terminal Domain Dynamics

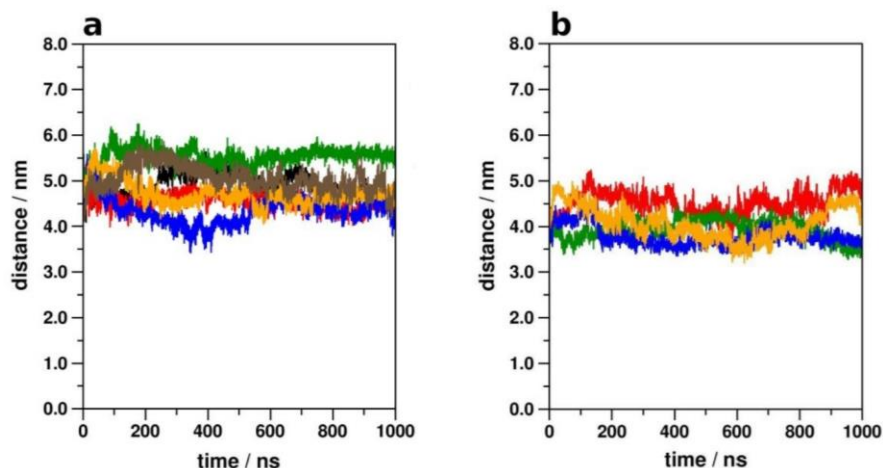
An initial challenge was to build simulation systems which approximate the physiological conditions of the GPCRs addressed in this study as closely as possible. The disposition of TM helices was proposed to have a central role in the activation mechanism therefore an important specific aim of this current study was to study the full sequence MOP,  $\beta_2$ AR and CB1 receptors in





**Figure 8.** Analysis data for a representative folding simulation of the N-terminal domain of the MOP. (a) Backbone RMSD with respect to the starting structure. (b) The radius of gyration. (c) The number of intramolecular H-bonds. (d) Evolution of secondary structure during the simulation. (White: unordered; green: bend; yellow:  $\beta$ -turn; black:  $\beta$ -bridge, grey:  $3_{10}$ -helix; blue:  $\alpha$ -helix). Inserts show the initial phase of the simulation.

order to take account of the drag posed by the mass of the relatively large N- and C-terminal domains on the TM helices and to see if that affects the internal dynamics of the TM domains (Latorraca et al. 2017, Marino et al. 2018, Zhou et al. 2019, Filipek 2019). In the absence of experimental structures of these or homologous domains, folding simulations were performed to create approximate structures. Parallel folding simulations performed for the N- and C-termini of the MOP,  $\beta_2$ AR and CB1 have provided convergent results in terms of the evolution of backbone RMSD, radius of gyration, number of H-bonds, and secondary structures (Figure 8). Although it is cannot be guaranteed that the folding of these domains was correct and complete in the available



**Figure 9.** Minimum distance between the N- and C-terminal domains in the (a) active and (b) inactive state MOP during production simulations. Black: MOP-G<sub>i</sub> protein complex without ligand, brown: MOP-G<sub>i</sub> protein complex with bound allosteric Na<sup>+</sup>, red: MOP-G<sub>i</sub> protein complex, green: MOP-β-arrestin-2 complex, blue: MOP-Nb39 nanobody complex, orange: MOP-T4-lysozyme fusion.

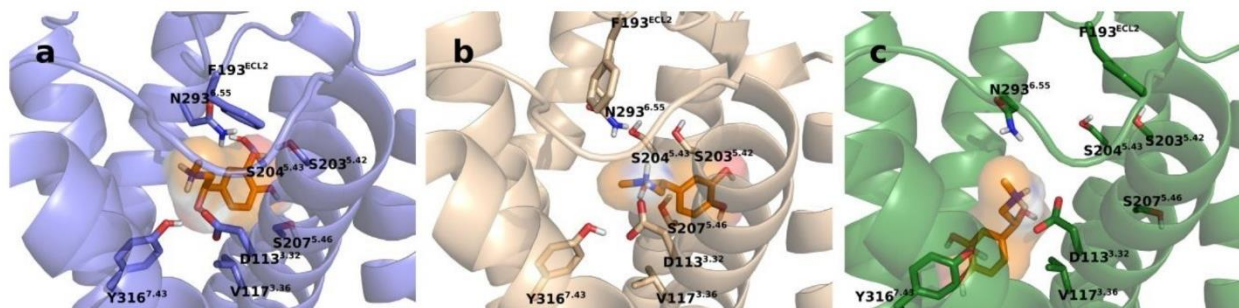
time frame, their effect on TM dynamics, primarily exerted by their mass is satisfactorily taken into account by using these resultant approximate structures.

The evolution of radii of gyration suggested partial unfolding of the N- and C-terminal domains during several of the production simulations. Nevertheless, the sizes of the periodic boxes employed for the different systems were always sufficiently large enough to keep periodic images of these domains from contacting each other and to produce artifacts. A representative analysis of the minimum distances between N- and C-terminal domains of the MOP receptor is shown in Figure 9, indicating, that these domains were always at least 3.5 nm apart, which is comparable with the thickness of the membrane bilayer.

#### 4.2.2. Dissociation of Molecular Parts

In the further analysis addressing the integrity of simulation systems dissociation of molecular components was checked in all MD trajectories. No dissociation of ligand or nucleotide (GTP or GDP) or notable displacement of protein components was observed for the MOP and CB1 receptors.

The systems including the β<sub>2</sub>AR were found to be intact for most simulations as no dissociation of epinephrine or the nucleotide or notable relative displacement of macromolecular

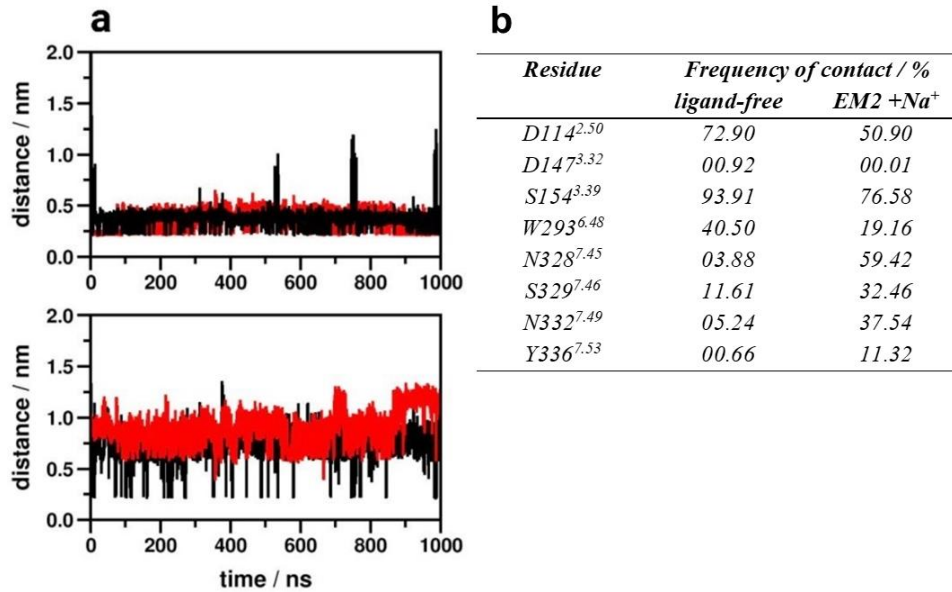


**Figure 10.** Ligand orientation in the 1<sup>st</sup> (b) and 2<sup>nd</sup> (c) replica simulations of the active  $\beta_2$ AR-G<sub>s</sub> protein-epinephrine complex in comparison with the crystallographic structure ((a), pdb code: 4LDO).

components were observed. However, epinephrine dissociated from the orthosteric binding pocket on two occasions during the production simulations. Epinephrine was first observed to depart from the binding pocket during the simulation of the inactive, G<sub>s</sub> protein-bound  $\beta_2$ AR after 600 ns, and then after 100 ns during the 3<sup>rd</sup> replica simulation of the active  $\beta_2$ AR bound to the G<sub>s</sub> protein. Interestingly, in the 2<sup>nd</sup> replica simulation of the active G<sub>s</sub> protein-bound  $\beta_2$ AR, the ligand took an opposite orientation in the binding pocket, compared to the x-ray crystallographic structure of the  $\beta_2$ AR-epinephrine complex (Figure 10). These simulations were not excluded from analysis but the results were interpreted accordingly. The instability of  $\beta_2$ AR-epinephrine complexes, observed during some of the simulations, may be attributed to the smaller size and remarkably lower affinity of epinephrine relative to the ligands used in previous simulations of the  $\beta_2$ AR ( $\mu$ M vs. pM range affinity, respectively) (Dror et al. 2011, Gregorio et al. 2017). In order to explain discrepancies emerging from spontaneous ligand dissociation during simulations, a reference simulation was carried out for the G<sub>s</sub> protein-bound active state  $\beta_2$ AR in which epinephrine was mildly restrained to the orthosteric binding pocket. A restraint with a force constant of 200 kJ mol<sup>-1</sup> nm<sup>-1</sup> sufficiently kept epinephrine from ejection.

#### 4.3. Allosteric Na<sup>+</sup> Binding

No penetration of Na<sup>+</sup> ions into the TM domain and the allosteric Na<sup>+</sup> binding site (D114<sup>2.50</sup>) was observed during the simulations of EM2-bound MOP receptors, while H<sub>2</sub>O molecules were observed to exchange frequently between the internal cavities of the TM domain and the bulk solvent phase. This suggests that for Na<sup>+</sup> ions the allosteric site is only accessible



**Figure 11.** (a) Minimum distance between Na<sup>+</sup> ions and the allosteric and orthosteric binding sites, D114<sup>2.50</sup> (top panel) and D147<sup>3.32</sup> (bottom panel), respectively, of the active state ligand-free MOP (black) and the MOP with bound by EM2 and allosteric Na<sup>+</sup> (red) during simulations. (b) The frequency of contact ( $d \geq 0.4$  nm) between Na<sup>+</sup> ions and polar amino acid side chains of the allosteric and orthosteric binding pockets.

through the orthosteric binding pocket and its entrance could be blocked by a bound ligand, whereas intracellular access to the TM domain is closed by the intracellular signalling proteins. Conversely, Na<sup>+</sup> rapidly localized at the allosteric site in the absence of bound EM2, following the transition of the receptor to an intermediate structural state. Interestingly, Na<sup>+</sup> binding to the orthosteric anchor residue, D147<sup>3.32</sup> was less frequent and often coincided the occasional dissociation of Na<sup>+</sup> from the allosteric site (Figure 11a). Spontaneous Na<sup>+</sup> penetration from the extracellular environment is in agreement with previously published MD simulation data (Yuan et al. 2013, Shang et al. 2014). Translocation of Na<sup>+</sup> ions through the active state MOP were also observed in previous MD simulations, but in the absence of bound ligands and intracellular proteins (Hu et al. 2019). The simulation of the active MOP-G<sub>i</sub> protein complex in the presence of EM2 at the orthosteric site, and Na<sup>+</sup> placed initially at the allosteric Na<sup>+</sup> binding site indicated, that the Na<sup>+</sup> ion in the allosteric site and its contacts with D114<sup>2.50</sup>, S154<sup>3.39</sup>, and W293<sup>6.48</sup> have become loose during the simulation, compared to the ligand-free system. Consequently, the frequency of close contacts with conserved polar/amphipathic residues further down towards the intracellular surface (N328<sup>7.45</sup>, N332<sup>7.49</sup>, and Y336<sup>7.53</sup>) have increased significantly (Figure 11b).

Similar results were obtained for the  $\beta_2$ AR.  $\text{Na}^+$  did not penetrate to interact with D113<sup>3,32</sup> and D79<sup>2,50</sup> in the presence of epinephrine. The localization of  $\text{Na}^+$  at the above sites/residues was observed only on the two occasions when ligand dissociation took place during the course of simulations. No  $\text{Na}^+$  insertion to the ortho- or allosteric sites was observed in any of the ligand- and  $\beta$ -arrestin-2-bound states. Frequent contacts were formed between  $\text{Na}^+$  and the ortho- and allosteric sites as well as residues of the conserved CWxP and NPxxY motifs in the ligand-free states, but no relevant trend of contact frequencies was identified which could be directly associated with the modulation of receptor activation by  $\text{Na}^+$  ions (Table 3). The presence of  $\text{Na}^+$  in the orthosteric pocket and the proximal CWxP motif was more prominent for the  $G_s$  protein and epinephrine-bound inactive state  $\beta_2$ AR. However, this was a clear consequence of ligand dissociation after approximately 600 ns of simulation time.

The orthosteric binding site of the CB1 receptor was continuously occupied by 2-AG through the simulations and the intercellular surface was stably bound by the  $G_i$  protein or  $\beta$ -arrestin-2 as it was reported in the previous subsection. Therefore, the entrance for  $\text{Na}^+$  ions was blocked and no  $\text{Na}^+$  penetration to the receptor core was observed. These observations are in complete agreement with previous MD simulation data of class A GPCRs and corroborate that the allosteric  $\text{Na}^+$  binding site is only accessible through the orthosteric binding pocket and the bound orthosteric ligand blocks the entrance of  $\text{Na}^+$  to the allosteric site. Intracellular access of  $\text{Na}^+$  ions through the TM domain is closed by the bound  $G_s$  protein complex or  $\beta$ -arrestin-2 (Yuan et al. 2013, Shang et al. 2014, Hu et al. 2019, Fleetwood et al. 2020).

#### 4.4. TM6 Dynamics

The disposition of TM6 of the transmembrane domain is a general indicator of the activation state of the class A GPCRs as it was shown by previous experimental results. Atomic displacement analysis of this TM helix of the EM2-bound MOP receptor revealed, that TM6 assumed intermediate conformations during simulations with minor changes from the corresponding starting structures (Figure 12a). This is in line with previous simulation results, where notable TM helix rearrangements were only observed at longer timescales and in the absence of bound signaling proteins (Dror et al. 2011). The largest disposition among the EM2-bound receptors was measured for the inactive MOP -  $\beta$ -arrestin-2 complex, which suggests that

**Table 3.** The frequency of contact ( $d \geq 0.4$  nm) between  $\text{Na}^+$  ions and polar amino acid side chains of the allosteric and orthosteric binding pockets and nearby conserved motifs

residue	frequency of contact / %							
	epinephrine-bound $\beta_2\text{AR}$				ligand-free $\beta_2\text{AR}$			
	G <sub>s</sub> protein complex		$\beta$ -arrestin-2 complex		G <sub>s</sub> protein complex		$\beta$ -arrestin-2 complex	
	active	inactive	active	inactive	active	inactive	active	inactive
<b>D</b> <sub>79</sub> <sup>2.50</sup>	0.0	0.0	0.0	0.0	52.15	93.41	25.91	26.04
<b>D</b> <sub>113</sub> <sup>3.32</sup>	< 0.1	32.33	0.0	0.0	9.22	45.06	5.02	72.03
<b>S</b> <sub>120</sub> <sup>3.39</sup>	0.0	0.0	0.0	0.0	44.95	96.36	5.55	25.86
<b>W</b> <sub>286</sub> <sup>6.48</sup>	0.0	6.88	0.0	0.0	3.36	19.51	29.16	10.92
<b>N</b> <sub>318</sub> <sup>7.45</sup>	0.0	0.0	0.0	0.0	2.78	0.46	46.14	0.22
<b>S</b> <sub>319</sub> <sup>7.46</sup>	0.0	0.0	0.0	0.0	30.52	14.39	79.75	24.69
<b>N</b> <sub>322</sub> <sup>7.49</sup>	0.0	0.0	0.0	0.0	45.87	51.12	8.17	20.06
<b>Y</b> <sub>326</sub> <sup>7.53</sup>	0.0	0.0	0.0	0.0	0.12	0.60	0.0	0.0

the active structural state of the receptor is preferred for  $\beta$ -arrestin-2 binding. TM6 of the ligand-free receptor, underwent much larger changes. (Figure 12b) This demonstrates remarkable stabilizing effect of the agonist, regardless of the activation state of the receptor.

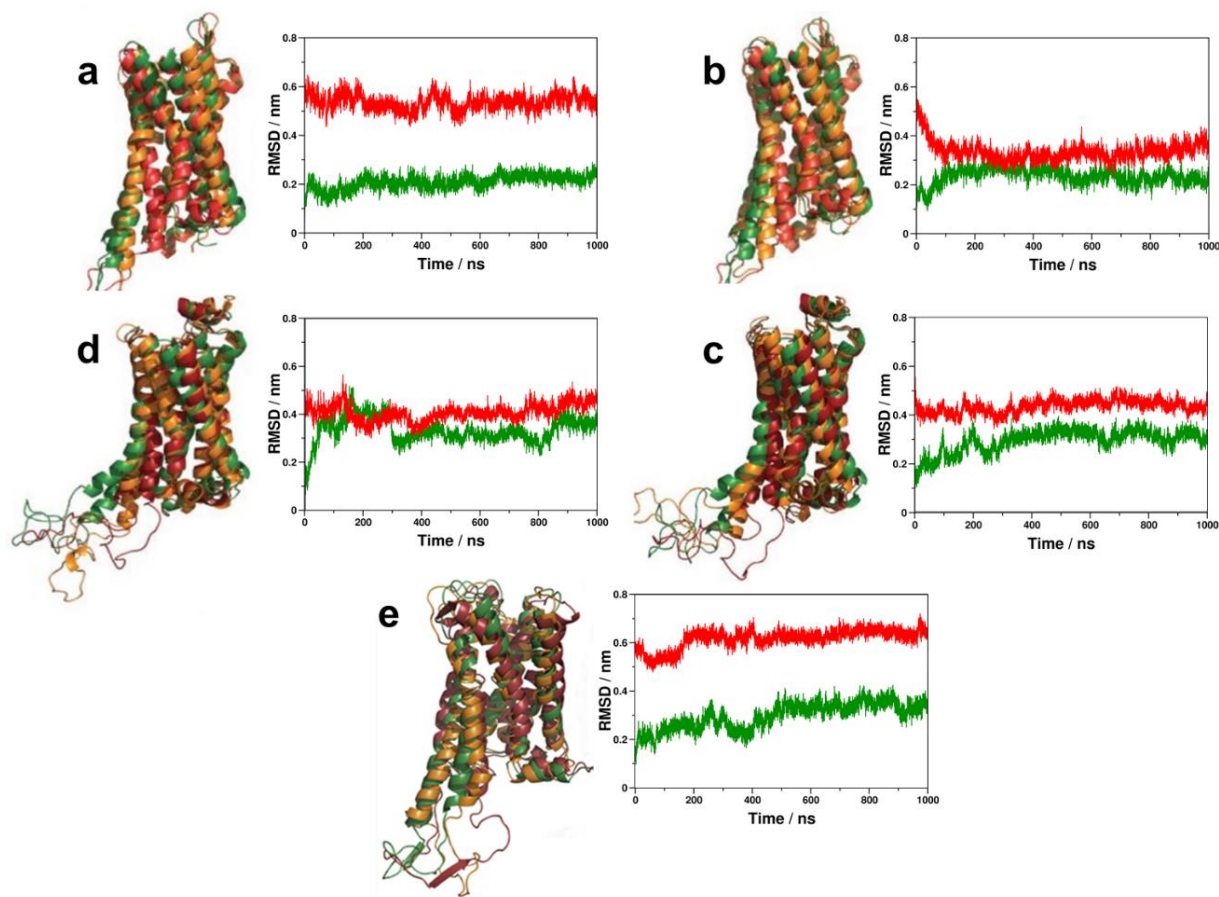
Similar results were observed for  $\beta_2\text{AR}$ , but the extent of TM6 displacement was more comparable in the epinephrine-bound and ligand-free states of the active receptor (Figure 12c-d). Therefore, the stabilizing effect of epinephrine appears to be less pronounced compared to that of EM2 bound to the MOP, which is acceptable considering the smaller molecular size and lower binding affinity of the former. In contrast with the results obtained for the corresponding MOP system, no significant TM6 disposition took place during the simulation of the  $\beta$ -arrestin-2-bound inactive  $\beta_2\text{AR}$  receptor. This suggests the preference of  $\beta$ -arrestin-2 for the inactive structure of  $\beta_2\text{AR}$ .

The above trends for the disposition of TM6 were followed by the CB1 receptor too, although no reference simulations have yet been done for the ligand-free CB1 (Figure 12e).

#### 4.5. Loop Dynamics and Intermolecular Interactions

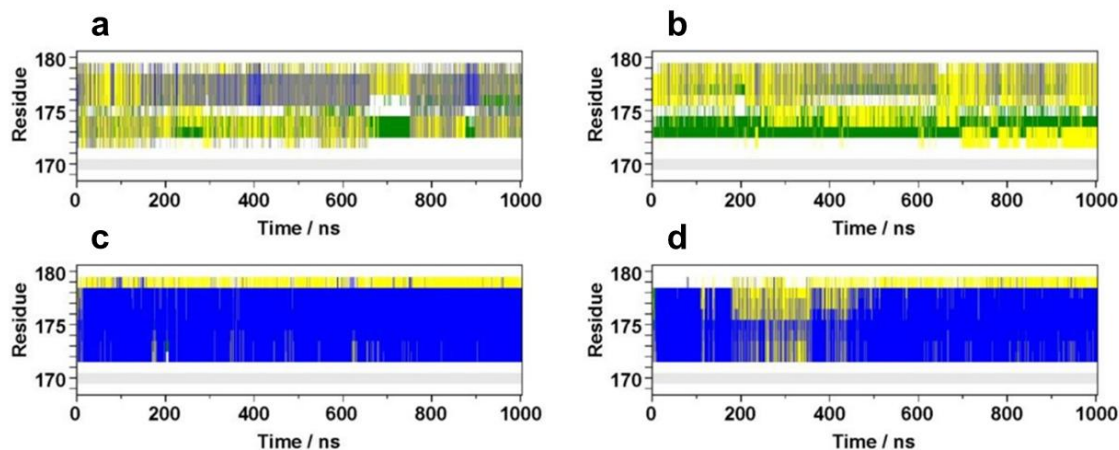
ICL1, ICL2, ICL3 and H8 form an interaction interface for G proteins on the intracellular surface of GPCRs. The evolution of secondary structure of these loops was assessed using the





**Figure 12.** Disposition of TM6 during simulations (orange) with respect to the active (green) and inactive (red) crystallographic structures of the (a) active, EM2 and  $G_i$  protein-bound MOP, (b) ligand-free  $G_i$  protein-bound MOP, (c) active, epinephrine and  $G_i$  protein-bound  $\beta_2AR$ , (d) active, ligand-free and  $G_i$  protein-bound  $\beta_2AR$ , (e) active, 2-AG and  $G_i$  protein-bound CB1.

DSSP method to reveal their structural preferences with regard to the bound signaling protein and the activation state of the receptor. ICL2 of the active MOP bound to EM2 were found to adopt stable  $\alpha$ -helical structure when bound to  $\beta$ -arrestin-2 and partially unfolded upon interaction with the  $G_i$  protein, independent of the state of the receptor (Figure 13). The presence of  $Na^+$  in the allosteric site along with the EM2 in the orthosteric site prevented this helix-to-coil transition. During the 1  $\mu s$  simulation of the epinephrine-bound  $\beta_2AR$ - $G_s$  protein complex, partial unfolding of ICL2 was observed for inactive state and  $\alpha$ -helical structure was maintained in the active state.

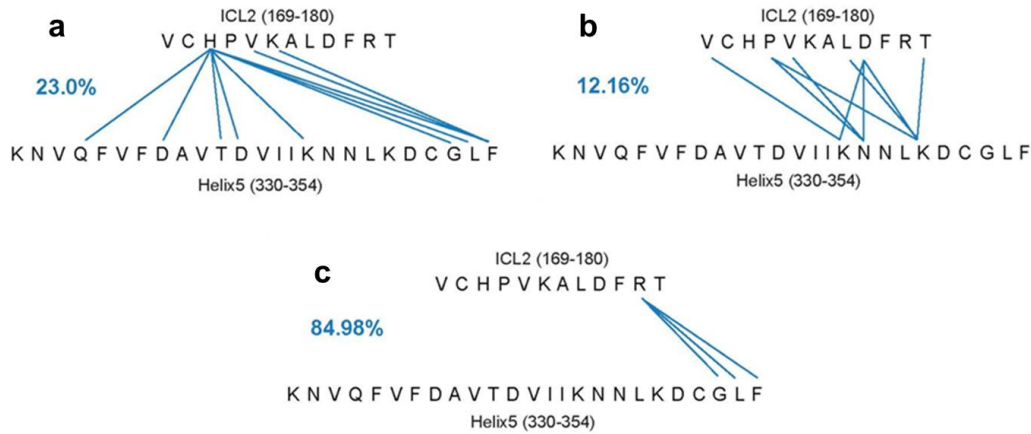


**Figure 13.** Evolution of the secondary structure of ICL2 during MOP simulations. (a) active receptor  $G_i$  protein complex (b) inactive receptor -  $G_i$  protein complex (c) active receptor -  $\beta$ -arrestin-2 complex (d) inactive receptor -  $\beta$ -arrestin-2 complex (white: unordered; green: bend; yellow:  $\beta$ -turn; black:  $\beta$ -bridge, grey:  $3_{10}$ -helix; blue:  $\alpha$ -helix).

These results obtained for the MOP and the  $\beta_2$ AR are both in agreement with recent experimental results that indicated that the  $\alpha$ -helical conformation of ICL2 of  $\beta_2$ AR is preserved when the receptor is interacting with the  $G_s$  protein complex and unfolds when  $\beta_2$ AR is bound to  $G_{i1}$  (Ma et al. 2020). Results presented here as well as the above cited experimental data suggest that ICL2 could be, at least partially responsible for G protein and/or signaling pathway specificity of different GPCRs, or GPCRs signaling through more than just one type of G protein complexes. Interestingly, ICL2 of the CB1 receptor showed partial unfolding during simulations of irrespective of receptor state and interacting partners. Further investigations of this receptor are needed to correlate these latter results with the above proposal.

The secondary structures of ICL1, ICL3 and H8 were maintained throughout the simulations of MOP,  $\beta_2$ AR and CB1 receptors with respect to their (active or inactive) starting structures. No major conformational transitions, initiated by the bound ligand and/or signaling proteins took place during the course of simulations. Only minor conformational fluctuations were observed mainly due to the internal dynamics of these molecular parts. The structures of ICL1, ICL3 and H8 of the  $\beta_2$ AR and CB1 receptors were highly similar independent of the functional state of the receptor and the presence of different signaling proteins. The biggest difference between these specific molecular parts was found for the ICL1 of the CB1 receptor, where ICL1 adopted a mostly  $3_{10}$ -helical conformation in the active state of CB1 and a more  $\alpha$ -helical in the





**Figure 14.** Frequency of H-bonds between ICL2 of MOP and H5G $\alpha$  expressed as percentages of the total conformational ensemble. (a) EM2-bound active receptor (b) EM2-bound active receptor in the presence of allosteric Na<sup>+</sup> (c) active receptor in the absence of EM2.

inactive state. The H8 of MOP was  $\alpha$ -helical in the inactive states and partially unfolded in the active states. These data suggest that the conformation of ICL1 and H8 may be controlled by the receptor state, whereas the structural adaptation of ICL2 is related to the bound signaling proteins and pathway specificity.

The ICL2 of the active, ligand-free, G<sub>i</sub> protein-bound MOP receptor, which maintained  $\alpha$ -helical conformation during simulation, also maintained strong contacts with the H5G $\alpha$  of the G<sub>i</sub> $\alpha$  subunit (Figure 14). H-bonds between the G<sub>i</sub> $\alpha$  subunit and the receptor involved only one participant residue of ICL2, as opposed to the EM2-bound systems where a partially unfolded ICL2 contributed more residues to this intermolecular interaction. Nevertheless, the frequency of that single H-bond in the ligand-free MOP indicated stronger and more specific interaction. This observation supports the hypothesis of pre-coupled GPCR-G protein complexes in the absence of ligands (Challiss et al. 2011). This also suggests that the lower frequency and specificity of intermolecular H-bonds in the EM2-bound active MOP may represent an intermediate complex state, which precedes G<sub>i</sub> protein dissociation during the signaling event (Challiss et al. 2011).

The ICL2-H5G<sub>s</sub> $\alpha$  contact was found to be very weak during simulations of the active state, epinephrine- and G<sub>s</sub> protein-bound  $\beta_2$ AR. Higher frequencies of H-bonds between ICL2 and H5G $\alpha$

were observed in inactive states and the absence of epinephrine. This may further support that the loss of interaction between ICL2 and H5G $\alpha$  in agonist-bound active states indicates the initiation of G<sub>s</sub> protein dissociation, although longer simulations would be needed to verify this assumption. Especially, if results obtained for the CB1 receptors are also considered, where ICL2 maintained H-bond interactions with the H5G $\alpha$  of the G<sub>i</sub> protein with high frequency in the active states, whereas the frequency of such contacts in the inactive state of the receptor were negligible. It has to be noted, however, that secondary structure analysis, presented above, did not indicate such potential specificity of ICL2 of CB1, unlike in the case of MOP and  $\beta_2$ AR.

A recent study of the neurotensin receptor 1 (NTSR1)- $\beta$ -arrestin-1 complex revealed, that the interface between  $\beta$ -arrestin-1 and NTSR1, including the finger loop (FL), is highly dynamic and the relative orientations captured by the Cryo-EM structure are likely to represent one of many conformational states (Yin et al. 2019). In this study ICL2 of the MOP was found to be the foremost participant of the MOP -  $\beta$ -arrestin-2 interaction, contacting the FL, the middle loop (ML) and the C-loop (CL) of  $\beta$ -arrestin-2. In case of the  $\beta_2$ AR all three intracellular loops of the receptor were found to contribute to the  $\beta_2$ AR- $\beta$ -arrestin-2 interaction, but no clear correlation between the activation state of the receptor and the pattern of these interactions was identified (Table 4). The overall frequency of interactions was highest for inactive, epinephrine-bound  $\beta_2$ AR suggesting that this structural state is the most preferred for  $\beta$ -arrestin-2 binding. However, taking into account that the ligand dissociated during the corresponding simulation, such an assumption cannot be taken. The second highest H-bond frequency between  $\beta_2$ AR and  $\beta$ -arrestin-2 was observed for the active, epinephrine-bound state. This latter apparent preference is corroborated by experimental data reporting the visual arrestin-bound and active, G<sub>T</sub> protein-bound structures of rhodopsin, which were almost identical (Zhan et al. 2011, Gao et al. 2019).

The interaction between the CB1 receptor and  $\beta$ -arrestin-2 was mainly mediated by ICL2 and the ML. The frequency of intermolecular salt bridges and H-bonds were the higher for the active state CB1, which suggests that  $\beta$ -arrestin-2 binds the active state CB1 prior to receptor deactivation.

**Table 4.** Representative table presenting the frequency of intermolecular H-bonds between  $\beta_2$ AR, the  $G_s\alpha$  protein and  $\beta$ -arrestin-2 expressed as percentages of the total conformational ensemble, generated by MD simulations.

interactions	residues involved, respectively	epinephrine-bound							ligand-free			
		active state				inactive state			active state		inactive state	
		$G_s$ protein complex				$\beta$ -arrestin-2	$G_s$ protein complex	$\beta$ -arrestin-2	$G_s$ protein complex	$\beta$ -arrestin-2	$G_s$ protein complex	$\beta$ -arrestin-2
		1	2	3	restrained							
H8-H8G $\alpha$	S329-L340; T369-L394	0.0	7.3	4.0	0.0	-	0.0	-	0.0	-	0.0	-
ICL1-H8G $\alpha$	F61-T66; T369-L394	0.2	0.1	0.0	0.2	-	0.6	-	30.3	-	40.0	-
ICL2-H8G $\alpha$	S137-T146; T369-L394	0.1	0.1	17.9	0.5	-	9.9	-	5.1	-	31.5	-
ICL3-H8G $\alpha$	E237-K267; T369-L394	90.5	92.5	97.5	96.3	-	77.1	-	76.2	-	76.7	-
H8-FL	S329-L340; G65-K78	-	-	-	-	33.4	-	0.0	-	0.0	-	0.0
ICL1-FL	F61-T66; G65-K78	-	-	-	-	25.6	-	58.8	-	22.9	-	49.3
ICL2-FL	S137-T146; G65-K78	-	-	-	-	8.7	-	30.4	-	77.8	-	35.6
ICL3-FL	E237-K267; G65-K78	-	-	-	-	85.5	-	71.7	-	14.8	-	14.0
H8-ML	S329-L340; P132-A140	-	-	-	-	0.0	-	0.0	-	0.0	-	0.0
ICL1-ML	F61-T66; P132-A140	-	-	-	-	0.9	-	0.1	-	0.0	-	0.0
ICL2-ML	S137-T146; P132-A140	-	-	-	-	7.1	-	4.2	-	9.3	-	30.3
ICL3-ML	E237-K267; P132-A140	-	-	-	-	0.0	-	0.0	-	0	-	0
H8-CL	S329-L340; V307-G317	-	-	-	-	0.0	-	0.0	-	0.0	-	0.0
ICL1-CL	F61-T66; V307-G317	-	-	-	-	0.0	-	0.0	-	0.0	-	0.0
ICL2-CL	S137-T146; V307-G317	-	-	-	-	4.6	-	58.4	-	0.2	-	8.9
ICL3-CL	E237-K267; V307-G317	-	-	-	-	0.0	-	0.0	-	4.0	-	0

#### 4.6. Intramolecular Interactions

The MD trajectories obtained in this study were analyzed for specific intramolecular salt bridges and H-bonds, involving conserved motifs, which were proposed previously to be relevant for class A GPCR activation. The analysis involved the following interactions: D<sup>3.49</sup>-R<sup>3.50</sup> (intra-

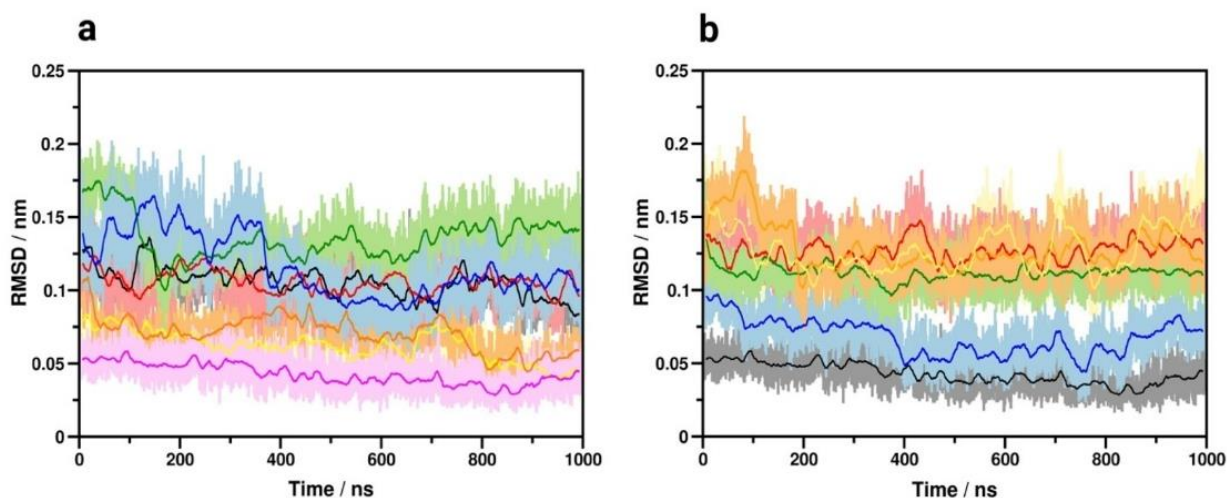
E/DRY), E/DRY motif-TM6 ionic lock, E/DRY-TM5, E/DRY-ICL2, CWxP-TM7, NPxxY-TM network (Cherezov et al. 2007, Wacker et al. 2010, Ring et al. 2013, Palczewski et al. 2000, Huang et al. 2015, Rasmussen et al. 2011, Jongejan et al. 2005). Interactions between the neighboring D<sup>3.49</sup> and R<sup>3.50</sup> of the E/DRY motif were generally more frequent in the inactive state and in the ligand-free MOP,  $\beta_2$ AR and CB1 receptors. On the other hand, considerable frequencies of salt bridges and H-bonds were observed in the active states of all three receptors too, although high-resolution experimental structures indicated coincidentally that this interaction is only present in inactive states and absent in active receptors. Formation of the ionic lock between the E/DRY motif (R<sup>3.50</sup>) and TM6 (E<sup>6.30</sup>) was only observed, in the inactive MOP and CB1 receptors. According to earlier proposals, this interaction acts as a constraint in the inactive state and gets disrupted upon receptor activation, followed by the release and disposition of TM6 (Palczewski et al. 2000). Mutation of residues of this ionic lock resulted in the elevated constitutional activity of the  $\beta_2$ AR, but the presence of this interaction was not corroborated by the crystallographic structures of this receptor (Ballesteros et al. 2001, Cherezov et al. 2007, Rasmussen et al. 2007). H-bonds between D<sup>3.49</sup> of the E/DRY motif and ICL2 residues were systematically present in all three receptors addressed in this study. In inactive state  $\beta_2$ AR D130<sup>3.49</sup> of the E/DRY motif was found to interact with S143 and L144 of ICL2, whereas in the active state ICL2 connected to the E/DRY motif through V141. This supports the discussion in the previous subsection, regarding the role of this loop in the activation mechanism. However, this specificity of interactions was only observed for this receptor. In the MOP and CB1 this interaction was found to be independent of the activation state of the receptors.

Frequent salt bridge formation was found between R165<sup>3.50</sup> (E/DRY) and D340<sup>8.47</sup> (H8) in the active state, EM2 and G<sub>i</sub> protein-bound MOP. This specific interaction was not described previously as, opposed to the previously discussed interactions, it was not evidently present in the reported high-resolution structures (Manglik et al. 2012, Huang et al. 2015, Koehl et al. 2018). Results presented here suggest, that this contact could be important for receptor activation and it is further supported by earlier mutation experiments (Liu et al. 2015). Opposed to this observation, intramolecular interactions between the E/DRY motif and H8 were missing from all  $\beta_2$ AR and CB1 systems examined here. Salt bridge formation is not facilitated between R<sup>3.50</sup> and S<sup>8.47</sup> in these receptors and no other potential, proximal partners were found in H8 that could participate in the formation of a salt bridge analogous to that between R165<sup>3.50</sup> and D340<sup>8.47</sup> of H8 in the active, G<sub>i</sub>

protein-bound MOP. No stable H-bond formation was indicated between the DRY motif and H8 either, hence this interaction is most likely a specific property of the MOP. No considerable trends were observed between the different receptor states and the frequencies of DRY-TM5, CWxP-TM7, and NPxxY-TM network interactions in any of the receptors within the time frame of simulations.

#### 4.7. Transmembrane Helix Properties

According to the hypothesis of this study, the activation of GPCRs is accompanied by a shift of macroscopic polarization in a shielded central duct of the TM domain, initiated by ligand binding and propagated to the intracellular G protein-binding interface. The inherent dipole moment of  $\alpha$ -helices can promote various conduction processes (electrons, protons, ions) along the helix axis (Hol 1985). Generally, the more ordered an  $\alpha$ -helical segment is, the higher its dipole moment. Therefore, to indirectly assess their potential role in the receptor activation, geometric features of TM helices were analyzed. The TM7 was found to be the most ordered among TM helices of the active MOP bound to the  $G_i$  protein and EM2 (Figure 15a). Furthermore, the helicity of TM7 was closest to ideal when the receptor was  $G_i$  protein-bound, and least ideal when it was complexed by  $\beta$ -arrestin-2, presumably providing TM7 with the highest dipole moment in the  $G_i$



**Figure 15.** Properties of transmembrane helices of the MOP receptor. (a) Deviation from ideal  $\alpha$ -helical geometry in the  $G_i$  protein-bound active state of MOP. Black: TM1, red: TM2, green: TM3, blue: TM4, yellow: TM5, orange: TM6, magenta: TM7. (b) Deviation of TM7 from ideal  $\alpha$ -helical geometry in the active,  $G_i$  protein-bound (black),  $\beta$ -arrestin-2 bound (red), Nb39 nanobody-bound (blue), T4-lysozyme-fused (green),  $G_i$  protein-bound, ligand free (orange) and  $G_i$  protein, EM2 and allosteric  $Na^+$  bound states (yellow).

protein-bound state, compared to all other receptor states (Figure 15b). This indirectly supports the involvement of the inherent dipole moment of TM7 in the activation mechanism and it is corroborated by previous reports (Fenalti et al. 2014, Bartuzi et al. 2016).

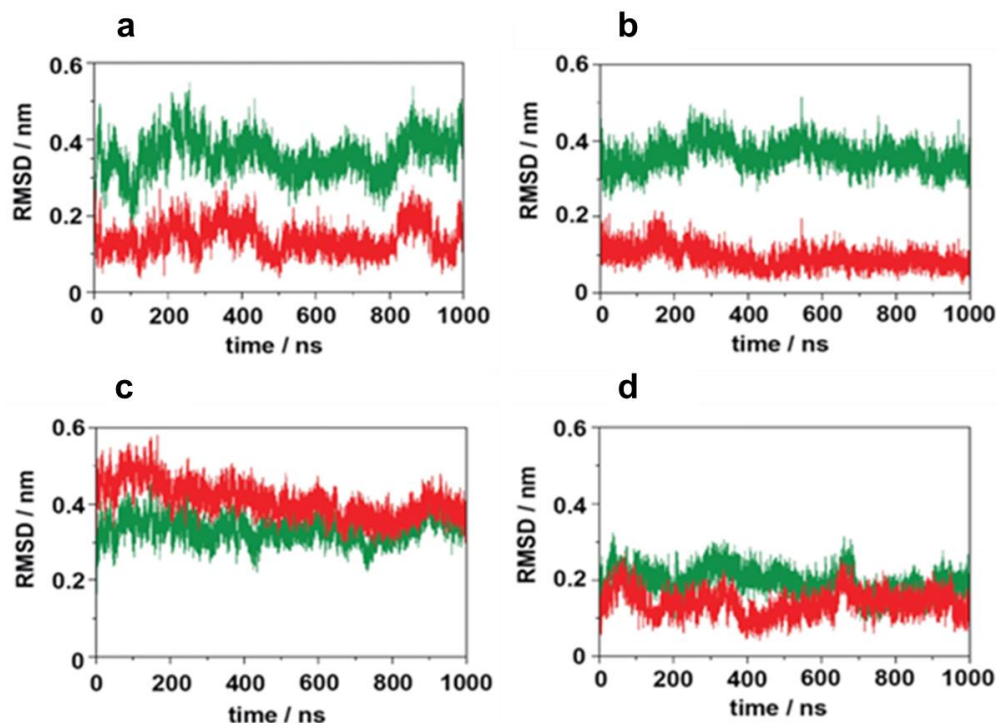
Such a proposed role of TM7 was not corroborated by the results obtained here for  $\beta_2$ AR and CB1 receptors. As opposed to the MOP, TM7 was found to be among the least ordered TM helices of the active, agonist and G protein-bound  $\beta_2$ AR and CB1 receptors. Furthermore, the slightly higher order of TM7 in the active, epinephrine- and  $G_s$  protein-bound  $\beta_2$ AR, relative to that of the other receptor states and complexes was not reproduced in the reference simulation of that system and was also matched by the  $\beta$ -arrestin-2 and epinephrine-bound receptor. These discrepancies suggest that the above-described potential role of TM7 is a specific property of MOP.

#### **4.8. NPxxY Motif Dynamics**

The conserved NPxxY motif in the junction of TM7 and H8 is associated with stabilizing intermediate conformations in ligand-free class A GPCRs that facilitate G protein insertion (Dror et al. 2011; Manglik et al. 2012, Huang et al. 2015, Liu et al. 2015). The surprisingly high disorder of TM7 observed in  $\beta_2$ AR (see pervious subsection), prompted the analysis of the disposition of this helical segment to see if the dynamics of this motif shows any correlation with the activation state of the receptor or with the presence of agonist and/or intracellular proteins. Thus, the displacement of this motif has been calculated in the  $\beta_2$ AR receptor simulation systems. Interestingly, relatively large dispositions ( $\sim 0.4$  nm RMSD of backbone atoms) of the NPxxY motif were found to coincide with the intense concerted dynamics of the second segment of a series of conserved polar amino acid side chains located close to the intracellular surface of the receptor (Figure 16). This suggests, that the elevated mobility of the NPxxY motif could be associated either with receptor activation or constitutional activity (discussed in detail in the next subsection).

#### **4.9. Correlated Side-chain Motions in the Transmembrane Domain**

Dynamic cross-correlation matrix (DCCM) analysis was first performed for the MOP, involving ECLs, ICLs and the TM domain. This analysis revealed that the orthosteric binding pocket is connected to the intracellular surface of the receptor through a channel of polar amino



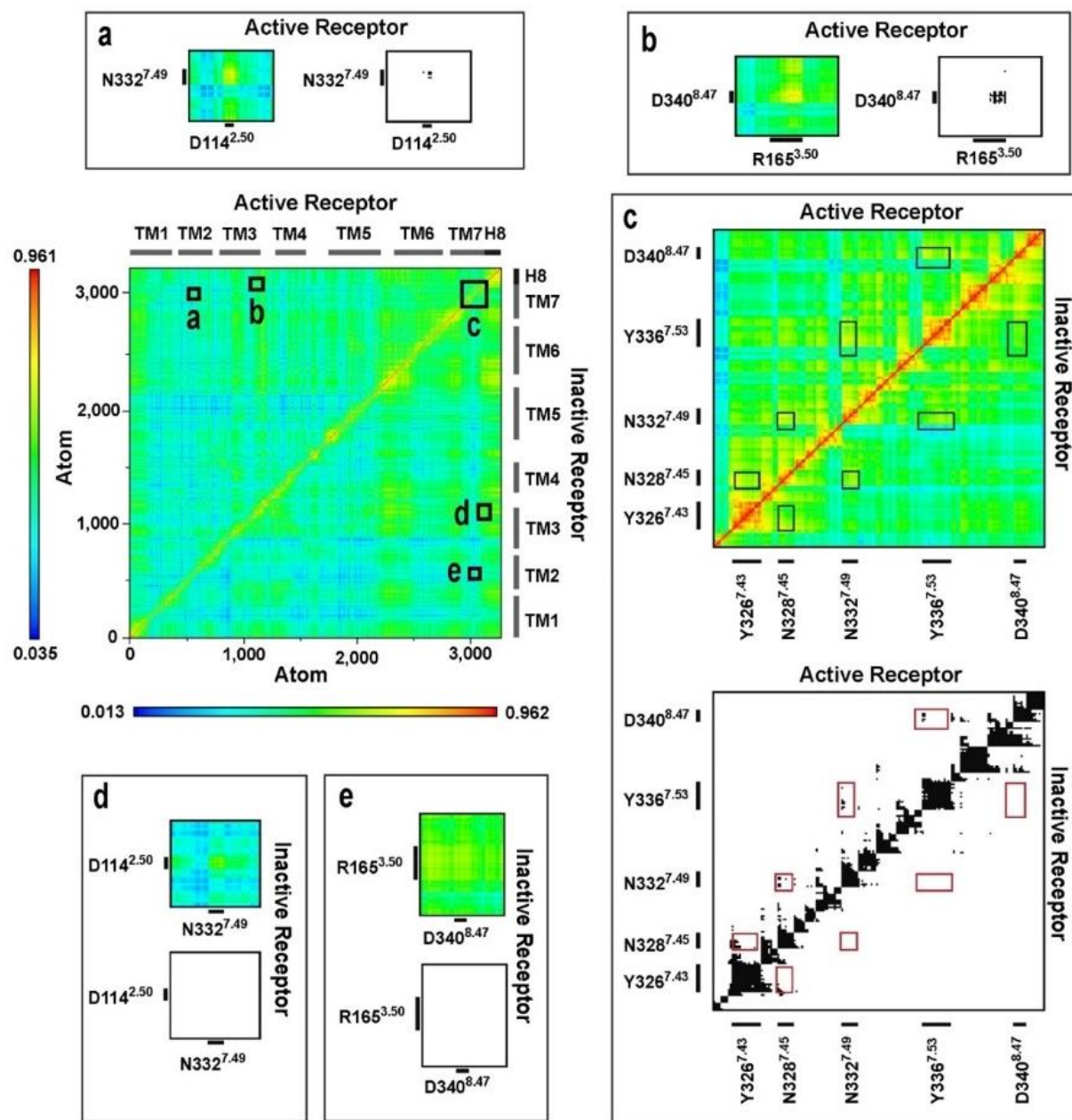
**Figure 16.** Disposition of the NPxxY motif during simulations with respect to the active (green) and inactive (red) crystallographic structures of  $\beta_2$ AR. (a) active  $\beta_2$ AR- $G_s$  protein-epinephrine complex, 1st replica; (b) inactive  $\beta_2$ AR- $G_s$  protein-epinephrine complex; (c) ligand-free active  $\beta_2$ AR- $G_s$  protein complex; (d) ligand-free inactive  $\beta_2$ AR- $G_s$  protein complex.

acid residues, of which motions are highly correlated (Figure 17-18). Such concerted motions were observed only for the active receptor- $G_i$  protein complex and none for the other reference systems. Interestingly, no such concerted motions of these polar amino acid side chains were found in the ligand-free receptor, although it was expected, considering that the TM domain demonstrated a higher degree of conformational flexibility, which may be associated with the constitutive activity of the receptor. The decoupling of these correlated motions may be caused by the notable disposition of transmembrane helices observed during the simulation of the ligand-free receptor. Furthermore, the absence of the positively charged EM2 in the orthosteric binding site and the simultaneous presence of a  $Na^+$  ion at the allosteric site represents a charge shift between the two sites, which could also lead to the freezing of motions of residues located in the proximity of the allosteric site (N328<sup>7.45</sup>, N332<sup>7.49</sup>). Residues of the above identified polar signaling channel are located mostly on the TM7, in the inner core of the TM helical bundle, shielded from the

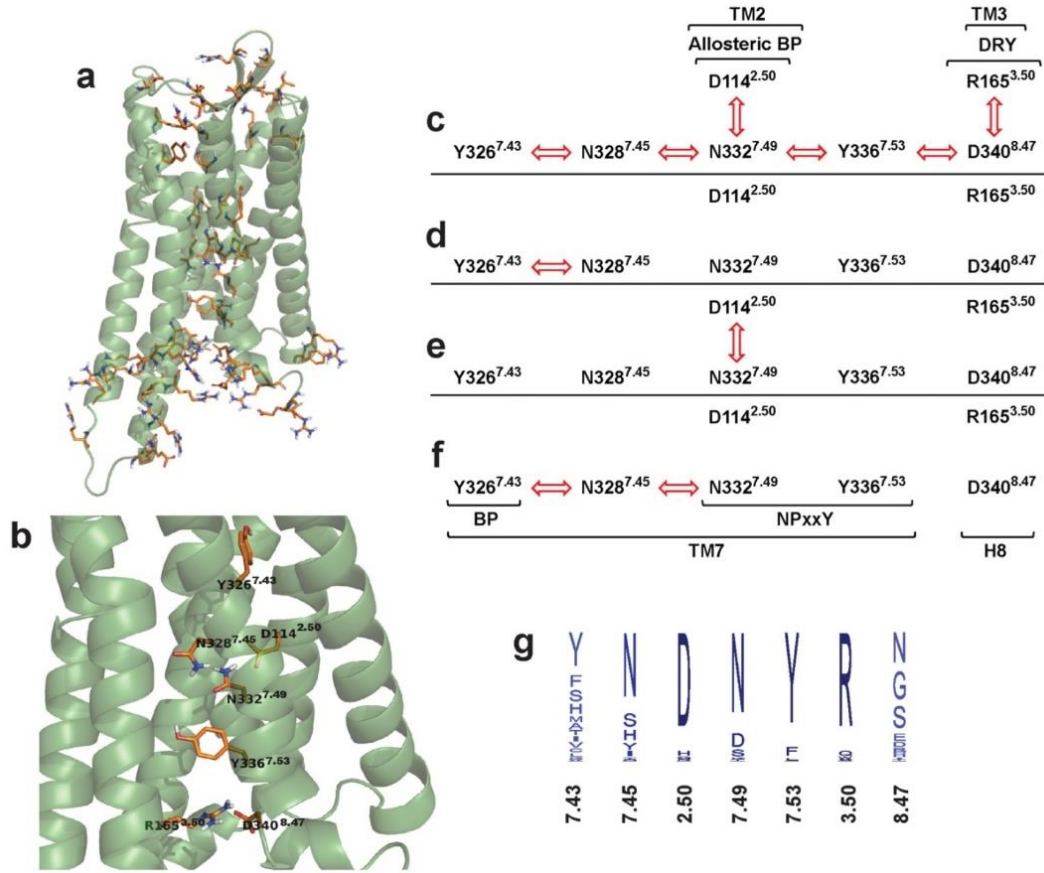
surrounding membrane environment. All of these channel residues are from highly conserved functional motifs, except for Y326<sup>7.43</sup> in the binding pocket and N340<sup>8.47</sup> at the G protein-binding interface. Variability of these two residues could be associated with ligand and G protein specificity, respectively. In class A GPCRs there is a strong complementary relationship between functional motifs, except for Y326<sup>7.43</sup> in the binding pocket and N340<sup>8.47</sup> at the G protein-binding interface. Variability of these two residues could be associated with ligand and G protein specificity, respectively. In class A GPCRs there is a strong complementary relationship between the X<sup>7.45</sup> residue of the binding pocket and the physico-chemical properties of the endogenous ligands, which provides further support for this assumption. Analysis of the individual dynamics of these specific side chains revealed that the observed movements are small and mostly occur without the transition between rotameric states. However, small conformational changes of polar or charged species could result in significant alterations in the local electronic structures, which could then propagate along the molecule. The involvement of protons and water molecules in the transmission of signal could also be presumed, since the orthosteric binding pocket and the G protein-binding interface is connected through a hydrated pathway and water molecules were observed to exchange rapidly between the internal cavities of the TM domain during simulations. In light of the above results and discussions, the initial hypothesis of this study proposing a parallel change of macroscopic polarization in a shielded central duct of the TM domain during GPCR activation could be deemed plausible. Classical force field methods cannot provide quantitative details of processes involving intramolecular polarization, but independent mutation data provides direct evidence for the interplay of these polar and charged amino acid side chains during receptor activation. Impaired G protein signaling, or elevation of constitutive activity, was observed for mutant receptors, where residues of the above of mentioned polar signaling channel were replaced, while receptor activity was preserved in double mutants, where the net charge of channel residues was kept intact (Fenalti et al. 2014, Jongejan et al. 2005, Liu et al. 2015, Hothersall et al. 2017, Sealton et al. 1995, Xu et al. 1999, Gales et al. 2000, Barak et al. 1994, Prioleau et al. 2002, Kalatskaya et al. 2004).

The high degree of conservation of polar signaling channel residues suggests that the above theory could be extended to other class A GPCRs. DCCM analysis of the MD simulation trajectories of  $\beta_2$ AR and CB1 receptors provided corroborating results (Figure 19-20). However, there are several differences between the channel residues of the three receptors. The first



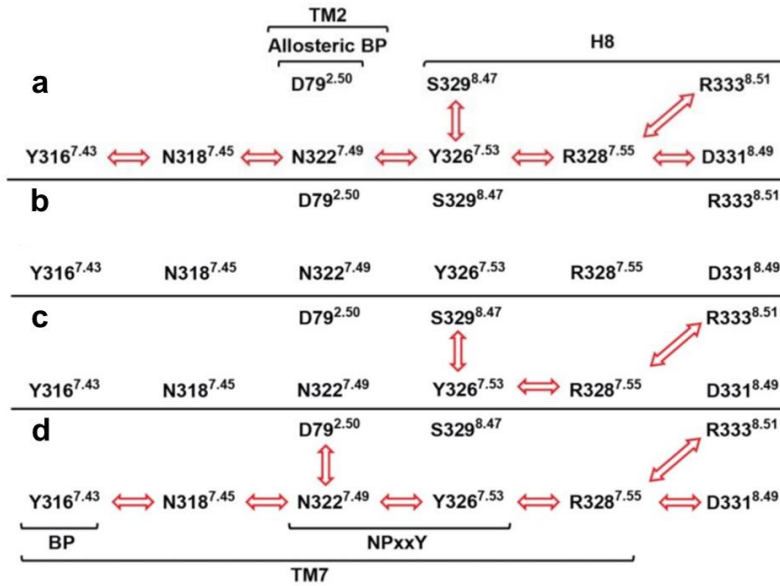


**Figure 17.** Dynamic cross-correlation matrices of the G protein-bound MOP in active and inactive states. Panels (a–e) are magnified views of regions of amino acid residues of interest. Black and white panels show correlations above the threshold value of 0.65 of the correlation coefficients.



**Figure 18.** The polar signaling channel of the MOP revealed by DCCM analysis. (a) Polar amino acids of which motions are correlated in the G<sub>i</sub> protein-bound active state, (b) polar amino acids of which motions are correlated and linking the BP and intracellular surface. Diagram of the channel residues in (c) active receptor-G<sub>i</sub> protein, (d) inactive receptor-G<sub>i</sub> protein, (e) active receptor-β-arrestin-2, (f) inactive receptor-β-arrestin-2 complexes. Red arrow indicates correlated motion of the respective amino acids. (g) Degree of conservation of polar signaling channel residues among class A GPCRs.

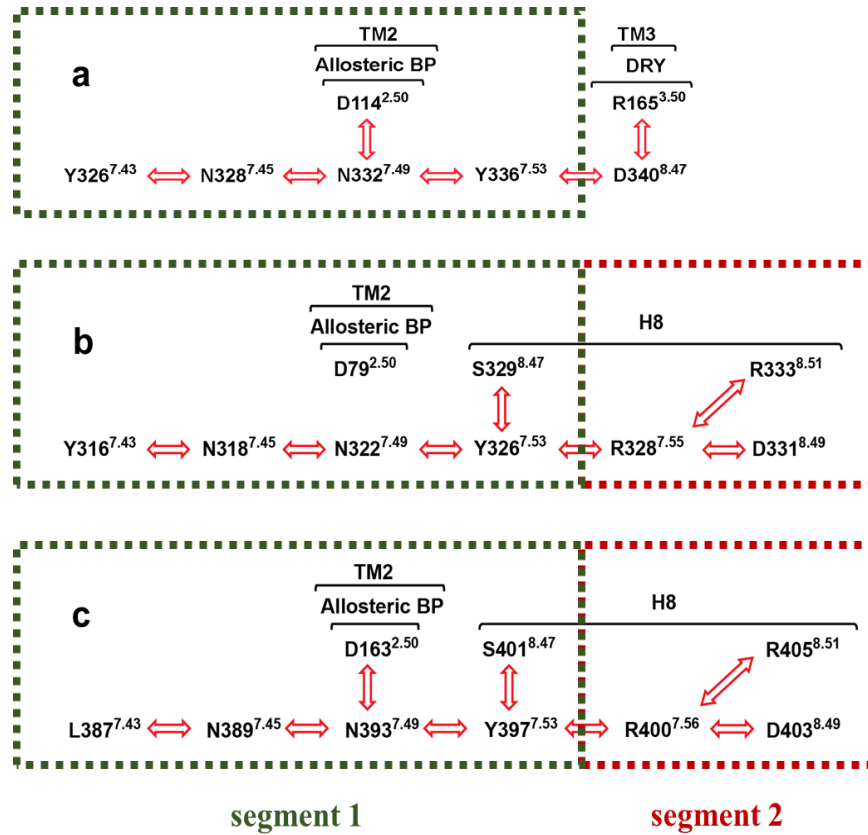
difference is that the residues of the E/DRY motif are not involved in correlated motions in the β<sub>2</sub>AR and CB1 receptors. In the active, EM2 and G<sub>i</sub> protein-bound MOP a salt bridge between R165<sup>3.50</sup> of the E/DRY motif and D340<sup>8.47</sup> of H8 was observed with high frequency. Consequently, the motions of these residues were in intense correlation. The analogous S329<sup>8.47</sup> and S401<sup>8.47</sup> residues of the β<sub>2</sub>AR and CB1 receptors, respectively, cannot form salt bridges, which may explain why the E/DRY motif is not coupled to the proposed signaling channel (see further discussion in subsection 4.6). Nevertheless, the coupling of motions of S329<sup>8.47</sup> and S401<sup>8.47</sup> to those of the NPxxY motif in were maintained in β<sub>2</sub>AR and CB1 receptors, respectively.



**Figure 19.** The polar signaling channel of the  $\beta_2$ AR identified by dynamic cross-correlation analysis. (a) Active  $\beta_2$  AR–Gs protein–epinephrine complex, 1st replica; (b) active  $\beta_2$  AR–Gs protein–epinephrine complex, 2nd replica; (c) active  $\beta_2$  AR–Gs protein–epinephrine complex, 3rd replica; (d) active  $\beta_2$  AR–Gs protein–epinephrine complex, restrained ligand. Red arrows indicate correlated motions of the respective amino acids.

Furthermore, D331<sup>8.49</sup> and R333<sup>8.51</sup> residues of the H8 of  $\beta_2$ AR and D403<sup>8.49</sup> and R405<sup>8.51</sup> residues of the H8 of CB1 showed a high degree of correlation. Among these residues R<sup>8.51</sup> is conserved and was suggested previously to be important for the G protein-coupling of the adenosine A<sub>2B</sub> receptor (Liu et al. 2015). A further difference observed between the MOP and  $\beta_2$ AR receptors is that the allosteric Na<sup>+</sup> binding site (D79<sup>2.50</sup>) is less intensely involved in the correlated motions of channel residues of  $\beta_2$ AR. However, this correlation was present during the reference simulation when epinephrine was mildly restrained to the orthosteric binding pocket, suggesting that the presence of a strongly bound, correctly oriented agonist initiates coupling of D79<sup>2.50</sup> to the signaling cascade.

A specific feature of the polar signaling channels of  $\beta_2$ AR and CB1 receptors is that they could be subdivided into two segments. The first segment spans the orthosteric (Y316<sup>7.43</sup>/L387<sup>7.43</sup>) and allosteric (D79<sup>2.50</sup>/D163<sup>2.50</sup>) binding pockets, N318<sup>7.45</sup>/N389<sup>7.45</sup> and the NPxxY motif



**Figure 25.** Comparison of the polar signaling channels in the (a) active MOP-G<sub>i</sub> protein-EM2 complex, the (b) active β<sub>2</sub>AR-G<sub>s</sub> protein-epinephrine complex, and the (c) active CB1-G<sub>i</sub> protein-2-AG complex.

(N322<sup>7.49</sup> and Y326<sup>7.53</sup>; N393<sup>7.49</sup> and Y397<sup>7.53</sup>), while the second segment shares the last residue of the NPxxY motif (Y326<sup>7.53</sup>/Y397<sup>7.53</sup>) and includes the tip of TM7 (R328<sup>7.55</sup>/R400<sup>7.56</sup>) and three H8 residues (S329<sup>8.47</sup>, D331<sup>8.49</sup>, and R333<sup>8.51</sup>; S401<sup>8.47</sup>, D331<sup>8.49</sup>, and R333<sup>8.51</sup>). The rationale behind this subdivision is given by the NPxxY disposition data obtained for β<sub>2</sub>AR, discussed earlier in subsection 4.8. The relatively large, approximately 0.4 nm (RMSD) disposition of the NPxxY motif coincided with intense concerted motions in the second segment of the polar signaling channel. Such movements were observed in inactive states and in the absence of ligand, whereas the full sequence of correlated motions was incomplete in those systems. This suggests that the concerted dynamics of this second segment indicate the constitutive activity of the β<sub>2</sub>AR and CB1 receptors. This was supported by that correlated motions of this second segment were decoupled upon β-arrestin-2 binding to β<sub>2</sub>AR and CB1 and/or if epinephrine was bound to β<sub>2</sub>AR in the wrong relative orientation (see subsection 4.2.2), presumably stabilizing a conformational

state of the receptor that is inappropriate for signaling. Further support is provided by a previous study, where a stable intermediate structure was identified in the absence of ligands, representing a receptor conformation that facilitates G<sub>s</sub> protein insertion and suggests that the structural changes that accompany the activation start at the intracellular side of the receptor. The role of the NPxxY motif in that process was highly emphasized (Dror et al. 2011).

## 5. Conclusions

The above results and independent literature data suggest that the potential contribution of the electrostatic balance in the TM domain, proposed in this thesis, is warranted for detailed, quantitative examination by means of experimental and more accurate theoretical methods. The general features of GPCR activation proposed here and the receptor-specific, characteristic details could provide alternative opportunities for the discovery of a new class of GPCR drugs. The extended perspective of the activation mechanism, if further pursued, may provide a deeper explanation for ligand-induced effects in multiple functional states and could help to identify and quantitatively assess specific physico-chemical properties of GPCR ligands that furnish different functional properties.

## 6. Summary of Findings

1. The structure of ICL2 has an important role in signaling protein specificity of the MOP and  $\beta_2$ AR receptors.
2. Localization  $\text{Na}^+$  ions in the allosteric binding site of class A GPCRs takes place from the extracellular side. The allosteric site is only accessible through the orthosteric binding pocket and its entrance is blocked by a bound ligand, whereas cytosolic access to the TM domain is closed by the bound intracellular signaling proteins.
3. The precise orientation of the agonist is crucial to initiate signal transduction of GPCRs. Minor displacements from the functional binding mode could arrest receptor function through stabilizing a conformational of state that is inappropriate for signaling.
4. The orthosteric binding pocket of the three GPCRs studied here is connected to the intracellular G protein-binding interface through a channel of conserved polar amino acid residues, of which motions correlate. This polar signaling channel may assist the shift of macroscopic polarization in a shielded central duct of the TM domain, leading to G protein activation.
5. The polar signaling channel could be subdivided into two segments. The first segment starting from the orthosteric binding pocket could be responsible for the transmission of the agonist induced signal, whereas the second segment, closer to the intracellular surface, could be related to the constitutive activity of the receptor.

## References

- Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015, 1-2: 19-25.
- Alegre KO, Paknejad N, Su M, Lou JS, Huang J, Jordan KD, Eng ET, Meyerson JR, Hite RK, Huang XY. Structural basis and mechanism of activation of two different families of G proteins by the same GPCR. *Nat Struct Mol Biol*. 2021, 28(11): 936-944.
- Attwood TK, Findlay JBC. Fingerprinting G-protein-coupled receptors. *Protein Engineering*. 1994, 7: 195–203.
- Ballesteros JA, Jensen AD, Liapakis G, Rasmussen SG, Shi L, Gether U, Javitch JA. Activation of the  $\beta$ 2 adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6. *J. Biol. Chem*. 2001, 276: 29171-29177.
- Barak LS, Tiberi M, Freedman NJ, Kwatra MM, Lefkowitz RJ, Caron MG. A Highly Conserved Tyrosine Residue in G Protein-Coupled Receptors is Required for Agonist-Mediated Beta 2-Adrenergic Receptor Sequestration. *J. Biol. Chem*. 1994, 269: 2790–2795.
- Bartuzi D, Kaczor AA, Matosiuk D. Interplay between two allosteric sites and their influence on agonist binding in human  $\mu$ -opioid receptor. *J. Chem. Inf. Model*. 2016, 56: 563–570.
- Brogi S, Tafi A, Désaubry L, Nebigil CG. Discovery of GPCR ligands for probing signal transduction pathways. *Front Pharmacol*. 2014, 5: 255.
- Buchoux S. FATS LIM: a fast and robust software to analyze MD simulations of membranes. *Bioinformatics*. 2017, 33(1): 133-134.
- Calizo RC, Scarlata S. A role for G-proteins in directing G-protein-coupled receptor-caveolae localization. *Biochemistry*. 2012, 51(47): 9513-9523.
- Challiss RA, Wess J. GPCR-G Protein Preassembly? *Nat. Chem. Biol*. 2011, 7: 657–658.
- Che T, Majumdar S, Zaidi SA, Ondachi P, McCorvy JD, Wang S, Mosier PD, Uprety R, Vardy E, Krumm BE, Han GW, Lee MY, Pardon E, Steyaert J, Huang XP, Strachan RT, Tribo AR, Pasternak GW, Carroll FI, Stevens RC, Cherezov V, Katritch V, Wacker D, Roth BL. Structure of the Nanobody-Stabilized Active State of the Kappa Opioid Receptor. *Cell*. 2018, 172(1-2): 55-67.

- Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC. High-resolution crystal structure of an engineered human  $\beta_2$ -adrenergic G protein-coupled receptor. *Science*. 2007, 318: 1258–1265.
- Chini B, Parenti M. G-protein coupled receptors in lipid rafts and caveolae: How, when and why do they go there? *J. Mol. Endocrinol.* 2004, 32: 325–338.
- de Oliveira PG, Ramos MLS, Amaro AJ, Dias RA, Vieira SI. Gi/o-Protein Coupled Receptors in the Aging Brain. *Front Aging Neurosci.* 2019, 11:89.
- Deganutti G, Cuzzolin A, Ciancetta A, Moro S. Understanding allosteric interactions in G protein-coupled receptors using supervised molecular dynamics: a prototype study analysing the human A3 adenosine receptor positive allosteric modulator. *Bioorg. Med. Chem.* 2015, 23: 4065–4071.
- Dror RO, Arlow DH, Borhani DW, Jensen MØ, Piana S, Shaw DE. Identification of two distinct inactive conformations of the  $\beta_2$ -adrenergic receptor reconciles structural and biochemical observations. *PNAS*. 2009, 106: 4689–4694.
- Dror RO, Arlow DH, Maragakis P, Mildorf TJ, Pan AC, Xu H, Borhani DW, Shaw DE. Activation mechanism of the  $\beta_2$ -adrenergic receptor. *PNAS*, 2011, 108: 18684–18689.
- Dror RO, Pan AC, Arlow DH, Borhani DW, Maragakis P, Shan Y, Xu H, Shaw DE. Pathway and mechanism of drug binding to G-protein-coupled receptors. *PNAS*. 2011, 108(32): 13118–13123.
- Farrens DL, Altenbach C, Yang K, Hubbell WL, Khorana HG. Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *Science*. 1996, 274(5288): 768–70.
- Fenalti G, Giguere PM, Katritch V, Huang XP, Thompson AA, Cherezov V, Roth BL, Stevens RC. Molecular control of  $\delta$ -opioid receptor signaling. *Nature*. 2014, 506: 191–196.
- Filipek, S. Molecular switches in GPCRs. *Curr. Opin. Struct. Biol.* 2019, 55: 114–120.
- Fleetwood O, Matricon P, Carlsson J, Delemotte L. Energy Landscapes Reveal Agonist Control of G Protein-Coupled Receptor Activation via Microswitches. *Biochemistry*. 2020, 59: 880–891.
- Galés C, Kowalski-Chauvel A, Dufour MN, Seva C, Moroder L, Pradayrol L, Vaysse N, Fourmy D, Silvente-Poirot S. Mutation of Asn-391 within the Conserved NPxxY Motif of the Cholecystokinin B Receptor Abolishes  $G_q$  Protein Activation without Affecting Its Association with the Receptor. *J. Biol. Chem.* 2000, 275: 17321–17327.



- Gao Y, Hu H, Ramachandran S, Erickson JW, Cerione RA, Skiniotis G. Structures of the Rhodopsin-Transducin Complex: Insights into G-Protein Activation. *Mol. Cell.* 2019, 75: 781–790.
- Ghanouni P, Steenhuis JJ, Farrens DL, Kobilka BK. Agonist-induced conformational changes in the G-protein-coupling domain of the beta 2 adrenergic receptor. *PNAS.* 2001, 98(11): 5997-6002.
- Gregorio GG, Masureel M, Hilger D, Terry DS, Juetten M, Zhao H, Zhou Z, Perez-Aguilar JM, Hauge M, Mathiasen S, Javitch JA, Weinstein H, Kobilka BK, Blanchard SC. Single-molecule analysis of ligand efficacy in  $\beta$ 2AR-G-protein activation. *Nature.* 2017, 547: 68–73.
- Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis.* 1997, 18: 2714-2723.
- Gupta R, Jung E, Brunak S. Prediction of N-glycosylation sites in human proteins. In preparation. <http://www.cbs.dtu.dk/services/NetNGlyc/>. 2004.
- Haga K, Kruse AC, Asada H, Yurugi-Kobayashi T, Shiroishi M, Zhang C, Weis WI, Okada T, Kobilka BK, Haga T, Kobayashi T. Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature.* 2012, 482: 547–551.
- Hausdorff WP, Bouvier M, O'Dowd BF, Irons GP, Caron MG, Lefkowitz RJ. Phosphorylation sites on two domains of the  $\beta$ 2-adrenergic receptor are involved in distinct pathways of receptor desensitization. *J. Biol. Chem.* 1989, 264(21): 12657-65.
- Hauser AS, Attwood MM, Rask-Andersen M, Schiöth HB, Gloriam DE. Trends in GPCR drug discovery: new agents, targets and indications. *Nat. Rev. Drug Discov.* 2017, 16: 829–842.
- Head BP, Patel HH, Roth DM, Lai NC, Niesman IR, Farquhar MG, Insel PA. G-protein-coupled Receptor Signaling Components Localize in Both Sarcolemmal and Intracellular Caveolin-3-associated Microdomains in Adult Cardiac Myocytes. *J. Biol. Chem.* 2005, 280: 31036–31044.
- Hol WG. Effects of the  $\alpha$ -helix dipole upon the functioning and structure of proteins and peptides. *Adv. Biophys.* 1985, 19: 133–165.
- Hothersall JD, Torella R, Humphreys S, Hooley M, Brown A, McMurray G, Nickolls SA. Residues W320 and Y328 within the Binding Site of the  $\mu$ -Opioid Receptor Influence Opiate Ligand Bias. *Neuropharmacology.* 2017, 118: 46–58.

- Howlett AC, Blume LC, Dalton GD. CB1 cannabinoid receptors and their associated proteins. *Curr. Med. Chem.* 2010, 17(14): 1382-93.
- Howlett AC, Champion-Dorow TM, McMahon LL, Westlake TM. The cannabinoid receptor: biochemical and cellular properties in neuroblastoma cells. *Pharmacol. Biochem. Behav.* 1991, 40(3): 565-9.
- Hu X, Wang Y, Hunkele A, Provasi D, Pasternak GW, Filizola M. Kinetic and thermodynamic insights into sodium ion translocation through the  $\mu$ -opioid receptor from molecular dynamics and machine learning analysis. *PLoS Comput. Biol.* 2019, 15: e1006689.
- Huang P, Chen C, Xu W, Yoon SI, Unterwald EM, Pintar JE, Wang Y, Chong PL, Liu-Chen LY. Brain region-specific N-glycosylation and lipid rafts association of the rat mu opioid receptors. *Biochem. Biophys. Res. Commun.* 2008, 365(1): 82-88.
- Huang W, Manglik A, Venkatakrishnan AJ, Laeremans T, Feinberg EN, Sanborn AL, Kato HE, Livingston KE, Thorsen TS, Kling RC, Granier S, Gmeiner P, Husbands SM, Traynor JR, Weis WI, Steyaert J, Dror RO, Kobilka BK. Structural insights into  $\mu$ -opioid receptor activation. *Nature.* 2015, 524(7565): 315-321.
- Humphrey W, Dalke A, Schulten K. VMD—Visual Molecular Dynamics. *J. Mol. Graph.* 1996, 14: 33–38.
- Inagaki S, Ghirlando R, White JF, Gvozdenovic-Jeremic J, Northup JK, Grisshammer R. Modulation of the interaction between neurotensin receptor NTS1 and Gq protein by lipid. *J. Mol. Biol.* 2012, 417: 95–111.
- Ingólfsson HI, Melo MN, van Eerden FJ, Arnarez C, Lopez CA, Wassenaar TA, Periole X, de Vries AH, Tieleman DP, Marrink SJ. Lipid organization of the plasma membrane. *J. Am. Chem. Soc.* 2014, 136: 14554–14559.
- Ishchenko A, Wacker D, Kapoor M, Zhang A, Han GW, Basu S, Patel N, Messerschmidt M, Weierstall U, Liu W, Katritch V, Roth BL, Stevens RC, Cherezov V. Structural insights into the extracellular recognition of the human serotonin 2B receptor by an antibody. *PNAS.* 2017, 114(31): 8223-8228.
- Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, Ijzerman AP, Stevens RC. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science.* 2008, 322(5905): 1211-7.
- Jin W, Brown S, Roche JP, Hsieh C, Cerver JP, Koo A, Chavkin C, Mackie K. Distinct domains of the CB1 cannabinoid receptor mediate desensitization and internalization. *J. Neurosci.* 1999, 19(10): 3773-80.

- Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: A Web-based Graphical User Interface for CHARMM. *J. Comput. Chem.* 2008, 29:1859–1865.
- Jongejan A, Bruysters M, Ballesteros JA, Haaksma E, Bakker RA, Pardo L, Leurs R. Linking agonist binding to histamine H1 receptor activation. *Nat. Chem. Biol.* 2005, 1(2): 98-103.
- Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 1983, 22(12): 2577–2637.
- Kalatskaya I, Schüssler S, Blaukat A, Müller-Esterl W, Jochum M, Proud D, Faussner A. Mutation of Tyrosine in the Conserved NPxxY Sequence Leads to Constitutive Phosphorylation and Internalization, but Not Signaling of the Human B<sub>2</sub> Bradykinin Receptor. *J. Biol. Chem.* 2004, 279: 31268–31276.
- Kaya AI, Thaker TM, Preininger AM, Iverson TM, Hamm HE. Coupling efficiency of rhodopsin and transducin in bicelles. *Biochemistry.* 2011, 50: 3193–3203.
- Koehl A, Hu H, Maeda S, Zhang Y, Qu Q, Paggi JM, Latorraca NR, Hilger D, Dawson R, Matile H, Schertler GFX, Granier S, Weis WI, Dror RO, Manglik A, Skinotis G, Kobilka BK. Structure of the  $\mu$ -opioid receptor-G<sub>i</sub> protein complex. *Nature.* 2018, 558(7711): 547–552.
- Kohlhoff KJ, Shukla D, Lawrenz M, Bowman GR, Konerding DE, Belov D, Altman RB, Pande VS. Cloud-based simulations on Google Exacycle reveal ligand modulation of GPCR activation pathways. *Nat. Chem.* 2014, 6 (1): 15–21.
- Kolakowski LFJ. GCRDb: a G-protein-coupled receptor database. *Recept. Channels.* 1994, 2(1): 7.
- Kruse AC, Ring AM, Manglik A, Hu J, Hu K, Eitel K, Hübner H, Pardon E, Valant C, Sexton PM, Christopoulos A, Felder CC, Gmeiner P, Steyaert J, Weis WI, Garcia KC, Wess J, Kobilka BK. Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature.* 2013, 504: 101–106.
- Kučerka N, Nieh MP, Katsaras J. Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochim. Biophys. A.* 2011, 1808: 2761–2771.
- Lange OF, Grubmüller H. Generalized correlation for biomolecular dynamics. *Proteins.* 2006, 62: 1053–1061.
- Latorraca NR, Venkatakrishnan AJ, Dror RO. GPCR Dynamics: Structures in Motion. *Chem Rev.* 2017, 117(1): 139-155.

- Lee Y, Warne T, Nehmé R, Pandey S, Dwivedi-Agnihotri H, Chaturvedi M, Edwards PC, García-Nafria J, Leslie AGW, Shukla AK, Tate CG. Molecular basis of  $\beta$ -arrestin coupling to formoterol-bound 1-adrenoceptor. *Nature*. 2020, 583: 862–866.
- Lefkowitz RJ, Shenoy SK. Transduction of receptor signals by beta-arrestins. *Science*. 2005, 308: 512–517.
- Liu R, Nahon D, le Roy B, Lenselink EB, Ijzerman AP. Scanning mutagenesis in a yeast system delineates the role of the NPxxY(x) (5,6) F motif and helix 8 of the adenosine A(2B) receptor in G protein coupling. *Biochem. Pharmacol.* 2015, 95(4): 290-300.
- Liu W, Chun E, Thompson AA, Chubukov P, Xu F, Katritch V, Han GW, Roth CB, Heitman LH, Ijzerman AP, Cherezov V, Stevens RC. Structural basis for allosteric regulation of GPCRs by sodium ions. *Science*. 2012, 337: 232–236.
- Lohse MJ, Andexinger S, Pitcher J, Trukawinski S, Codina J, Faure JP, Caron MG, Lefkowitz RJ. Receptor-specific desensitization with purified proteins. Kinase dependence and receptor specificity of beta-arrestin and arrestin in the beta 2-adrenergic receptor and rhodopsin systems. *J. Biol. Chem.* 1992, 267(12): 8558-64.
- Ma X, Hu Y, Batebi H, Heng J, Xu J, Liu X, Niu X, Li H, Hildebrand PW, Jin C, Kobilka BK. Analysis of  $\beta_2$ -AR- $G_s$  and  $\beta_2$ -AR- $G_i$  complex formation by NMR spectroscopy. *PNAS*. 2020, 117 (37): 23096-23105.
- Mahaut-Smith MP, Martinez-Pinna J, Gurung IS. A role for membrane potential in regulating GPCRs? *Trends Pharmacol. Sci.* 2008, 29: 421–429.
- Manglik A, Kruse AC, Kobilka TS, Thian FS, Mathiesen JM, Sunahara RK, Pardo L, Weis WI, Kobilka BK, Granier S. Crystal structure of the  $\mu$ -opioid receptor bound to a morphinan antagonist. *Nature*. 2012, 485(7398): 321-326.
- Mann A, Illing S, Miess E, Schulz S. Different mechanisms of homologous and heterologous  $\mu$ -opioid receptor phosphorylation. *Br. J. Pharmacol.* 2015, 172(2): 311-316.
- Marino KA, Shang Y, Filizola M. Insights into the function of opioid receptors from molecular dynamics simulations of available crystal structures. *Br. J. Pharmacol.* 2018, 175(14): 2834-2845.
- Mialet-Perez J, Green SA, Miller WE, Liggett SB. A primate-dominant third glycosylation site of the  $\beta_2$ -adrenergic receptor routes receptors to degradation during agonist regulation. *J. Biol. Chem.* 2004, 279(37): 38603-7.

- Moore CA, Milano SK, Benovic JL. Regulation of receptor trafficking by GRKs and arrestins. *Annu. Rev. Physiol.* 2007, 69: 451–482.
- Niemelä PS, Ollila S, Hyvönen MT, Karttunen M, Vattulainen I. Assessing the nature of lipid raft membranes. *PLoS Comp. Biol.* 2007, 3: e34.
- Oddi S, Dainese E, Sandiford S, Fezza F, Lanuti M, Chiurchiù V, Totaro A, Catanzaro G, Barcaroli D, De Laurenzi V, Centonze D, Mukhopadhyay S, Selent J, Howlett AC, Maccarrone M. Effects of palmitoylation of Cys415 in helix 8 of the CB1 cannabinoid receptor on membrane localization and signalling. *Br. J. Pharmacol.* 2012, 165(8): 2635-51.
- O'Dowd BF, Hnatowich M, Caron MG, Lefkowitz RJ, Bouvier M. Palmitoylation of the human  $\beta_2$ -adrenergic receptor. Mutation of Cys341 in the carboxyl tail leads to an uncoupled nonpalmitoylated form of the receptor. *J. Biol. Chem.* 1989, 264(13): 7564-9.
- Oh P, Schnitzer JE. Segregation of heterotrimeric G proteins in cell surface microdomains. Gq binds caveolin to concentrate in caveolae, whereas Gi and Gs target lipid rafts by default. *Mol. Biol. Cell.* 2001, 12: 685–698.
- Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* 2000, 289 (548): 739-745.
- Pert CB, Pasternak G, Snyder SH. Opiate agonists and antagonists discriminated by receptor binding in brain. *Science* 1973, 182: 1359–1361.
- Pike LJ, Han X, Chung KN, Gross RW. Lipid rafts are enriched in arachidonic acid and plasmalogen ethanolamine and their composition is independent of caveolin-1 expression: a quantitative electrospray ionization/mass spectrometric analysis. *Biochemistry.* 2002, 41: 2075–2088.
- Prioleau C, Visiers I, Ebersole BJ, Weinstein H, Sealfon SC. Conserved Helix 7 Tyrosine Acts as a Multistate Conformational Switch in the 5-HT<sub>2C</sub> Receptor. Identification of a Novel “LOCKED-ON” Phenotype and Double Revertant Mutations. *J. Biol. Chem.* 2002, 277: 36577–36584.
- Provasi D, Filizola M. Putative active states of a prototypic g-protein-coupled receptor from biased molecular dynamics. *Biophys. J.* 2010, 98: 2347–2355.
- Rapino C, Castellucci A, Lizzi AR, Sabatucci A, Angelucci CB, Tortolani D, Rossi G, D'Andrea G, Maccarrone M. Modulation of Endocannabinoid-Binding Receptors in Human Neuroblastoma Cells by Tunicamycin. *Molecules.* 2019. 24(7): 1432.

- Rasmussen SG, Choi HJ, Fung JJ, Pardon E, Casarosa P, Chae PS, Devree BT, Rosenbaum DM, Thian FS, Kobilka TS, Schnapp A, Konetzki I, Sunahara RK, Gellman SH, Pautsch A, Steyaert J, Weis WI, Kobilka BK. Structure of a nanobody-stabilized active state of the  $\beta_2$ -adrenoceptor. *Nature*. 2011, 469: 175–180.
- Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VR, Sanishvili R, Fischetti RF, Schertler GF, Weis WI, Kobilka BK. Crystal structure of the human  $\beta_2$ -adrenergic G-protein-coupled receptor. *Nature*. 2007, 450: 383–387.
- Rasmussen SG, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, Thian FS, Chae PS, Pardon E, Calinski D, Mathiesen JM, Shah ST, Lyons JA, Caffrey M, Gellman SH, Steyaert J, Skinotis G, Weis WI, Sunahara RK, Kobilka BK. Crystal structure of the  $\beta_2$ -adrenergic receptor- $G_s$  protein complex. *Nature*. 2011, 477: 549–555.
- Ring AM, Manglik A, Kruse AC, Enos MD, Weis WI, Garcia KC, Kobilka BK. Adrenaline-activated structure of  $\beta_2$ -adrenoceptor stabilized by an engineered nanobody. *Nature*. 2013, 502: 575–579.
- Rosenbaum DM, Zhang C, Lyons JA, Holl R, Aragao D, Arlow DH, Rasmussen SG, Choi HJ, Devree BT, Sunahara RK, Chae PS, Gellman SH, Dror RO, Shaw DE, Weis WI, Caffrey M, Gmeiner P, Kobilka BK. Structure and function of an irreversible agonist- $\beta_2$ -adrenoceptor complex. *Nature*. 2011, 469: 236–240.
- Sansom MSP, Weinstein H. Hinges, swivels and switches: the role of prolines in signaling via transmembrane alpha helices. *TIPS* 2000, 21:445–451.
- Scheer A, Fanelli F, Costa T, Benedetti PGD, Cotecchia S. Constitutively active mutants of the  $\alpha_1B$ -adrenergic receptor: role of highly conserved polar amino acids in receptor activation. *EMBO J*. 1996, 15: 3566–3578.
- Schiöth HB, Fredriksson R. The GRAFS classification system of G-protein coupled receptors in comparative perspective. *General and Comparative Endocrinology*. 2004, 142(1-2): 94–101.
- Schneider J, Korshunova K, Musiani F, Alfonso-Prieto M, Giorgetti A, Carloni P. Predicting ligand binding poses for low-resolution membrane protein models: Perspectives from multiscale simulations. *Biochem. Biophys. Res. Commun*. 2018, 498(2): 366–374.
- Sealfon SC, Chi L, Ebersole BJ, Rodic V, Zhang D, Ballesteros JA, Weinstein H. Related Contribution of Specific Helix 2 and 7 Residues to Conformational Activation of the Serotonin 5-HT<sub>2A</sub> Receptor. *J. Biol. Chem*. 1995, 270: 16683–16688.

- Seeber M, Benedetti PGD, Fanelli F. Molecular dynamics simulations of the ligand induced chemical information transfer in the 5-HT<sub>1A</sub> receptor. *J. Chem. Inf. Comp. Sci.* 2003, 43: 1520–1531.
- Selent J, Sanz F, Pastor M, De Fabritiis G. Induced Effects of Sodium Ions on Dopaminergic G-Protein Coupled Receptors. *PLoS Comput. Biol.* 2010, 6: e1000884.
- Shang Y, LeRouzic V, Schneider S, Bisignano P, Pasternak GW, Filizola M. Mechanistic Insights into the Allosteric Modulation of Opioid Receptors by Sodium Ions. *Biochemistry.* 2014, 53 (31): 5140-5149.
- Shao Z, Yin J, Chapman K, Grzemska M, Clark L, Wang J, Rosenbaum DM. High-resolution crystal structure of the human CB1 cannabinoid receptor. *Nature.* 2016, 540: 602–606.
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 2011, 7: 539.
- Song C, Howlett AC. Rat brain cannabinoid receptors are N-linked glycosylated proteins. *Life Sci.* 1995, 56(23-24): 1983-9.
- Sounier R, Mas C, Steyaert J, Laeremans T, Manglik A, Huang W, Kobilka BK, D  m  n   H, Granier S. Propagation of conformational changes during  $\mu$ -opioid receptor activation. *Nature.* 2015, 524(7565): 375-8.
- Strahs D, Weinstein H. Comparative modeling and molecular dynamics studies of the delta, kappa and mu opioid receptors. *Protein Eng.* 1997, 10(9): 1019–1038.
- Strathmann MP, Simon MI. G alpha 12 and G alpha 13 subunits define a fourth class of G protein alpha subunits. *PNAS.* 1991, 88 (13): 5582–6.
- Strohman MJ, Maeda S, Hilger D, Masureel M, Du Y, Kobilka BK. Local membrane charge regulates  $\beta_2$ -adrenergic receptor coupling to G<sub>i3</sub>. *Nat. Commun.* 2019, 10: 2234.
- Sugiura T, Kondo S, Sukagawa A, Nakane S, Shinoda A, Itoh K, Yamashita A, Waku K. 2-Arachidonoylglycerol: a possible endogenous cannabinoid receptor ligand in brain. *Biochem. Biophys. Res. Commun.* 1995, 215(1): 89-97.
- Venkatakrishnan AJ, Ma AK, Fonseca R, Latorraca NR, Kelly B, Betz RM, Asawa C, Kobilka BK, Dror RO. Diverse GPCRs exhibit conserved water networks for stabilization and activation. *PNAS.* 2019, 116(8): 3288-3293.

- Verkleij AJ, Post JA. Membrane phospholipid asymmetry and signal transduction. *J. Membr. Biol.* 2000, 178: 1–10.
- Vukoti K, Kimura T, Macke L, Gawrisch K, Yeliseev A. Stabilization of functional recombinant cannabinoid receptor CB2 in detergent micelles and lipid bilayers. *PLoSOne*. 2012, 7: e46290.
- Wacker D, Fenalti G, Brown MA, Katritch V, Abagyan R, Cherezov V, Stevens RC. Conserved binding mode of human beta2 adrenergic receptor inverse agonists and antagonist revealed by X-ray crystallography. *J. Am. Chem. Soc.* 2010, 132(33): 11443-5.
- Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. 2009, 25: 1189–1191.
- Webb B, Sali A. Comparative Protein Structure Modeling Using Modeller. *Curr. Protoc. Bioinformatics*. 2016, 54: 5.6.1-5.6.37.
- Xiang Y. Caveolar localization dictates physiologic signaling of beta 2-adrenoceptors in neonatal cardiac myocytes. *J. Biol. Chem.* 2002, 277: 34280–34286.
- Xu W, Ozdener F, Li JG, Chen C, de Riel JK, Weinstein H, Liu-Chen LY. Functional Role of the Spatial Proximity of Asp114<sup>2.50</sup> in TMH 2 and Asn332<sup>7.49</sup> in TMH 7 of the Mu Opioid Receptor. *FEBS Lett.* 1999, 447: 318–324.
- Yao X, Parnot C, Deupi X, Ratnala VR, Swaminath G, Farrens D, Kobilka B. Coupling ligand structure to specific conformational switches in the beta2-adrenoceptor. *Nat. Chem. Biol.* 2006, 2(8): 417-22.
- Yen HY, Hoi KK, Liko I, Hedger G, Horrell MR, Song W, Wu D, Heine P, Warne T, Lee Y, Carpenter B, Plückthun A, Tate CG, Sansom MSP, Robinson CV. PtdIns (4,5) P2 stabilizes active states of GPCRs and enhances selectivity of G-protein coupling. *Nature*. 2018, 559: 423–427.
- Yin W, Li Z, Jin M, Yin YL, de Waal PW, Pal K, Yin Y, Gao X, He Y, Gao J, Wang X, Zhang Y, Zhou H, Melcher K, Jiang Y, Cong Y, Edward Zhou X, Yu X, Eric Xu H. A complex structure of  $\beta$ -arrestin-2 bound to a G protein-coupled receptor. *Cell Res.* 2019, 29: 971–983.
- Yuan S, Vogel H, Filipek S. The Role of Water and Sodium Ions in the Activation of the m-Opioid Receptor *Angew. Chem. Int. Ed. Engl.* 2013, 52 (38): 10112-10115.
- Zadina JE, Hackler L, Ge LJ, Kastin AJ. A potent and selective endogenous agonist for the mu-opiate receptor. *Nature*. 1997, 386(6624): 499-502.



- Zamah AM, Delahunty M, Luttrell LM, Lefkowitz RJ. Protein kinase A-mediated phosphorylation of the  $\beta_2$ -adrenergic receptor regulates its coupling to  $G_s$  and  $G_i$ . demonstration in a reconstituted system. *J. Biol. Chem.* 2002, 277(34): 31249-56.
- Zhan X, Gimenez LE, Gurevich VV, Spiller BW. Crystal Structure of Arrestin-3 Reveals the Basis of the Difference in Receptor Binding Between Two Non-Visual Subtypes. *J. Mol. Biol.* 2011, 406: 467–478.
- Zhang C, Srinivasan Y, Arlow DH, Fung JJ, Palmer D, Zheng Y, Green HF, Pandey A, Dror RO, Shaw DE, Weis WI, Coughlin SR, Kobilka BK. High-resolution crystal structure of human protease-activated receptor 1. *Nature.* 2012, 492: 387–392.
- Zheng H, Pearsall EA, Hurst DP, Zhang Y, Chu J, Zhou Y, Reggio PH, Loh HH, Law PY. Palmitoylation and membrane cholesterol stabilize  $\mu$ -opioid receptor homodimerization and G protein coupling. *BMC Cell Biol.* 2012, 13(6): 1-18.
- Zhou Q, Yang D, Wu M, Guo Y, Guo W, Zhong L, Cai X, Dai A, Jang W, Shakhnovich EI, Liu ZJ, Stevens RC, Lambert NA, Babu MM, Wang MW, Zhao S. Common activation mechanism of class A GPCRs. *eLife.* 2019, 8: e50279.
- Zia SR, Gaspari R, Decherchi S, Rocchia W. Probing hydration patterns in class-A GPCRs via biased MD: the  $A_{2A}$ -receptor. *J. Chem. Theor. Comput.* 2016, 12: 6049–6061.

# I

**Mitra, A** and Sarkar, A, Szabó, MR and Borics, A.

*Correlated motions of conserved polar motifs lay out a plausible mechanism of G Protein-Coupled Receptor activation.*





BIOMOLECULES, 11 (5), 670 (2021).

doi: 10.3390/biom11050670.

IF: 4.879 [Q2]

## Article

# Correlated Motions of Conserved Polar Motifs Lay out a Plausible Mechanism of G Protein-Coupled Receptor Activation

Argha Mitra <sup>1</sup>, Arijit Sarkar <sup>1</sup>, Márton Richárd Szabó <sup>1,2</sup> and Attila Borics <sup>1,\*</sup>

<sup>1</sup> Laboratory of Chemical Biology, Institute of Biochemistry, Biological Research Centre, Szeged, 62. Temesvári krt., H-6726 Szeged, Hungary; argha.mitra@brc.hu (A.M.); sarkararajit@brc.hu (A.S.); szabo.marton@med.u-szeged.hu (M.R.S.)

<sup>2</sup> Department of Biochemistry, Faculty of Medicine, University of Szeged, 9 Dóm sq., H-6720 Szeged, Hungary

\* Correspondence: borics.attila@brc.hu; Tel.: +36-62-599-600 (ext. 430)



**Citation:** Mitra, A.; Sarkar, A.; Szabó, M.R.; Borics, A. Correlated Motions of Conserved Polar Motifs Lay out a Plausible Mechanism of G Protein-Coupled Receptor Activation. *Biomolecules* **2021**, *11*, 670. <https://doi.org/10.3390/biom11050670>

Academic Editor: Terry Hébert

Received: 23 March 2021

Accepted: 28 April 2021

Published: 30 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Recent advancements in the field of experimental structural biology have provided high-resolution structures of active and inactive state G protein-coupled receptors (GPCRs), a highly important pharmaceutical target family, but the process of transition between these states is poorly understood. According to the current theory, GPCRs exist in structurally distinct, dynamically interconverting functional states of which populations are shifted upon binding of ligands and intracellular signaling proteins. However, explanation of the activation mechanism, on an entirely structural basis, gets complicated when multiple activation pathways and active receptor states are considered. Our unbiased, atomistic molecular dynamics simulations of the  $\mu$  opioid receptor (MOP) revealed that transmission of external stimulus to the intracellular surface of the receptor is accompanied by subtle, concerted movements of highly conserved polar amino acid side chains along the 7th transmembrane helix. This may entail the rearrangement of polar species and the shift of macroscopic polarization in the transmembrane domain, triggered by agonist binding. Based on our observations and numerous independent indications, we suggest amending the widely accepted theory that the initiation event of GPCR activation is the shift of macroscopic polarization between the ortho- and allosteric binding pockets and the intracellular G protein-binding interface.

**Keywords:** GPCR; opioid; activation mechanism; signal transduction; molecular dynamics

## 1. Introduction

G protein-coupled receptors (GPCRs) are located on cell surfaces and act as communication interfaces for external stimuli exerted by structurally diverse molecules. Upon activation, GPCRs initiate signal transduction through interactions with G proteins and arrestins, and control a variety of intracellular processes. Owing to this, approximately 34% of all prescription pharmaceuticals target members of this receptor family [1]. However, application of such drugs is often limited by a number of unwanted side effects due to non-selective activation of multiple GPCRs, or multiple signaling pathways associated with one receptor. The most recent challenge of rational drug design is, therefore, to develop signaling pathway-specific, or in other words “functionally selective” GPCR agonists. To address this challenge, complete understanding of the structural mechanism of GPCR activation is necessary. Opposed to the high diversity of external activators, signaling is mediated by only a few types of G proteins, advocating that GPCR activation may follow a general mechanism.

The structure of GPCRs consists of a conserved bundle of seven transmembrane (TM)  $\alpha$ -helices, and highly dynamic extracellular and cytosolic domains of various lengths. High-resolution experimental structures are available for many GPCRs both in the active and inactive states (for a comprehensive collection visit <http://gpcrdb.org>, accessed on

28 April 2021) [2], but the mechanism of transition between these forms is intensely debated. The most conspicuous difference between the active and inactive class A GPCR structures, published to this date, is a notable disposition of the 6th transmembrane helix (TM6) [3,4]. However, such large dispositions were shown to occur even in the absence of a bound ligand, due to the inherent dynamics of the receptor structure, or may originate from the applied conditions of crystallographic structure determination, namely the attachment of fusion proteins or the application of crystallization chaperones [5–7]. Apart from TM6 disposition, a possible role of intracellular loop 1 (ICL1) and the cytosolic helix (H8) in the activation mechanism was highlighted by dynamic NMR measurements of the  $\mu$ -opioid receptor (MOP) [8]. Comparison of the structures of active and inactive state MOP and  $\delta$ -opioid (DOP) receptors suggested that an extended network of polar amino acids and water molecules connects the orthosteric ligand binding pocket to the cytosolic domains, which may be functionally relevant. The highly conserved polar functional motifs, E/DRY, NPxxY, and CWxP, have been specified to participate in the activation mechanism of class A GPCRs and the known effect of elevated concentrations of  $\text{Na}^+$  that prevents the agonist-induced activation of opioid receptors [9] and related GPCRs was attributed to a conserved allosteric  $\text{Na}^+$  binding site [10–14]. Conceivably, activation signal is transmitted to the intracellular surface of the receptor through the interplay of these polar microswitches, however, no direct evidence of such integral mechanism has yet been given. Real-time observation of such processes using conventional experimental techniques is unattainable.

A significant part of the now widely accepted theory of GPCR activation was provided by landmark molecular dynamics (MD) simulation studies [3,4,13,14]. According to this theory, GPCRs exist as a dynamic ensemble of multiple active, inactive, and intermediate states. The populations of active states are increased by agonist binding and the stabilization of an active structure facilitates the insertion of G proteins [13]. The growing amount of evidence of pre-coupled GPCR-G protein complexes in the absence of ligands, however, presents a challenge to the above hypothesis [15]. In general, explanations given on an entirely structural basis are often diffuse and fail to provide unequivocal suggestion for a possible structural mechanism of GPCR activation, especially when multiple active states, or structurally similar but functionally different ligands are considered. Further limitations of previous MD studies are that simulation systems were confined to the TM region of GPCRs, embedded in very simplistic representations of the cell membrane. Most recently, special effects of charged interfacial lipids on  $\beta_2$ -adrenergic receptor signaling was demonstrated, drawing attention to the importance of accurate membrane representation [16].

The MOP is one of the most extensively studied GPCRs, therefore it appropriately represents the general structural features of class A (rhodopsin-like) GPCRs. In order to get a deeper insight, we have performed all-atom MD simulations of the full sequence MOP, including the N- and C-terminal domains, on a  $\mu\text{s}$  timescale. To better approximate the physiological conditions of the activation mechanism, simulations of the active and inactive receptors were executed in caveolar membrane environment [17], in the presence of the endogenous agonist endomorphin-2 (EM2, H-Tyr-Pro-Phe-Phe-NH<sub>2</sub>) [18] and the G<sub>i</sub> protein complex, or beta-arrestin-2. In addition, reference simulations were carried out in the presence of allosterically bound  $\text{Na}^+$  or in the absence of EM2. Further control simulations were performed with fused T4-lysozyme or intracellularly bound Nb39 nanobody, representing the previously applied crystallization conditions [5,7]. Our simulations were intended to gather information about how the N- and C-terminal domains, intracellular proteins, and crystallization chaperones affect the internal dynamics of the TM domain and consequently the activation mechanism. The analysis of trajectories was aimed at confirming the integrity of the simulation systems, comparing structural properties of our systems to previously published simulation data, as well as introducing new perspectives.

## 2. Methods

### 2.1. System Building

The crystallographic structures used in this study were downloaded from the Brookhaven Protein Data Bank (<http://www.rcsb.org>, accessed on 28 April 2021): active MOP (pdb code: 5C1M), inactive MOP (pdb code: 4DKL), heterotrimeric G<sub>i</sub> protein complex (pdb code: 1GP2), Nb39 nanobody (pdb code: 5C1M), T4-lysozyme (pdb code: 4DKL),  $\beta_2$ -adrenergic receptor complexed with G<sub>s</sub> protein (pdb code: 3SN6), and rhodopsin complexed with beta-arrestin-2 (pdb code: 4ZWJ). These latter two structures were used as templates to orient the G<sub>i</sub> protein complex and beta-arrestin-2 to the active and inactive MOP. The full sequence of the murine MOP (UniProtKB-P42866-OPRM1) was obtained from UniProt (<http://www.uniprot.org>, accessed on 28 April 2021), and the coordinates of the membrane orientation from the OPM server (<http://opm.phar.umich.edu>, accessed on 28 April 2021). The crystallization chaperone and fusion protein (Nb39 nanobody and T4-lysozyme, respectively) were removed from the crystallographic structures. The Swiss-PdbViewer was used to retrieve all missing, modified, or mutated residues of the transmembrane (TM) domain of the receptor. GTP was generated in CHARMM-GUI [19] and edited manually to replace GDP in the G<sub>i</sub> complex.

To model the missing N- and C-terminal domains, 10 ns folding simulations of the N- and C-terminal domains were performed using the GROMACS ver. 5.1.4 program package [20], the AMBER ff99SB-ILDN-NMR [21] force field, and the GB/SA implicit solvation model [22]. During MD simulations, the system temperature was set to 310 K and maintained by the v-rescale algorithm [23]. Ten parallel simulations were run for both the N- and C-terminal domains from where the resultant, folded structures were evaluated and selected based on the compactness, accessibility of post-translational modification, and TM region attachment sites. (For further details, see the “MD trajectory analysis” subsection below.) Glycosylation sites were predicted using the NetNGlyc 1.0 online server [24]. The selected N- and C-terminal domain structures were linked to the TM region using Pymol ver. 2.1.0. Four intracellular partners were used in this study, namely the heterotrimeric G<sub>i</sub> protein, beta-arrestin-2, Nb39 nanobody, and T4-lysozyme. Among them the last two were used for reference simulation systems. The first three proteins were attached non-covalently to the receptor, while T4-lysozyme was fused with the receptor replacing the third intracellular loop (ICL3), similar to that in the crystallographic structure of the inactive MOP (pdb code: 4DKL). Cryo-electron microscopic structure of the G<sub>i</sub> protein-bound MOP [6], published later, have verified the adequacy of the corresponding model built in this study (Figure S1).

CHARMM-GUI was used to include various post-translational modifications, as well as to build membrane bilayers. Complex type glycans were added to the N-terminal domain, containing a common core (Man $\alpha$ 1-3 (Man $\alpha$ 1-6) Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc $\beta$ 1-N) and sialic acid (N-acetylneuraminic acid) at glycosylation prone N9, N31, and N38 residues of the N-terminal domain [25]. Phosphorylation of S363 and T370 were done for all the complexes, while S375, T376, and T379 sites at the C-terminal domain were phosphorylated in addition for the arrestin complexes [26]. The C170 residue of ICL2 was palmitoylated [27].

A caveolar membrane environment, considered to be the physiological environment of the MOP, was built using the membrane builder tool of CHARMM-GUI. CHARMM36 parameters were used to build complex, multicomponent membrane systems, which included cholesterol (CHL-32.8%), 1-palmitoyl-2-oleoyl-glycero-3-phospho-choline (POPC-14.9%), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE-27.8%), 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine (POPS-3.6%), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoinositol (POPI2-6%), palmitoyl-sphingomyelin (PSM-9.9%), and monosialodihexosylganglioside (GM3-5%) [17]. GM3 gangliosides were generated separately using the glycoprotein builder tool of CHARMM-GUI, and then added manually to the membrane. The asymmetric upper and the lower leaflet membrane compositions were specified in a most probable ratio [28]. The membrane builder was also used to embed the glycosylated, palmitoylated,



and phosphorylated full sequence model of the MOP into the membrane. Systems were then solvated explicitly with TIP3P water molecules in a hexagonal shaped periodic box, and sodium and chloride ions (0.15 M) were added to neutralize the net charge and to attain physiological ionic strength. System coordinates and topologies were generated in GROMACS format.

EM2 [18], a peptide agonist of the MOP was built manually, using Pymol ver. 2.1.0. The binding site was confirmed by flexible docking of this ligand to the active state MOP crystallographic structure (pdb code: 5C1M), using the Autodock ver. 4.2 software [29] and the Lamarckian genetic algorithm. All  $\phi$ ,  $\psi$ , and  $\chi^1$  ligand torsions, as well as receptor side chains in contact with the bound ligand [5] were kept flexible. Docking of EM2 was performed in an 8.0 nm  $\times$  8.0 nm  $\times$  8.0 nm grid volume, large enough to cover the whole binding pocket of the receptor region accessible from the extracellular side. The spacing of grid points was set at 0.0375 nm and 1000 dockings were done. The resultant ligand-receptor complexes were clustered and ranked according to the corresponding binding free energies. The lowest energy bound state was selected for simulations, in which specific ligand-receptor interactions observed in the crystallographic structures were present. Cryo-electron microscopic structure of the MOP and the peptide agonist DAMGO, published later, have confirmed the correct localization and analogous orientation of pharmacophores of EM2 [6]. Two additional control simulation systems were built for the active state,  $G_i$  protein-bound receptor, either with the exclusion of bound EM2, or with the inclusion of EM2 in the orthosteric site together with a  $Na^+$  ion placed in the proximity of D114<sup>2,50</sup> (Ballesteros-Weinstein numbering is indicated in the upper index) in the allosteric binding pocket [10–14].

## 2.2. MD Simulations

All equilibration and production MD simulations were performed using the GROMACS ver. 5.1.4. molecular dynamics program package. Ten independent simulations were performed, four each for inactive and active MOP, complexed with heterotrimeric  $G_i$  protein, beta-arrestin-2, Nb39 nanobody, and T4-lysozyme. Additionally, reference simulations were run for the ligand-free, active state,  $G_i$  protein-bound receptor, and for the EM2-bound active receptor- $G_i$  protein complex in the presence of an allosteric  $Na^+$  ion. After orienting and adding EM2, where applicable, the resultant complex systems were energy minimized thoroughly performing 5000 steps steepest descent, followed by 5000 steps conjugate gradient minimization with convergence criteria of 1000 kJ/mol nm<sup>-1</sup> in both cases. After minimization, systems were subjected to a six-step equilibration protocol, supplied by CHARMM-GUI. According to this protocol, positionally restrained MD simulations were executed in the canonical (NVT) and then, after 2 steps, in the isobaric-isothermal (NPT) ensemble at 303.15 K and 1 bar, having the positional restraints on the heavy atoms of the proteins and membrane constituents decreasing gradually. The first three equilibration MD runs were done for 25 ps in 1 fs time steps, and the following two were continued for 100 ps in 2 fs time steps. The sixth step of the equilibration protocol was run for 50 ns in 2 fs time steps. The following, further parameters were applied: the LINCS algorithm was used to constrain all bonds to their correct length, temperature was regulated by the v-rescale [23] algorithm with a coupling constant of 1 ps, semi-isotropic Berendsen pressure coupling [30] was applied with a coupling constant of 5 ps and compressibility of  $4.5 \times 10^{-5}$  bar<sup>-1</sup>. The Particle Mesh Ewald (PME) method was used and a twin-range cutoff was applied to calculate energy contributions from electrostatic and van der Waals interactions, respectively. All cut-off values were set to 1.2 nm. After equilibration, production simulations were performed for 1  $\mu$ s at 310 K in the NPT ensemble, with other parameters same as above. The coordinates were stored in every 5000th steps yielding trajectories of 100.000 snapshots.

### 2.3. MD Trajectory Analysis

Analysis of MD trajectories was performed using the analysis suite of the GROMACS 5.1.4 package. Generally, the analyses of MD trajectories were performed to evaluate membrane properties and protein conformational changes, stability of the molecular complexes, as well as to investigate previously described interactions and their role in different activation states of the receptor.

Folding simulations were assessed through the analysis of root mean square deviation (RMSD) of backbone atom positions with respect to the starting structures. Furthermore, the radius of gyration of the N- and C-terminal domains were measured, and the number of intramolecular H-bonds were calculated along the trajectories using the gmx gyrate and gmx hbond utilities, respectively. The evolution of secondary structure was monitored using the DSSP (Define Secondary Structure of Proteins) method [31]. Membrane thickness and area per lipid head-group values were calculated using the FATSlim 0.2.1 program [32].

For each production MD trajectory RMSD, calculations were carried out to assess the structural stability of the complex and demonstrate significant displacements of structural components as a function of time. RMSD values of protein backbone atoms were calculated in comparison with the active and inactive state starting structures. The dynamics of terminal domains during simulations was examined by monitoring the radii of gyration. Conformational fluctuations of specific amino acid side chains were analyzed by measuring side chain  $\chi^1$  angles and calculating the frequency of transitions between rotameric states using gmx chi. The gmx helix utility was used to calculate helix properties. Secondary structure assignment was, again, done using the DSSP method.

The occurrence and frequency of intra- and intermolecular H-bonds were calculated using the gmx hbond utility. The donor-acceptor distance and donor-hydrogen-acceptor angle cut-offs for H-bond assignments were set to 0.35 nm and 30.0 degrees, respectively. The presence of salt bridges was monitored by measuring distance and angle between the corresponding acidic and basic side chain functional groups, using gmx distance and gmx angle, respectively. The distance threshold for salt bridge assignment was 0.4 nm and the angle threshold was 90.0 degrees. Where applicable, penetration of Na<sup>+</sup> ions into the allosteric Na<sup>+</sup> binding site, D114<sup>2.50</sup>, was checked using the gmx mindist utility.

The extent of correlation of atomic displacements was examined by dynamic cross-correlation matrix analysis (DCCM) integrated into an earlier version of the GROMACS suite (g\_correlation, ver. 3.3) [33]. The GIMP ver. 2.8 software was used for image analysis of the obtained DCCM maps, where the extent of correlation was demonstrated by color intensity. The threshold of assignment of correlation was red color intensity corresponding to >0.7 MI (mutual information). Amino acid side chains having at least 4 atoms participating in correlated motions were considered. The threshold of 4 atoms have been set in order to exclude irrelevant sidechain motions, such as torsional rotations of methyl groups.

Systems were visualized using Pymol ver. 2.1.0 or VMD ver. 1.9.3. software [34] and graphs were prepared using the Xmgrace ver. 5.1.25 program.

### 2.4. Sequence Alignment and Conservation Analysis

244 sequences of class A mouse GPCRs (without orphan and olfactory receptors) were retrieved from the UniProt database in FASTA format. The Clustal Omega program [35] was used to align those multiple sequences and the results were analyzed using Jalview ver. 2.10.5. [36] The OPRM\_MOUSE (P42866) sequence was set as reference. The sequences were compared based on percentage of identity.

## 3. Results and Discussion

### 3.1. System Building

An initial challenge in this study was to build simulation systems which approximate the physiological conditions of the MOP as closely as possible. The N- and C-terminal domains of the receptor were included to account for the drag posed by the mass of these

domains and its effect on the dynamics of transmembrane helices, of which the central role in the activation mechanism was widely emphasized by previous proposals [3–8,12–14]. In the absence of atomic resolution structures of these or homologous domains, folding simulations were performed to create approximate structures. Although the evolution of backbone RMSD, radius of gyration, number of H-bonds, and secondary structure indicated convergence of folding simulations in most cases (Figure S2), it is ambiguous that the folding of N- and C-terminal domains were correct and complete in the given time frame. Nevertheless, the aforementioned primary purpose of the inclusion of these domains was sufficiently fulfilled by the generated structures. Partial unfolding of the N- and C-terminal domains was observed during some of the production simulations (Figure S3), but the size of the periodic box, varying between 17.64 nm and 24.00 nm in the Z dimension was large enough to keep periodic images of these domains from contacting each other and to produce artifacts (Figure S4).

### 3.2. Membrane Properties

Immediate vertical contraction of the membrane was observed in the equilibration phase, approaching the thickness of homogeneous phospholipid bilayers [37], whereas lateral contraction was more gradual (Figure S5). Then, after the removal of all restraints, a slight vertical expansion of the membrane was observed in the initial production phase (first 20–30 ns), then the membrane stabilized at approximately 4.30 nm thickness, a value more appropriate for lipid rafts [38]. Lateral contraction of the membrane continued in the initial production phase until stabilization at approximately 0.49 nm<sup>2</sup> after 30 ns. It is important to note, that the initial imbalance of membrane parameters in the production phase emerge from that positional restraints on the protein were released only at the beginning of the production phase. Since simulations were intended to monitor the process of activation, of which timescale could not be estimated, the possibility of structural changes in the equilibration phase had to be kept minimal. Therefore, positional restraints on the heavy atoms of the protein were decreased step-by-step, but maintained throughout the equilibration. No artificial trends reflecting this initial membrane imbalance were perceived in the results. The average values for bilayer thickness and area per lipid headgroup, having excluded the values of the first 40 ns, were 4.328 nm and 0.493 nm<sup>2</sup>, respectively.

### 3.3. Allosteric Na<sup>+</sup> Binding

Penetration of Na<sup>+</sup> ions into the allosteric Na<sup>+</sup> binding site (D114<sup>2.50</sup>) did not happen during the simulations of EM2-bound receptors (Figure S6), regardless of the receptor state. While H<sub>2</sub>O molecules were observed to exchange between the internal cavities of the transmembrane domain as well as the bulk solvent phase, Na<sup>+</sup> ions did not enter the TM domain, neither from the extracellular nor the intracellular side. This suggests that for Na<sup>+</sup> ions the allosteric site is only accessible through the orthosteric binding pocket, and its entrance could be blocked by a bound ligand, whereas intracellular access to the TM domain is closed by the G<sub>i</sub> protein. Conversely, Na<sup>+</sup> quickly localized at the allosteric site when EM2 was not present, following the transition of the receptor to an intermediate structural state. Interestingly, Na<sup>+</sup> binding to the orthosteric anchor residue, D147<sup>3.32</sup>, was significantly less frequent, rather occasional in the ligand-free receptor. Moreover, binding of Na<sup>+</sup> to the orthosteric site and dissociation of Na<sup>+</sup> from the allosteric site was often coincidental (Figure S7a). The very low frequency of simultaneous occupation of ortho- and allosteric sites, however, may be simply the consequence of the applied Na<sup>+</sup> concentration and more frequent simultaneous occupation could be achieved at higher ionic strengths. The observed spontaneous Na<sup>+</sup> penetration from the extracellular environment is in agreement with previously published MD simulation data [39,40], but this and the occupancy data presented above is insufficient to provide explanation for the modulation of receptor activation. The localization of Na<sup>+</sup> ions in the transmembrane region and their effect on the activation of class A GPCRs have been extensively studied previously, both by experimental and theoretical methods [10–12,39–41]. Translocation of

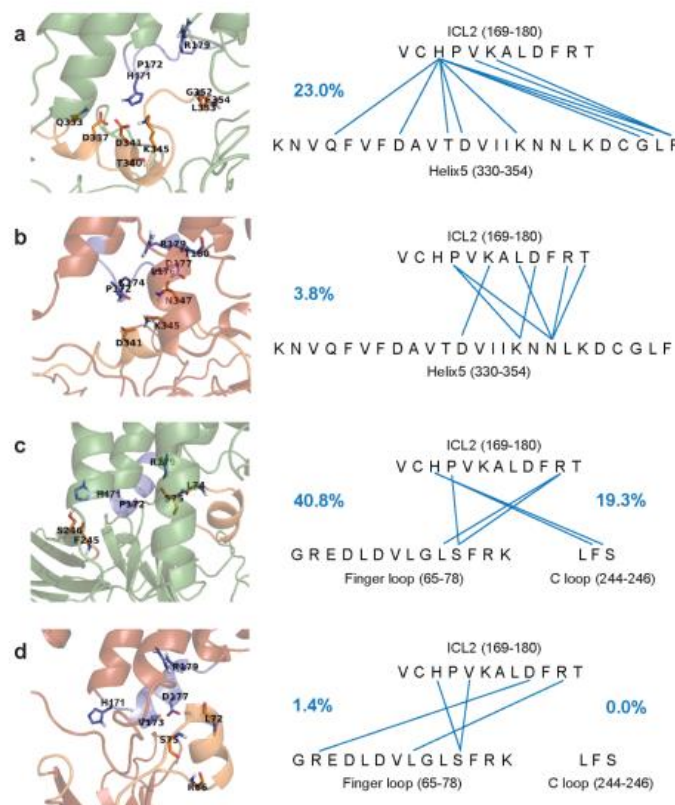


Na<sup>+</sup> ions through the active state MOP was observed in previous MD simulations, but in the absence of bound ligands and intracellular proteins [41]. According to the current state-of-the-art receptor activation is accompanied by the “collapse” of the allosteric Na<sup>+</sup> binding site, which results in the ejection of the bound Na<sup>+</sup> and its migration towards the cytosol. [41]. Our simulation of the active MOP-G protein complex in the presence of EM2 at the orthosteric binding site, and Na<sup>+</sup> placed initially at the allosteric Na<sup>+</sup> binding site, have provided corroborating results. Compared to the ligand-free system, the localization of the Na<sup>+</sup> ion in the allosteric site and its contacts with D114<sup>2.50</sup>, S154<sup>3.39</sup>, and W293<sup>6.48</sup> have become loose during the simulation. Consequently, the frequency of close contacts with conserved polar/amphipathic residues further down towards the intracellular surface (N328<sup>7.45</sup>, N332<sup>7.49</sup>, and Y336<sup>7.53</sup>) have increased significantly (Figure S7b).

### 3.4. TM Helix and Loop Dynamics

Atomic displacement analysis of transmembrane helical backbones of the EM2-bound receptors indicated, that TM6 assumed intermediate conformations during simulations with minor changes from the corresponding starting structures (Figure S8). This is in line with previous simulation results, where notable TM helix rearrangements were only observed at longer timescales and in the absence of bound intracellular proteins [13].

The largest disposition was measured for the inactive receptor beta-arrestin-2 complex, suggesting a preference of beta-arrestin-2 for the active structural state of the receptor. TM6 of the ligand-free receptor, on the other hand, underwent much larger changes. This demonstrates remarkable stabilizing effect of the agonist, regardless of activation state. Initial conformations of ICL1 and H8 were maintained throughout simulations in each receptor state, regardless of the intracellular interacting partners (Figures S9 and S10). The second intracellular loop (ICL2), on the other hand, adopted a stable  $\alpha$ -helical structure when bound by beta-arrestin-2 and partially unfolded upon interaction with the G<sub>i</sub> subunit, independent of the state of the receptor (Figure S11). This latter, however, was only observed in the presence of EM2 in the orthosteric site and the absence of Na<sup>+</sup> in the allosteric site. Apparently, this structural transition of ICL2 was prevented by the bound Na<sup>+</sup> ion. In the active states, increased frequency of intermolecular hydrogen bonds was observed involving ICL2, helix 5 of G<sub>i $\alpha$</sub> , and the finger- and C-loops of beta-arrestin-2 (Figure 1, Table S1). These observations indicate that the conformations of ICL1 and H8 is controlled by the receptor state whereas, in the presence of an agonist, ICL2 adapts its structure to the bound signaling proteins. Therefore, ICL2 may be partially responsible for signaling pathway specificity. Such dynamics of ICL2 was not indicated by the published high resolution structures of this receptor [5–7]. Supporting evidence was, however, provided by a most recent NMR spectroscopic study of the  $\beta_2$ -adrenergic receptor. Distinct conformations of ICL2 were indicated when the receptor was bound either by the G<sub>s</sub> or the G<sub>i1</sub> protein complex and, similar to the results presented here, G<sub>i1</sub> did not promote the formation of an alpha helix in ICL2 [42]. Peculiar results were obtained for the active, ligand-free, G<sub>i</sub> protein-bound receptor. ICL2 have folded into an  $\alpha$ -helix during simulation, but maintained strong contact with the G<sub>i $\alpha$</sub>  subunit (Figures S11 and S12). Although H-bonds between the G<sub>i $\alpha$</sub>  subunit and the receptor involved only one participant residue of ICL2, the frequency of this H-bond indicated stronger and more specific interaction. This observation supports the hypothesis of pre-coupled GPCR-G protein complexes in the absence of ligands [15], and suggests that the lower frequency and specificity of intermolecular H-bonds in the EM2-bound active receptor may represent an intermediate complex state, which precedes G<sub>i</sub> protein dissociation during the signaling event. Secondary structure analysis of the control systems with Nb39 or T4-lysozyme fusion suggests that these systems better represent the arrestin-bound state of the receptor (Figure S11).



**Figure 1.** Frequency and donor and acceptor sites of intermolecular H-bonds between ICL2 of the MOP and the  $G_i$  protein or beta-arrestin-2. Frequency of H-bonds are expressed as percentages of the total structural ensemble and indicated by blue numbers. ICL2 is shown as blue cartoon and sticks. Helix 5 of the  $G_{i\alpha}$  subunit and the finger- and C loops of beta-arrestin-2 are shown as orange cartoon and sticks. (a) Active receptor and the  $G_{i\alpha}$  subunit. (b) Inactive receptor and the  $G_{i\alpha}$  subunit. (c) Active receptor and beta-arrestin-2. (d) Inactive receptor and beta-arrestin-2.

### 3.5. Specific Intramolecular Interactions

Analysis of intramolecular salt bridges and H-bonds between conserved motifs (Table 1) indicated, in agreement with previous proposals [7], that interactions between D164<sup>3,49</sup> and R165<sup>3,50</sup> of the DRY motif were more frequent in the inactive states and in the ligand-free receptor. The specific role of the previously reported DRY-TM5 [7], DRY-TM6 [7], and NPxxY-TM [5] contacts in the activation mechanism could not be deduced from our simulation results. No systematic connection was found between the frequency of those interactions and physiologically relevant receptor states and complexes within the time frame of simulations. The frequency of CWxP-TM7 interactions was, however, significantly higher in the ligand-free receptor (Table 1). Furthermore, a salt bridge between R165<sup>3,50</sup> (DRY) and D340<sup>8,47</sup> (H8) was found to be present only in the active state and most frequent in the presence of the  $G_i$  protein complex. This latter specific interaction was not described previously as, opposed to the aforementioned contacts, it was not evidently present in the reported high-resolution structures [5–7]. Our data suggest that this contact could be important for receptor activation and it is further supported by earlier mutation experiments [43].

**Table 1.** Frequency of intramolecular salt bridges and H-bonds expressed as percentages of the total conformational ensemble, generated by MD simulations.

Interactions	Residues Involved	Active State					Inactive State				
		G <sub>i</sub> Protein Complex, No Ligand	G <sub>i</sub> Protein Complex, Allosteric Na <sup>+</sup>	G <sub>i</sub> Protein Complex	Beta-arrestin-2	Nb39 Nanobody	Fused T4-lysozyme	G <sub>i</sub> Protein Complex	Beta-arrestin-2	Nb39 Nanobody	Fused T4-lysozyme
Salt bridges											
DRY-H8	R165 <sup>3.30</sup> , D340 <sup>8.47</sup>	6.9	34.12	53.5	8.8	29.4	53.0	0.0	0.0	0.0	0.0
	D164 <sup>3.49</sup> , R165 <sup>3.50</sup>	9.1	0.96	0.1	5.4	0.1	0.3	8.9	10.9	17.9	6.4
H-bonds											
intra-DRY	D164 <sup>3.49</sup> , R165 <sup>3.50</sup>	63.7	1.25	0.2	19.3	0.0	0.9	21.7	35.1	82.6	11.0
DRY-ICL2	D164 <sup>3.49</sup> , R179	0.4	99.0	98.1	99.6	100.0	100.0	100.0	100.0	100.0	95.2
DRY-TM5	R165 <sup>3.50</sup> , Y252 <sup>5.58</sup>	0.0	0.2	0.0	5.7	5.7	3.1	0.0	0.0	0.0	0.0
DRY-TM6	R165 <sup>3.50</sup> , T279 <sup>6.34</sup>	0.0	0.0	0.0	0.0	0.0	0.0	37.7	2.4	98.8	91.0
CWXP-TM7	C292 <sup>6.47</sup> , W293 <sup>6.48</sup> , N328 <sup>7.45</sup>	64.5	0.2	18.4	0.8	7.3	33.3	0.0	4.6	0.1	0.7
NPxxY-TM network											
	N332 <sup>7.49</sup> , Y336 <sup>7.53</sup> , L158 <sup>3.43</sup> , Y252 <sup>5.58</sup> , V285 <sup>6.40</sup>	0.6	44.4	5.5	3.4	40.7	7.1	0.3	0	0.1	0.0

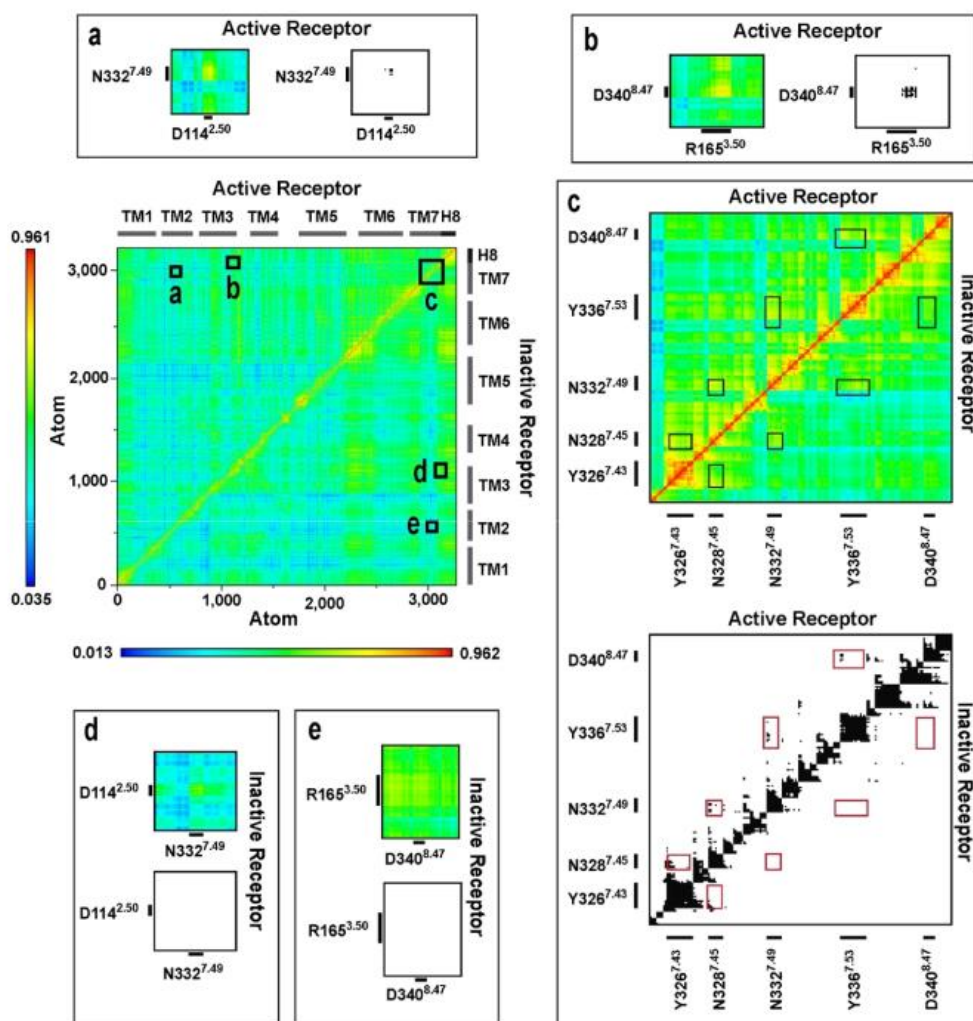
BP = orthosteric binding pocket of the mu opioid receptor; EM2 = endomorphin-2; ICL2 = 2nd intracellular loop of the mu opioid receptor; H5 = helix 5 of the G<sub>i</sub> protein α subunit; FL / ML / CL = finger loop / middle loop / C loop of beta-arrestin-2; Nb39 = Nb39 nanobody. Ball-and-stick representation of residues is indicated in superscript. <sup>†</sup> Described in reference [5].

### 3.6. Correlated Side-Chain Motions and the Polar Signaling Channel

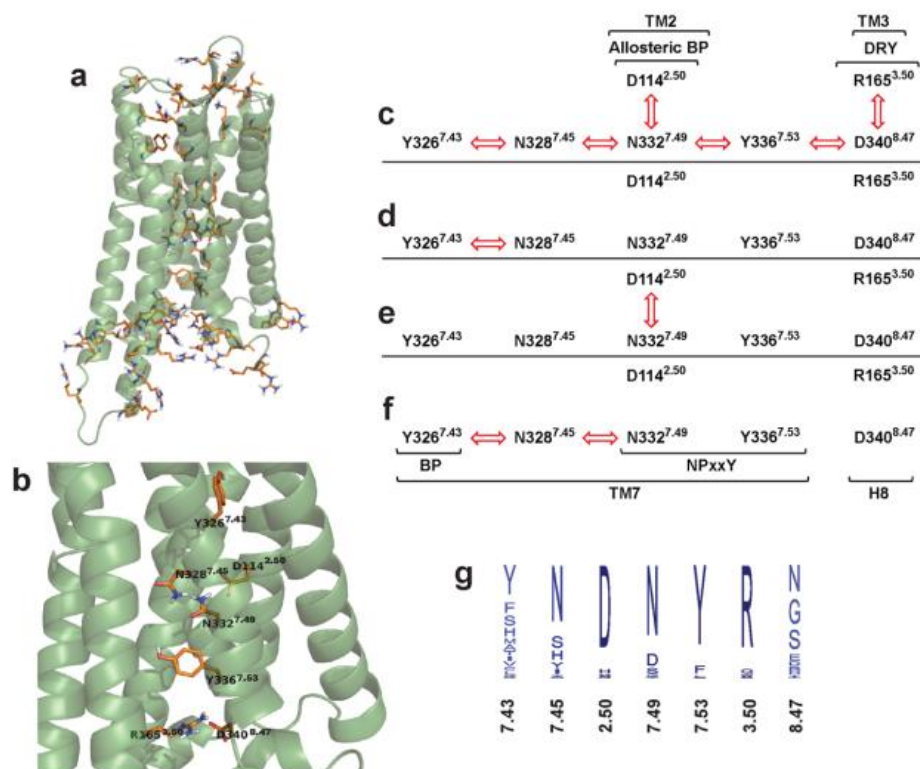
Dynamic cross-correlation analysis of the transmembrane domain and the extra- and intracellular loops provided the most conspicuous results (Figure 2). According to those, the orthosteric binding pocket is connected to the intracellular surface through a channel of polar amino acid residues, of which motions are highly correlated (Figure 3). Such concerted motions were observed only for the active receptor-G<sub>i</sub> protein complex (Figures 3 and S13–S18), suggesting that this phenomenon has a fundamental role in G protein-mediated signaling. Interestingly, no such concerted motions of these polar amino acid side chains were found in the ligand-free receptor either, although it was expected on the basis of higher conformational flexibility and consequential constitutive activity. The freezing or decoupling of these correlated motions may be attributed to the shift of a positive charge from the orthosteric binding site to the allosteric Na<sup>+</sup> binding site, or to the notable disposition of transmembrane helices observed during the simulation of the ligand-free receptor. The former potential reason is further supported by the apparent lack of correlated motions in the active MOP-G<sub>i</sub> protein-EM2-allosteric Na<sup>+</sup> system, where the Na<sup>+</sup> ion was shown to connect frequently with the aforementioned channel of polar residues (Figure S7b). On the other hand, both reasons could be debated, considering that neither such charge transfer nor similar TM helix disposition was observed in the other reference systems. Residues of the above identified polar signaling channel are located mostly on the 7<sup>th</sup> transmembrane helix (TM7), in the inner region of the transmembrane helical bundle, most distant from the surrounding membrane environment. All channel residues are parts of highly conserved functional motifs and allosteric Na<sup>+</sup> binding sites, except for Y326<sup>7,43</sup> of the binding pocket and N340<sup>8,47</sup> at the G protein-binding interface. The increased variability of these two residues is associated with ligand and G protein specificity, respectively. Special attention should be paid to residue Y326<sup>7,43</sup> of the binding pocket. Similar to highly conserved residues of the binding pocket, such as D147<sup>3,32</sup>, often referred to as the “anchor residue”, this variable residue is always found to be in close contact with the bound ligands in various class A GPCRs. Furthermore, there is a strong complementary relationship between the chemical properties of the endogenous class A GPCR ligands and the chemical properties of the residue in this position. Most interestingly, the corresponding K296<sup>7,43</sup> of rhodopsin is covalently linked to retinal, a cofactor with a highly delocalized electronic system [43]. Considering, that rhodopsin is activated upon the interaction between retinal and a single photon, high importance of this delocalized electronic system and the interplay of quantum effects in the initiation of signal transduction could be assumed, at least for this GPCR. In such mechanism K296<sup>7,43</sup>, the first unit of the polar signaling channel, would act as a sensor residue. The high degree of conservation of polar signaling channel residues suggest that this theory could be extended to other class A GPCRs, but indeed, further verification is needed. Analysis of the individual dynamics of these specific side chains revealed that the observed movements are small and mostly occur without the transition between rotameric states (Table S2). Considering that the orientation of amino acid side chains in the orthosteric binding pocket of the MOP are nearly identical in the agonist- [5,6] and antagonist-bound states [7], our results suggest that the underlying event of receptor activation is the parallel change of macroscopic polarization in a shielded central duct of the transmembrane domain. Small, simultaneous changes of side chain orientations of polar signaling channel residues could not necessarily result in direct interactions between them. Small conformational changes of polar or charged residues, however, could result in significant alterations in the local electronic structures, which could then propagate along the structure. The involvement of protons and water molecules in the transmission of signal could also be presumed, since the orthosteric binding pocket and the G protein-binding interface is connected by a hydrated pathway [5,6], and water molecules were observed to exchange rapidly between the internal cavities of the TM domain during simulations (Figure S19, Table S3). Although classical force field methods cannot provide quantitative details of such mechanism, independent mutation data provides direct evidence for the interplay of these polar and charged amino



acid side chains during receptor activation. Impaired G protein signaling, or elevation of constitutional activity, was observed for mutant receptors, where residues of the above mentioned polar signaling channel were replaced [12,44–52], while receptor activity was preserved in double mutants, where the net charge of channel residues was kept intact [48].



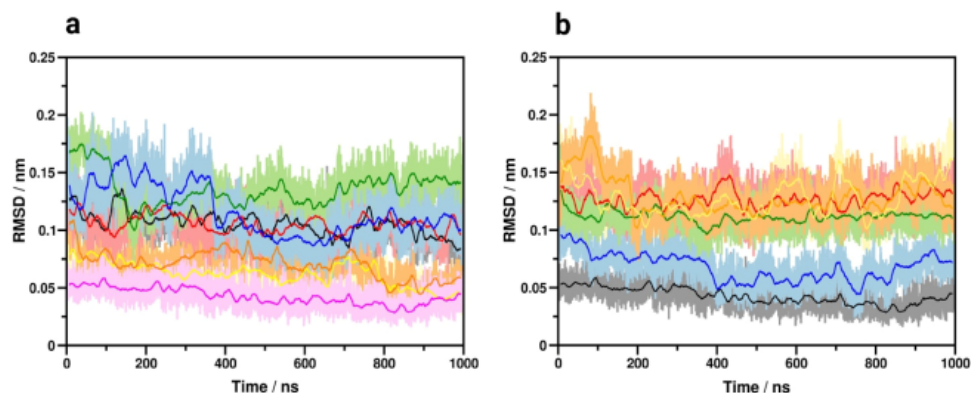
**Figure 2.** Dynamic cross-correlation matrices of the  $G_1$  protein-bound MOP in active and inactive states. Panels (a–e) are magnified views of regions of amino acid residues of interest. Black and white panels show correlations above the threshold of 0.7 MI.



**Figure 3.** The polar signaling channel of the MOP revealed by dynamic cross-correlation analysis. (a) Polar amino acids of which motions are correlated in the G<sub>i</sub> protein-bound active state. (b) Polar amino acids of which motions are correlated and connecting the orthosteric binding pocket to the G protein-binding interface. Diagrams of channel residues in (c) the active receptor-G<sub>i</sub> protein, (d) inactive receptor-G<sub>i</sub> protein, (e) active receptor-beta-arrestin-2 and (f) inactive receptor-beta-arrestin-2 complexes. Red arrows indicate correlated motions of the respective amino acids. (g) Degree of conservation of polar signaling channel residues of class A GPCRs. Non-polar hydrogens are omitted for clarity.

### 3.7. TM7 Dipole Moment

The shift of macroscopic polarization may be assisted by the inherent dipole moments of TM helices. Generally, a more ordered  $\alpha$ -helical segment possesses a higher dipole moment, which can participate in various conduction processes [53]. Further indirect indication of the plausibility of the proposed mechanism is provided by the analysis of the evolution of helix properties during simulations. Results of such analysis revealed that TM7 is the most ordered among the TM helices of the active G<sub>i</sub> protein-bound receptor. Furthermore, the helicity of TM7 is closest to ideal when the receptor is G<sub>i</sub> protein-bound, and least ideal when complexed by beta-arrestin-2, presumably providing TM7 with the highest dipole moment in the G<sub>i</sub> protein-bound state, compared to all other receptor states (Figure 4). This accentuates the role of TM7 in the activation mechanism and it is corroborated by previous reports [12,54]. The role of electrostatic forces and the importance of charge balance is further supported by the known effect of elevated concentrations of Na<sup>+</sup> ions [9] and the concept of voltage sensing [55,56]. According to this latter, changes in the transmembrane electrostatic potential ( $V_m$ ) resulting from the rearrangement of charged species and polar membrane components elicit functional effects in GPCRs.



**Figure 4.** Properties of transmembrane helices. (a) Deviation from ideal  $\alpha$ -helical geometry in the  $G_i$  protein-bound active state. Black: TM1, red: TM2, green: TM3, blue: TM4, yellow: TM5, orange: TM6, magenta: TM7. (b) Deviation of TM7 from ideal  $\alpha$ -helical geometry in the active,  $G_i$  protein-bound (black), beta-arrestin-2 bound (red), Nb39 nanobody-bound (blue), T4-lysozyme-fused (green),  $G_i$  protein-bound, ligand free (orange) and  $G_i$  protein, EM2 and allosteric  $Na^+$ -bound states (yellow).

#### 4. Conclusions

The above presented results and considerations, as well as comparison to published mutation data have led us to suggest that large scale structural rearrangements may not be the key event of receptor activation. We suggest that the current theory of GPCR activation could be extended to include that the signal transduction mechanism may be initiated by the perturbation of the electrostatic balance within the binding pocket. Such perturbation is then propagated to the intracellular G protein-binding interface through the minuscule rearrangement of polar amino acid side chains of highly conserved structural motifs, located along TM7, while assisted by the inherent dipole moment of that helical segment. This alternative perspective of the activation mechanism, corroborated by a number of earlier indications [44–52], may lead to a more accurate explanation of ligand induced effects in multiple functional states. More importantly, this could highlight certain physico-chemical properties of ligands with different functional properties and may provide a new perspective for medicinal chemists in the pursuit of a new generation of GPCR drugs.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/biom11050670/s1>, Figure S1: Comparison of the constructed peptide agonist-bound active MOP- $G_i$  protein to the recently published cryo-electron microscopic structure, Figure S2: Analysis data for a representative folding simulation of the N-terminal domain of the MOP, Figures S3 and S4: Analysis of the dynamics of N- and C-terminal domains, Figure S5: Analysis of membrane properties during equilibration and production simulations, Figures S6 and S7: Analysis of  $Na^+$  ion presence in the allosteric and orthosteric binding sites, Figure S8: TM6 disposition analysis, Figure S9: Secondary structure of ICL1, Figure S10: Secondary structure of H8, Figure S11: Secondary structure of ICL2, Figure S12: Intermolecular H-bonds between ICL2 and the  $G_i$  protein, Figures S13–S17: Dynamic cross-correlation matrices, Figure S18: The polar signaling channel of the  $G_i$  protein or beta-arrestin-2-bound MOP, Figure S19: Transmembrane water network in the active,  $G_i$  protein-bound MOP, Table S1: Intermolecular salt bridges and H-bonds, Table S2: Side chain rotameric states and transitions of the residues of the polar signaling channel, Table S3: Water occupancy in the transmembrane domain of the active,  $G_i$  protein-bound MOP.

**Author Contributions:** Conceptualization, A.B.; methodology, A.M. and A.B.; formal analysis, A.M., A.S., M.R.S., and A.B.; investigation, A.M., M.R.S., and A.B.; writing—original draft preparation, A.M. and A.B.; writing—review and editing, A.S. and M.R.S.; visualization, A.M. and A.B.; supervision,



A.B.; project administration, A.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** Scholarship for A.S. was provided by the ‘Stipendium Hungaricum’ programme of the Hungarian Ministry of Foreign Affairs and Trade and the Tempus Public Foundation. M.R.S. was supported by the ÚNKP-19-3-SZTE-269 New National Excellence Program of the Ministry for Innovation and Technology.

**Acknowledgments:** Computing resources were provided by the Government Agency of Information Technology Development, Hungary.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Hauser, A.S.; Attwood, M.M.; Rask-Andersen, M.; Schiöth, H.B.; Gloriam, D.E. Trends in GPCR Drug Discovery: New Agents, Targets and Indications. *Nat. Rev. Drug Discov.* **2017**, *16*, 829–842. [CrossRef] [PubMed]
- Munk, C.; Mutt, E.; Isberg, V.; Nikolajsen, L.F.; Bibbe, J.M.; Flock, T.; Hanson, M.A.; Stevens, R.C.; Deupi, X.; Gloriam, D.E. An Online Resource for GPCR Structure Determination and Analysis. *Nat. Methods* **2019**, *16*, 151–162. [CrossRef] [PubMed]
- Latorraca, N.R.; Venkatakrishnan, A.J.; Dror, R.O. GPCR Dynamics: Structures in Motion. *Chem. Rev.* **2017**, *117*, 139–155. [CrossRef] [PubMed]
- Kapoor, A.; Martinez-Rosell, G.; Provati, D.; de Fabritiis, G.; Filizola, M. Dynamic and kinetic elements of  $\mu$ -opioid receptor functional selectivity. *Sci. Rep.* **2017**, *7*, 11255. [CrossRef] [PubMed]
- Huang, W.; Manglik, A.; Venkatakrishnan, A.J.; Laeremans, T.; Feinberg, E.N.; Sanborn, A.L.; Kato, H.E.; Livingston, K.E.; Thorsen, T.S.; Kling, R.C.; et al. Structural Insights into  $\mu$ -Opioid Receptor Activation. *Nature* **2015**, *524*, 315–321. [CrossRef]
- Koehl, A.; Hu, H.; Maeda, S.; Zhang, Y.; Qu, Q.; Paggi, J.M.; Latorraca, N.R.; Hilger, D.; Dawson, R.; Matile, H.; et al. Structure of the  $\mu$ -Opioid Receptor-G<sub>i</sub> Protein Complex. *Nature* **2018**, *558*, 547–552. [CrossRef] [PubMed]
- Manglik, A.; Kruse, A.C.; Kobilka, T.S.; Thian, F.S.; Mathiesen, J.M.; Sunahara, R.K.; Pardo, L.; Weis, W.I.; Kobilka, B.K.; Granier, S. Crystal Structure of the  $\mu$ -Opioid Receptor Bound to a Morphinan Antagonist. *Nature* **2012**, *485*, 321–326. [CrossRef]
- Sounier, R.; Mas, C.; Steyaert, J.; Laeremans, T.; Manglik, A.; Huang, W.; Kobilka, B.K.; Dénéné, H.; Granier, S. Propagation of Conformational Changes During  $\mu$ -Opioid Receptor Activation. *Nature* **2015**, *524*, 375–378. [CrossRef]
- Pert, C.B.; Pasternak, G.; Snyder, S.H. Opiate agonists and antagonists discriminated by receptor binding in brain. *Science* **1973**, *182*, 1359–1361. [CrossRef]
- Liu, W.; Chun, E.; Thompson, A.A.; Chubukov, P.; Xu, F.; Katritch, V.; Han, G.W.; Roth, C.B.; Heitman, L.H.; IJzerman, A.P.; et al. Structural basis for allosteric regulation of GPCRs by sodium ions. *Science* **2012**, *337*, 232–236. [CrossRef]
- Zhang, C.; Srinivasan, Y.; Arlow, D.H.; Fung, J.J.; Palmer, D.; Zheng, Y.; Green, H.F.; Pandey, A.; Dror, R.O.; Shaw, D.E.; et al. High-resolution crystal structure of human protease-activated receptor 1. *Nature* **2012**, *492*, 387–392. [CrossRef] [PubMed]
- Fenalti, G.; Giguere, P.M.; Katritch, V.; Huang, X.P.; Thompson, A.A.; Cherezov, V.; Roth, B.L.; Stevens, R.C. Molecular Control of  $\delta$ -Opioid Receptor Signaling. *Nature* **2014**, *506*, 191–196. [CrossRef]
- Dror, R.O.; Arlow, D.H.; Maragakis, P.; Mildorf, T.J.; Pan, A.C.; Xu, H.; Borhani, D.W.; Shaw, D.E. Activation Mechanism of the  $\beta_2$ -Adrenergic Receptor. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 18684–18689. [CrossRef]
- Marino, K.A.; Shang, Y.; Filizola, M. Insights into the Function of Opioid Receptors from Molecular Dynamics simulations of Available Crystal Structures. *Br. J. Pharmacol.* **2018**, *175*, 2834–2845. [CrossRef]
- Challiss, R.A.; Wess, J. GPCR-G Protein Preassembly? *Nat. Chem. Biol.* **2011**, *7*, 657–658. [CrossRef]
- Strohman, M.J.; Maeda, S.; Hilger, D.; Masureel, M.; Du, Y.; Kobilka, B.K. Local Membrane Charge Regulates  $\beta_2$ -Adrenergic Receptor Coupling to G<sub>B</sub>. *Nat. Commun.* **2019**, *10*, 2234. [CrossRef]
- Pike, L.J.; Han, X.; Chung, K.N.; Gross, R.W. Lipid rafts are enriched in arachidonic acid and plasmalogen ethanolamine and their composition is independent of caveolin-1 expression: a quantitative electrospray ionization/mass spectrometric analysis. *Biochemistry* **2002**, *41*, 2075–2088. [CrossRef] [PubMed]
- Zadina, J.E.; Hackler, L.; Ge, L.J.; Kastin, A.J. A Potent and Selective Endogenous Agonist for the Mu-Opioid Receptor. *Nature* **1997**, *386*, 499–502. [CrossRef] [PubMed]
- Jo, S.; Kim, T.; Iyer, V.G.; Im, W. CHARMM-GUI: A Web-based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859–1865. [CrossRef]
- Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. [CrossRef]
- Li, D.-W.; Bruschweiler, R. NMR-based protein potentials. *Angew. Chem. Int. Ed.* **2010**, *49*, 6778–6780. [CrossRef]
- Qiu, D.; Shenkin, P.; Hollinger, F.; Still, W. The GB/SA Continuum Model for Solvation. A Fast Analytical Method for the Calculation of Approximate Born Radii. *J. Phys. Chem. A* **1997**, *101*, 3005–3014. [CrossRef]
- Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101. [CrossRef]
- Gupta, R.; Jung, E.; Brunak, S. Prediction of N-glycosylation Sites in Human Proteins. In Preparation. 2004. Available online: <http://www.cbs.dtu.dk/services/NetNGlyc/> (accessed on 28 April 2021).



25. Huang, P.; Chen, C.; Xu, W.; Yoon, S.-I.; Unterwald, E.M.; Pintar, J.E.; Wang, Y.; Chong, P.L.-G.; Liu-Chen, L.-Y. Brain region-specific N-glycosylation and lipid rafts association of the rat mu opioid receptors. *Biochem. Biophys. Res. Commun.* **2008**, *365*, 82–88. [\[CrossRef\]](#)
26. Mann, A.; Illing, S.; Miess, E.; Schulz, S. Different mechanisms of homologous and heterologous  $\mu$ -opioid receptor phosphorylation. *Br. J. Pharmacol.* **2015**, *172*, 311–316. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Zheng, H.; Pearsall, E.A.; Hurst, D.P.; Zhang, Y.; Chu, J.; Zhou, Y.; Reggio, P.H.; Loh, H.A.; Law, P.-Y. Palmitoylation and membrane cholesterol stabilize  $\mu$ -opioid receptor homodimerization and G protein coupling. *BMC Cell Biol.* **2012**, *13*, 1–18. [\[CrossRef\]](#)
28. Ingólfsson, H.I.; Melo, M.N.; van Eerden, F.J.; Arnarez, C.; Lopez, C.A.; Wassenaar, T.A.; Periole, X.; de Vries, A.H.; Tieleman, D.P.; Marrink, S.J. Lipid organization of the plasma membrane. *J. Am. Chem. Soc.* **2014**, *136*, 14554–14559. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. Autodock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comp. Chem.* **2009**, *16*, 2785–2791. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Berendsen, H.J.C.; Postma, J.P.M.; DiNola, A.; Haak, J.R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690. [\[CrossRef\]](#)
31. Kabsch, W.; Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577–2637. [\[CrossRef\]](#)
32. Buchoux, S. FATSlim: A fast and robust software to analyze MD simulations of membranes. *Bioinformatics* **2017**, *33*, 133–134. [\[CrossRef\]](#)
33. Lange, O.F.; Grubmüller, H. Generalized correlation for biomolecular dynamics. *Proteins* **2006**, *62*, 1053–1061. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Humphrey, W.; Dalke, A.; Schulten, K. VMD—Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [\[CrossRef\]](#)
35. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.A.; Clamp, M.; Barton, G.J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189–1191. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Kučerka, N.; Nieh, M.P.; Katsaras, J. Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochim. Biophys. A* **2011**, *1808*, 2761–2771. [\[CrossRef\]](#)
38. Niemelä, P.S.; Ollila, S.; Hyvönen, M.T.; Karttunen, M.; Vattulainen, I. Assessing the nature of lipid raft membranes. *PLoS Comp. Biol.* **2007**, *3*, e34. [\[CrossRef\]](#)
39. Yuan, S.; Vogel, H.; Filipek, S. The Role of Water and Sodium Ions in the Activation of the  $\mu$ -Opioid Receptor. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 10112–10115. [\[CrossRef\]](#)
40. Shang, Y.; LeRouzic, V.; Schneider, S.; Bisignano, P.; Pasternak, G.W.; Filizola, M. Mechanistic Insights into the Allosteric Modulation of Opioid Receptors by Sodium Ions. *Biochemistry* **2014**, *53*, 5140–5149. [\[CrossRef\]](#)
41. Hu, X.; Wang, Y.; Hunkele, A.; Provasi, D.; Pasternak, G.W.; Filizola, M. Kinetic and thermodynamic insights into sodium ion translocation through the  $\mu$ -opioid receptor from molecular dynamics and machine learning analysis. *PLoS Comput. Biol.* **2019**, *15*, e1006689. [\[CrossRef\]](#)
42. Ma, X.; Hu, Y.; Batebi, H.; Heng, J.; Xu, J.; Liu, X.; Niu, X.; Li, H.; Hildebrand, P.W.; Jin, C.; et al. Analysis of  $\beta_2$ AR-G<sub>s</sub> and  $\beta_2$ AR-G<sub>i</sub> complex formation by NMR spectroscopy. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 23096–23105. [\[CrossRef\]](#)
43. Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C.A.; Motoshima, H.; Fox, B.A.; Le Trong, I.; Teller, D.C.; Okada, T.; Stenkamp, R.E.; et al. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **2000**, *289*, 739–745. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Jongejans, A.; Bruysters, M.; Ballesteros, J.A.; Haaksma, E.; Bakker, R.A.; Pardo, L.; Leurs, R. Linking Agonist Binding to Histamine H1 Receptor Activation. *Nat. Chem. Biol.* **2005**, *1*, 98–103. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Liu, R.; Nahon, D.; le Roy, B.; Lenselink, E.B.; Ijzerman, A.P. Scanning Mutagenesis in a Yeast System Delineates the Role of the NPxxY(x)<sub>5</sub>6F Motif and Helix 8 of the Adenosine A<sub>2B</sub> Receptor in G Protein Coupling. *Biochem. Pharmacol.* **2015**, *95*, 290–300. [\[CrossRef\]](#)
46. Hothersall, J.D.; Torella, R.; Humphreys, S.; Hooley, M.; Brown, A.; McMurray, G.; Nickolls, S.A. Residues W320 and Y328 within the Binding Site of the  $\mu$ -Opioid Receptor Influence Opiate Ligand Bias. *Neuropharmacology* **2017**, *118*, 46–58. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Sealfon, S.C.; Chi, L.; Ebersole, B.J.; Rodic, V.; Zhang, D.; Ballesteros, J.A.; Weinstein, H. Related Contribution of Specific Helix 2 and 7 Residues to Conformational Activation of the Serotonin 5-HT<sub>2A</sub> Receptor. *J. Biol. Chem.* **1995**, *270*, 16683–16688. [\[CrossRef\]](#)
48. Xu, W.; Ozdener, F.; Li, J.G.; Chen, C.; de Riel, J.K.; Weinstein, H.; Liu-Chen, L.Y. Functional Role of the Spatial Proximity of Asp114<sup>2.50</sup> in TMH 2 and Asn332<sup>7.49</sup> in TMH 7 of the Mu Opioid Receptor. *FEBS Lett.* **1999**, *447*, 318–324. [\[CrossRef\]](#)
49. Galés, C.; Kowalski-Chauvel, A.; Dufour, M.N.; Seva, C.; Moroder, L.; Pradayrol, L.; Vaysse, N.; Fourmy, D.; Silvente-Poirot, S. Mutation of Asn-391 within the Conserved NPxxY Motif of the Cholecystokinin B Receptor Abolishes C<sub>q</sub> Protein Activation without Affecting Its Association with the Receptor. *J. Biol. Chem.* **2000**, *275*, 17321–17327. [\[CrossRef\]](#)
50. Barak, L.S.; Tiberi, M.; Freedman, N.J.; Kwatra, M.M.; Lefkowitz, R.J.; Caron, M.G. A Highly Conserved Tyrosine Residue in G Protein-Coupled Receptors is Required for Agonist-Mediated Beta 2-Adrenergic Receptor Sequestration. *J. Biol. Chem.* **1994**, *269*, 2790–2795. [\[CrossRef\]](#)

51. Prioleau, C.; Visiers, I.; Ebersole, B.J.; Weinstein, H.; Sealfon, S.C. Conserved Helix 7 Tyrosine Acts as a Multistate Conformational Switch in the 5-HT<sub>2C</sub> Receptor. Identification of a Novel “LOCKED-ON” Phenotype and Double Revertant Mutations. *J. Biol. Chem.* **2002**, *277*, 36577–36584. [[CrossRef](#)]
52. Kalatskaya, I.; Schüssler, S.; Blaukat, A.; Müller-Esterl, W.; Jochum, M.; Proud, D.; Faussner, A. Mutation of Tyrosine in the Conserved NPxxY Sequence Leads to Constitutive Phosphorylation and Internalization, but Not Signaling of the Human B<sub>2</sub> Bradykinin Receptor. *J. Biol. Chem.* **2004**, *279*, 31268–31276. [[CrossRef](#)]
53. Hol, W.G. Effects of the  $\alpha$ -Helix Dipole Upon the Functioning and Structure of Proteins and Peptides. *Adv. Biophys.* **1985**, *19*, 133–165. [[CrossRef](#)]
54. Bartuzi, D.; Kaczor, A.A.; Matosiuk, D. Interplay between two allosteric sites and their influence on agonist binding in human  $\mu$  opioid receptor. *J. Chem. Inf. Model.* **2016**, *56*, 563–570. [[CrossRef](#)]
55. Vickery, O.N.; Machtens, J.P.; Tamburrino, G.; Seeliger, D.; Zachariae, U. Structural mechanisms of voltage sensing in G protein-coupled receptors. *Structure* **2016**, *24*, 997–1007. [[CrossRef](#)] [[PubMed](#)]
56. Mahaut-Smith, M.P.; Martinez-Pinna, J.; Gurung, I.S. A role for membrane potential in regulating GPCRs? *Trends Pharmacol. Sci.* **2008**, *29*, 421–429. [[CrossRef](#)] [[PubMed](#)]

## II

**Mitra, A** and Sarkar, and Borics, A.

*Universal properties and specificities of the  $\beta_2$ -adrenergic receptor- $G_s$  protein complex activation mechanism revealed by all-atom molecular dynamics simulations.*

INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES, 22(19), 10423 (2021). doi:  
10.3390/ijms221910423

IF: 5.923 [Q1]



Article

# Universal Properties and Specificities of the $\beta_2$ -Adrenergic Receptor- $G_s$ Protein Complex Activation Mechanism Revealed by All-Atom Molecular Dynamics Simulations

Argha Mitra <sup>1,2</sup>, Arijit Sarkar <sup>1,2</sup> and Attila Borics <sup>1,\*</sup>

<sup>1</sup> Laboratory of Chemical Biology, Institute of Biochemistry, Biological Research Centre, 62. Temesvári krt., H-6726 Szeged, Hungary; argha.mitra@brc.hu (A.M.); sarkar.arajit@brc.hu (A.S.)

<sup>2</sup> Theoretical Medicine Doctoral School, Faculty of Medicine, University of Szeged, 97. Tisza L. krt., H-6722 Szeged, Hungary

\* Correspondence: borics.attila@brc.hu; Tel.: +36-62-599-600 (ext. 430)



**Citation:** Mitra, A.; Sarkar, A.; Borics, A. Universal Properties and Specificities of the  $\beta_2$ -Adrenergic Receptor- $G_s$  Protein Complex Activation Mechanism Revealed by All-Atom Molecular Dynamics Simulations. *Int. J. Mol. Sci.* **2021**, *22*, 10423. <https://doi.org/10.3390/ijms221910423>

Academic Editor: Francisco Ciruela

Received: 30 August 2021

Accepted: 23 September 2021

Published: 27 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** G protein-coupled receptors (GPCRs) are transmembrane proteins of high pharmacological relevance. It has been proposed that their activity is linked to structurally distinct, dynamically interconverting functional states and the process of activation relies on an interconnecting network of conformational switches in the transmembrane domain. However, it is yet to be uncovered how ligands with different extents of functional effect exert their actions. According to our recent hypothesis, based on indirect observations and the literature data, the transmission of the external stimulus to the intracellular surface is accompanied by the shift of macroscopic polarization in the transmembrane domain, furnished by concerted movements of highly conserved polar motifs and the rearrangement of polar species. In this follow-up study, we have examined the  $\beta_2$ -adrenergic receptor ( $\beta_2$ AR) to see if our hypothesis drawn from an extensive study of the  $\mu$ -opioid receptor (MOP) is fundamental and directly transferable to other class A GPCRs. We have found that there are some general similarities between the two receptors, in agreement with previous studies, and there are some receptor-specific differences that could be associated with different signaling pathways.

**Keywords:** GPCR; adrenergic; activation mechanism; signal transduction; molecular dynamics

## 1. Introduction

G protein-coupled transmembrane receptors (GPCRs) constitute one of the largest and most important protein superfamilies of the human genome. Their importance mainly derives from their remarkably high pharmacological relevance [1–3]. GPCRs are classified into six sub-families (A–F) based on their sequence and function. From the structural perspective GPCRs share similar architecture in their transmembrane domains and possess high sequential and structural diversity of their extra- and intracellular loops and the extracellular (N-terminal) and cytosolic (C-terminal) domains. These variable domains are proposed to be responsible for ligand and G protein/arrestin specificity, whereas the transmembrane (TM) domain controls the transmission of external signals to the intracellular surface of the protein. The variety of G proteins that mediate GPCR signaling is very low relative to the diversity of GPCRs and their external activators. Therefore, the activation of GPCRs is suggested to follow a general structural mechanism.

Adrenergic receptors, responsible for the homeostasis between stressful and resting conditions of the body, belong to the most populated class (A) of GPCRs. The  $\beta_2$ -adrenergic receptor ( $\beta_2$ AR) is linked to respiratory, as well as other smooth muscle relaxation. These receptors are stimulated by endogenous neurotransmitters, such as epinephrine and norepinephrine and the signal transduction of  $\beta_2$ AR is mediated predominantly by the  $G_s$  (stimulatory) protein complex [4]. The agonist- $\beta_2$ AR- $G_s$  signaling triad is an important target of drug development activities to treat severe respiratory conditions, such as chronic

obstructive pulmonary disease (COPD) or asthma. Remarkable efforts have been invested in the optimization of GPCR targeting drugs to reduce their undesired side effects. Such efforts necessitated the in-depth structural analysis of target GPCR structures [2].

The  $\beta$ 2AR was among the first GPCRs of which a three-dimensional structure was solved at atomic resolution [5] and numerous experimental studies have been conducted to elucidate the structural details of activation of this TM receptor [6–15]. To date, the  $\beta$ 2AR is the most widely investigated class A GPCR and is frequently used as a universal model for the study of the structural mechanism of activation for class A GPCRs [16]. Recent developments in the field of experimental structural biology led to the rapid accumulation of experimental data of GPCR structures which are now readily available in the Brookhaven Protein Data Bank as well as in GPCRdb, a specific, comprehensive collection of all GPCR structures published to date (<http://gpcrdb.org>, accessed on 23 August 2021) [15]. A general structural difference between active and inactive state class A GPCRs, including  $\beta$ 2AR, is the position of the sixth helix of the TM domain (TM6) [5,7,8]. The capabilities of conventional experimental techniques to provide information about the mechanism of transition between these structural states are limited. The current theory of the structural mechanism of class A GPCR activation was framed on the basis of results acquired through the application of state-of-the-art molecular dynamics (MD) simulation hardware and techniques [17–19]. According to this theory, GPCRs exist in a dynamic ensemble of multiple active, inactive, and intermediate states, even in the absence of ligands. The populations of these states are shifted upon ligand binding, depending on the functional properties of the bound ligand. The fifth, sixth, and seventh TM helices (TM5, TM6, TM7, respectively) have been emphasized for their foremost interplay in signal transduction [16,20].

As well as the rearrangement of TM helices, highly conserved polar functional motifs, E/DRY, NPxxY, and CWxP, have been appointed as participants of the activation mechanism [21,22]. Specific rearrangements of intramolecular interactions involving these motifs, the conserved allosteric Na<sup>+</sup> binding site [23,24], and the extended network of water molecules in the internal cavities connecting the orthosteric ligand-binding pocket to the cytosolic domains have been proposed recently as general machinery of signal transmission in GPCRs [25,26]. In order to respond to the most recent challenge of rational drug design and to develop high-affinity, high-efficacy, and functionally selective GPCR ligands a quantitative model of the activation mechanism is necessary. Models built exclusively on a structural basis have limited capabilities to differentiate between ligands with similar structure, physico-chemical properties, and binding affinity, but different efficacy. Consequently, the creation of a directly transferable model would need the introduction of new perspectives.

Our recent results of extensive MD simulations of the  $\mu$ -opioid receptor (MOP) indicated that the dynamic motions of polar amino acid side chains of conserved motifs are highly correlated [27]. Such concerted motions were only observed during the simulations of the agonist- and G<sub>i</sub> protein-bound active state MOP, suggesting that this phenomenon could be associated with the signal transduction process, corroborating the above-cited proposal [25,26]. These polar amino acid side chains of the orthosteric and allosteric binding pockets, the NPxxY and E/DRY motifs and the cytosolic helix (H8), form a polar signaling channel connecting the binding pocket to the intracellular G protein-binding surface. Frequent transitions between rotameric states, however, were not observed for these specific side chains, which casts doubt on the channel's operation as a sequential conformational switch. According to our recent proposal, receptor activation is accompanied by a shift of macroscopic polarization in a shielded central duct of the TM domain, initiated by ligand binding and propagated by the minuscule rearrangements of polar amino acid side chains along TM7. TM7 was further implied in the activation mechanism as a potential conductor owing to its inherent dipole moment. Evidently, MD simulations employing fixed point charge force fields cannot reveal exact or quantitative details of processes involving



charge shift. Nevertheless, independent mutation data provided convincing support for the interplay of the above-mentioned polar species [27].

Here we present an extensive, unbiased, atomistic MD simulation study of the full sequence  $\beta$ 2AR, embedded in a native-like caveolar membrane bilayer, in the presence of an endogenous agonist and the Gs protein complex or  $\beta$ -arrestin-2. Simulations were started from the active and inactive structural form of the receptor, revealed by previous X-ray crystallographic studies [5,7]. Analyses were conducted according to the above-mentioned novel perspective, to examine that our previously proposed indirect hypothesis could be extended to other class A GPCRs.

## 2. Results and Discussion

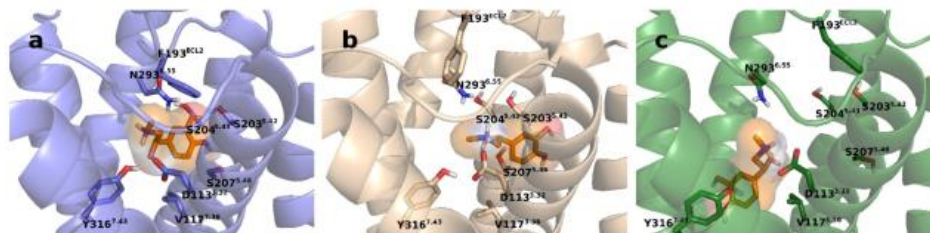
### 2.1. Simulation System Integrity

On the grounds that the disposition of TM helices was proposed to have a pivotal role in the activation mechanism [16,20,25,26], an important specific aim of this current study was to study the full sequence  $\beta$ 2AR in order to take account of the pull of the N- and C-terminal domains posed on the TM helices and to see if that affects the internal dynamics of the TM domain. These highly variable and flexible domains are generally omitted in the experimental structures of GPCRs and, consequently, from the corresponding MD simulation studies. Here, approximate structures of the terminal domains were generated through folding simulations (see Section 3). Even if parallel folding simulations provide convergent results, the correct and complete folding of these domains in the available time frame could not be guaranteed. Nevertheless, their effect on TM dynamics, primarily exerted by their mass is satisfactorily taken into account by using these approximate structures.

Unfolding of N- and C-terminal domains during simulations could result in contacts formed between the neighboring periodic images of these unfolded domains which could lead to artifacts. The evolution of the radii of gyration of N- and C-terminal domains indicated partial unfolding during some of the production simulations (Figure S1), but the minimum distance between the N- and C-terminal domains, was never below 1.4 nm (Figure S2). Therefore, the possibility of artificial contact between these domains could be excluded.

With regard to the stability of simulation systems, no dissociation or notable relative displacement of macromolecular components were observed during the simulations. Epinephrine, however, dissociated from the orthosteric binding pocket on two occasions. First, it was ejected from the binding pocket during the first 100 ns of one of the three replica simulations of the active Gs protein-bound  $\beta$ 2AR. The second time it was observed for the inactive Gs protein-bound  $\beta$ 2AR, when epinephrine left the orthosteric binding site after approximately 600 ns of simulation time.

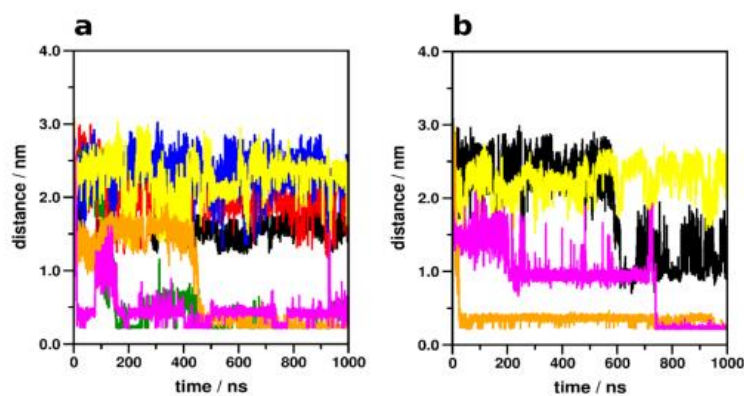
In the second replica simulation of the active Gs protein-bound  $\beta$ 2AR, the ligand took an opposite orientation in the binding pocket, compared to the X-ray crystallographic structure [8] of the  $\beta$ 2AR-epinephrine complex (Figure 1). These simulations were not excluded from analysis but the results were interpreted accordingly. A reference simulation of the active G protein-bound  $\beta$ 2AR in which epinephrine was restrained to the binding pocket was performed in order to explain discrepancies emerging from ligand dissociation. The instability of  $\beta$ 2AR-epinephrine complexes observed during some of the simulations may be explained by the smaller size and remarkably lower affinity of epinephrine relative to the ligands used in previous simulations of the  $\beta$ 2AR ( $\mu$ M vs. pM range affinity, respectively) [17,28].



**Figure 1.** Ligand position in the orthosteric binding pocket in the 1st (b) and 2nd (c) replica simulations of the active  $\beta_2$ AR- $G_s$  protein-epinephrine complex in comparison with the crystallographic structure ((a), PDB code: 4LDO).

## 2.2. Allosteric $Na^+$ Binding

$Na^+$  penetration to the allosteric  $Na^+$  binding site (D79<sup>2.50</sup>) (Figure 2, Table 1) was not observed for the epinephrine-bound  $\beta_2$ AR, regardless of the state of the receptor, except when epinephrine dissociated from the orthosteric binding site during the course of the simulation. Furthermore,  $Na^+$  entrance from the cytosolic side did not occur in any of the systems. In the absence of epinephrine, however, localization of  $Na^+$  at the ortho- (D113<sup>3.32</sup>) (Figure S3) and allosteric sites took place. Frequent contacts were formed between  $Na^+$  and the ortho- and allosteric sites as well as residues of the conserved CWxP and NPxxY motifs, but no relevant trend of contact frequencies was identified which could be directly associated with the modulation of receptor activation by  $Na^+$  ions (Table 1). Interestingly,  $Na^+$  was occasionally present in the orthosteric binding pocket of the  $G_s$  protein- and epinephrine-bound, active state  $\beta_2$ AR. This phenomenon may be less relevant since it was only noticeable in one of the simulation replicas (Figure S3) and the frequency of contact between  $Na^+$  and D113<sup>3.32</sup> was negligible (Table 1). The presence of  $Na^+$  in the orthosteric pocket and the proximal CWxP motif was more prominent for the  $G_s$  protein- and epinephrine-bound inactive state  $\beta_2$ AR. However, this was a clear consequence of ligand dissociation after approximately 600 ns of simulation time. No  $Na^+$  insertion to the ortho- or allosteric sites was observed in any of the ligand- and  $\beta$ -arrestin-2-bound states.



**Figure 2.** Minimum distance between  $Na^+$  ions and the allosteric  $Na^+$  binding site, D79<sup>2.50</sup> of the active (a) and inactive state (b)  $\beta_2$ AR during simulations. Black:  $\beta_2$ AR- $G_s$  protein-epinephrine complex, 1st replica; Red:  $\beta_2$ AR- $G_s$  protein-epinephrine complex, 2nd replica; Green:  $\beta_2$ AR- $G_s$  protein-epinephrine complex, 3rd replica; Blue:  $\beta_2$ AR- $G_s$  protein-epinephrine complex, with restrained epinephrine and GDP; Yellow:  $\beta_2$ AR- $\beta$ -arrestin-2-epinephrine complex; Orange: ligand-free  $\beta_2$ AR- $G_s$  protein complex; Magenta: ligand-free  $\beta_2$ AR- $\beta$ -arrestin-2 complex.

**Table 1.** The frequency of contact ( $d \leq 0.4$  nm) between  $\text{Na}^+$  ions and polar amino acid side chains of the allosteric and orthosteric binding pockets and nearby conserved motifs.

Residue	Frequency of Contact/%							
	Epinephrine-Bound $\beta_2\text{AR}$				Ligand-Free $\beta_2\text{AR}$			
	$\text{G}_s$ Protein Complex		$\beta$ -Arrestin-2 Complex		$\text{G}_s$ Protein Complex		$\beta$ -Arrestin-2 Complex	
	Active	Inactive	Active	Inactive	Active	Inactive	Active	Inactive
D <sub>79</sub> <sup>2.50</sup>	0.0	0.0	0.0	0.0	52.15	93.41	25.91	26.04
D <sub>113</sub> <sup>3.32</sup>	<0.1	32.33	0.0	0.0	9.22	45.06	5.02	72.03
S <sub>120</sub> <sup>3.39</sup>	0.0	0.0	0.0	0.0	44.95	96.36	5.55	25.86
W <sub>286</sub> <sup>6.48</sup>	0.0	6.88	0.0	0.0	3.36	19.51	29.16	10.92
N <sub>318</sub> <sup>7.45</sup>	0.0	0.0	0.0	0.0	2.78	0.46	46.14	0.22
S <sub>319</sub> <sup>7.46</sup>	0.0	0.0	0.0	0.0	30.52	14.39	79.75	24.69
N <sub>322</sub> <sup>7.49</sup>	0.0	0.0	0.0	0.0	45.87	51.12	8.17	20.06
Y <sub>326</sub> <sup>7.53</sup>	0.0	0.0	0.0	0.0	0.12	0.60	0.0	0.0

These observations are in complete agreement with previous MD simulation data of this receptor [29], and other class A GPCRs [27,30–32], and corroborate that the allosteric  $\text{Na}^+$  binding site is only accessible through the orthosteric binding pocket and the bound orthosteric ligand blocks the entrance of  $\text{Na}^+$  to the allosteric site. Intracellular access of  $\text{Na}^+$  ions through the TM domain is closed by the bound  $\text{G}_s$  protein complex or  $\beta$ -arrestin-2. Translocation of  $\text{Na}^+$  ions through the TM region was observed in previous MD simulations of the active state MOP in the absence of bound ligands and intracellular proteins, suggesting that it takes place during the process of receptor deactivation [33].

### 2.3. Transmembrane Helix and Loop Dynamics

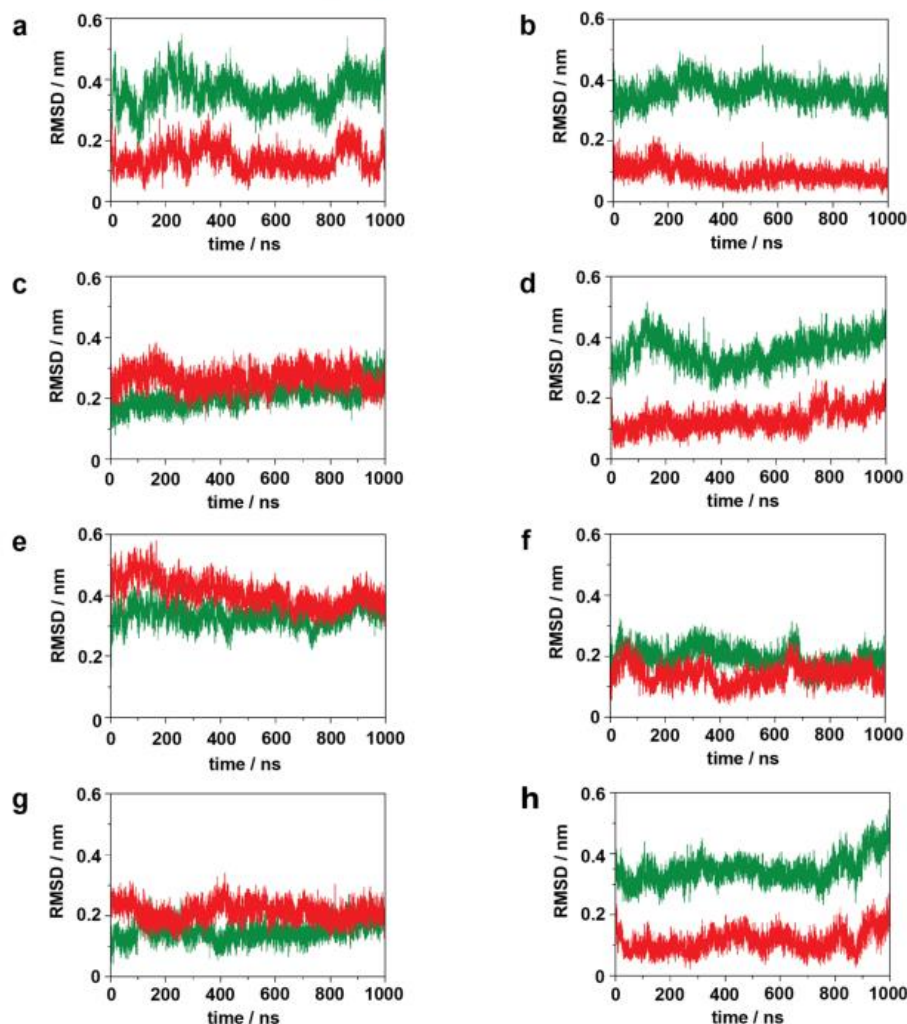
In our previous study of the MOP, TM7 was found to be the most ordered among the TM helices of the active state  $\text{G}_i$  protein-bound receptor. Furthermore, the helical conformation of TM7 was indicated to be the closest to ideal when the receptor was bound to the  $\text{G}_i$  protein complex. Conversely, the lowest degree of order of TM7 was observed in the  $\beta$ -arrestin-2-bound state. Based on these geometric features it was assumed that TM7 possesses the highest dipole moment in the  $\text{G}_i$  protein-bound state, which could facilitate electron, proton, or ion conduction along the helix axis [34]. Such conduction capacity could be relevant for the proposed model, which involves the shift of charge balance between the orthosteric binding pocket and the intracellular G protein-binding interface during class A GPCR activation [27]. Such a proposed role of TM7 was not corroborated by the results obtained here for  $\beta_2\text{AR}$  (Figure S4). As opposed to the MOP, TM7 was among the least ordered TM helices of the  $\beta_2\text{AR}$ . Furthermore, the relatively high order of TM7 in the active, epinephrine- and  $\text{G}_s$  protein-bound state was not reproduced in the reference simulation of that system and was also matched by the  $\beta$ -arrestin-2 and epinephrine-bound receptor (Figure S4). This suggests that the above-described trend is a specific property of MOP and/or class A GPCRs signaling through the  $\text{G}_i$  protein complex.

Atomic displacement analysis of the TM6 of epinephrine-bound  $\beta_2\text{AR}$  indicated moderate dispositions from the corresponding starting structures (Figures S5 and S6), similar to our previous simulation results for the MOP [27]. Remarkable rearrangements of TM helices have occurred during previous simulations of the  $\beta_2\text{AR}$ , but at longer timescales and in the absence of bound intracellular proteins [17]. Slightly larger dispositions of TM6 were observed in the ligand-free systems, confirming the stabilizing effect of the bound agonist, again, in agreement with our previous results [27]. One exception was the  $\beta$ -arrestin-2-bound inactive receptor in which, similar to the ligand-bound systems, no significant TM6 disposition took place. This, in contrast with our previous results, suggests the preference of  $\beta$ -arrestin-2 for the inactive structure of this receptor.

The unexpected disorder of TM7, discussed above, prompted us to analyze the disposition of this helical segment to see if the dynamics of the NPxxY motif show any correlation with the activation state or with the presence of agonist and/or intracellular signaling



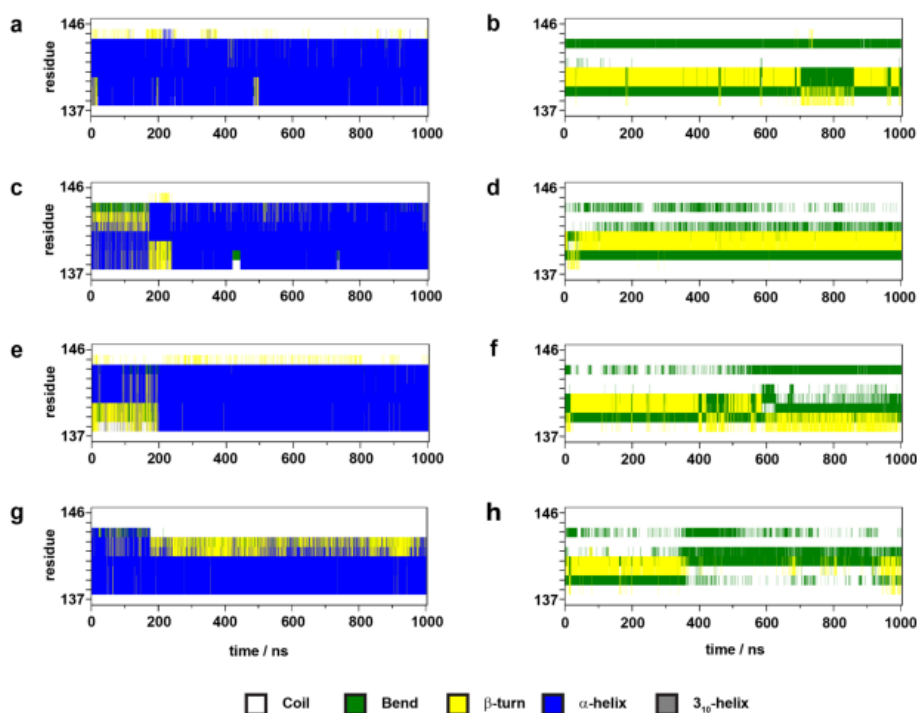
proteins. Surprisingly, large dispositions ( $\sim 0.4$  nm RMSD) of the NPxxY motif were found, relative to the active state structure and to the corresponding TM7 dispositions in several simulations (Figures 3 and S7).



**Figure 3.** Disposition of the NPxxY motif during simulations with respect to the active (green) and inactive (red) crystallographic structures of the  $\beta_2$ AR. (a) active  $\beta_2$ AR- $G_s$  protein–epinephrine complex, 1st replica; (b) inactive  $\beta_2$ AR- $G_s$  protein–epinephrine complex; (c) active  $\beta_2$ AR- $\beta$ -arrestin-2–epinephrine complex; (d) inactive  $\beta_2$ AR- $\beta$ -arrestin-2–epinephrine complex. (e) active, ligand-free  $\beta_2$ AR- $G_s$  protein complex; (f) inactive, ligand-free  $\beta_2$ AR- $G_s$  protein complex; (g) active, ligand-free  $\beta_2$ AR- $\beta$ -arrestin-2 complex; (h) inactive, ligand-free  $\beta_2$ AR- $\beta$ -arrestin-2 complex.

Most interestingly, this large disposition of the NPxxY motif coincides with the intensive concerted dynamics of the second segment of the polar signaling channel, which could be associated either with receptor activation or constitutional activity (see data and discussion below). It should be noted, that the results of the epinephrine and G protein-bound inactive state receptor and the ligand-free G protein-bound inactive  $\beta_2$ AR are very

similar, due to epinephrine dissociation after 600 ns. The results of secondary structure analysis indicated that ICL1, ICL3, and H8 maintain their secondary structures in all receptor states and only minor, reversible changes occur, resulting from internal dynamics (Figures S8–S12). ICL2, on the other hand, adopted an  $\alpha$ -helical structure in active states and got partially unfolded in inactive states (Figure 4, Figure S13). ICL2 of the active  $\beta_2$ AR was recently shown to be  $\alpha$ -helical when the receptor is bound by  $G_s$  and partially unfolded when the  $G_i$  protein complex is attached [35]. Similar signaling protein selectivity was observed previously for the MOP receptor [27]. However, in contrast with the present results, the structure of ICL2 did not demonstrate any dependence on the activation state of the MOP. Nevertheless, the present simulation results are in agreement with extensive experimental data reporting an  $\alpha$ -helical structure of ICL2 in the active [7,8,36,37] and unfolded structure in the inactive state of  $\beta_2$ AR [5,11].



**Figure 4.** Evolution of the secondary structure of ICL2 during simulations. (a) active  $\beta_2$ AR- $G_s$  protein-epinephrine complex, 1st replica; (b) inactive  $\beta_2$ AR- $G_s$  protein-epinephrine complex; (c) active  $\beta_2$ AR- $\beta$ -arrestin-2-epinephrine complex; (d) inactive  $\beta_2$ AR- $\beta$ -arrestin-2-epinephrine complex; (e) active, ligand-free  $\beta_2$ AR- $G_s$  protein complex; (f) inactive, ligand-free  $\beta_2$ AR- $G_s$  protein complex; (g) active, ligand-free  $\beta_2$ AR- $\beta$ -arrestin-2 complex; (h) inactive, ligand-free  $\beta_2$ AR- $\beta$ -arrestin-2 complex.

#### 2.4. Correlated Side-Chain Motions in the Transmembrane Domain

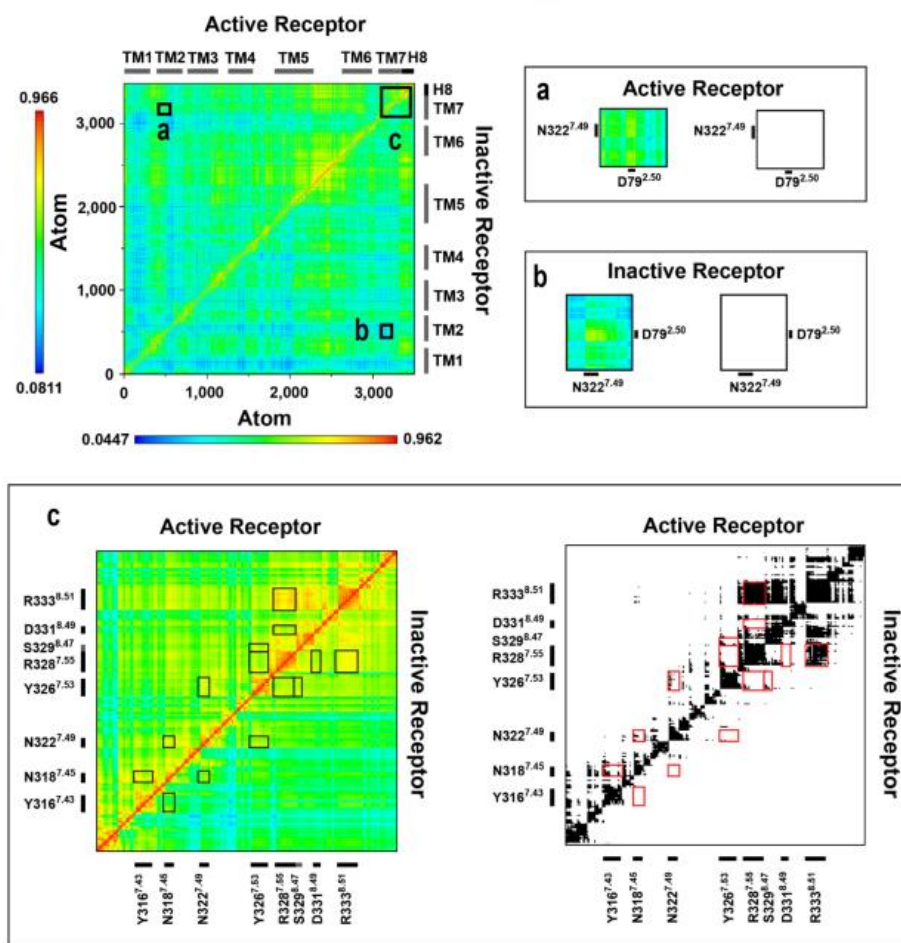
Dynamic cross-correlation matrix (DCCM) analysis (Figures 5 and S14) of the transmembrane domain and the extra- and intracellular loops indicated, that similar to that observed for the MOP receptor previously [27], the orthosteric binding pocket and the G protein-binding interface of the  $\beta_2$ AR are connected through a channel of polar amino acids which are engaged in concerted motions in the active, or constitutively active states of the receptor (Figures 6 and S15). However, there are several differences between the channel residues of the MOP and the  $\beta_2$ AR. The first difference is that, unlike in the case of

MOP, the residues of the DRY motif are not involved in correlated motions. In the active  $G_i$  protein-bound MOP, R165<sup>3,50</sup> of the DRY and D340<sup>8,47</sup> of H8 were frequently connected through a salt bridge and, consequently, their motions were in intense correlation. In  $\beta_2$ AR no such salt bridge could be formed by the analogous S329<sup>8,47</sup> of H8 and the occurrence of H-bonds was also very low, which may provide an explanation for the missing involvement of the DRY motif (See further discussion below). Nevertheless, S329<sup>8,47</sup> was found to move in accord with the NPxxY motif in  $\beta_2$ AR, but the occurrence of that connection was also not as pronounced as in the case of MOP. Instead, D331<sup>8,49</sup> and R333<sup>8,51</sup> residues of the H8 showed a high degree of correlation. R333<sup>8,51</sup> is conserved (Figure 6) and was suggested previously to be important for the G protein-coupling of the adenosine  $A_{2B}$  receptor [38]. While such occasional similarities with other class A GPCRs may support the importance of H8 residues, their increased variability compared to the other residues of the channel could also be associated with G protein specificity. A further difference observed between the MOP and  $\beta_2$ AR is that the allosteric  $Na^+$  binding site (D79<sup>2,50</sup>) is less intensely involved in the correlated motions of channel residues of  $\beta_2$ AR. However, this correlation was present during the reference simulation when epinephrine was mildly restrained to the orthosteric binding pocket, suggesting that the presence of a strongly bound, correctly oriented ligand initiates coupling of D79<sup>2,50</sup> to the signaling cascade (Figure S15). An interesting, specific feature of the polar signaling channel of  $\beta_2$ AR is that it could be subdivided into two segments, suggested by the results of NPxxY disposition and DCCM analysis. The first segment spans the orthosteric (Y316<sup>7,43</sup>) and allosteric (D79<sup>2,50</sup>) binding pockets, N318<sup>7,45</sup> and the NPxxY motif (N322<sup>7,49</sup> and Y326<sup>7,53</sup>), while the second segment shares the last residue of the NPxxY motif (Y326<sup>7,53</sup>) and includes the tip of TM7 (R328<sup>7,55</sup>) and the three H8 residues (S329<sup>8,47</sup>, D331<sup>8,49</sup>, and R333<sup>8,51</sup>) (Figures 6 and S15). The rationale behind this subdivision is provided by NPxxY disposition data as was already mentioned earlier in this report. The relatively large, approximately 0.4 nm (RMSD) disposition of the NPxxY motif resulted in intense concerted motions in the above-mentioned second segment of the polar signaling channel. Such movements were observed in inactive states and in the absence of ligand, whereas the full sequence of correlated motions was incomplete in those systems. This suggests that the elevated concerted dynamics of the second segment could also be associated with the constitutive activity of  $\beta_2$ AR. This presumption was supported by that correlated motions of this second segment were decoupled upon  $\beta$ -arrestin-2 binding and/or if epinephrine was bound in the wrong relative orientation, stabilizing a conformational state which is inappropriate for signaling. Further support is provided by results of a previous extensive study, where an intermediate structure was identified in the absence of ligands, which may represent a receptor conformation that facilitates  $G_s$  protein insertion and suggests that the activation process, in terms of structural changes, starts at the intracellular side of the receptor. The role of the NPxxY motif in the formation of such intermediate was also emphasized [17].

## 2.5. Intramolecular Interactions

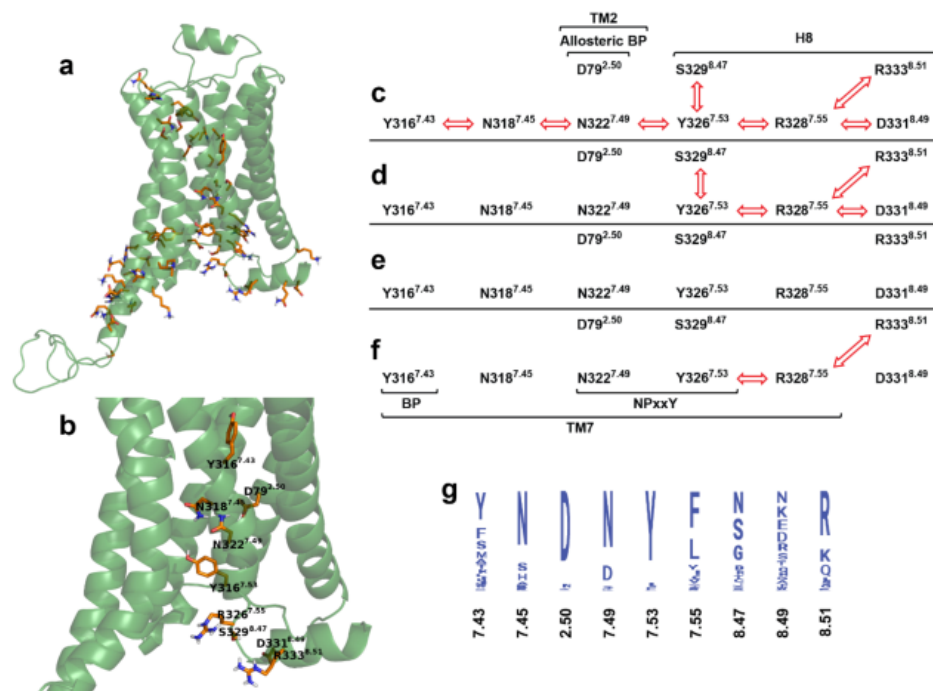
The frequencies of intramolecular salt bridges and H-bonds, previously proposed to be relevant for class A GPCR activation, are summarized in Table 2. As opposed to observations taken previously for the MOP [27], intramolecular interactions between the DRY motif and H8 were missing from all  $\beta_2$ AR systems examined here. Salt bridge formation is not facilitated between R131<sup>3,50</sup> and S329<sup>8,47</sup> and no other potential, proximal partners were found in H8 that could participate in the formation of a salt bridge analogous to that between R165<sup>3,50</sup> of the DRY and D340<sup>8,47</sup> of H8 in the active,  $G_i$  protein-bound MOP. No stable H-bond formation was indicated between the DRY motif and H8 either. This specific interaction was first described in our previous report [27] as it was not present in the high-resolution experimental structures of the MOP. Since this interaction is not facilitated in  $\beta_2$ AR, it is most likely a specific property of the MOP. Considerable frequencies of salt bridges and H-bonds were observed between the neighboring D130<sup>3,49</sup> and R131<sup>3,50</sup> of the DRY motif in the active states, although high-resolution experimental structures indicated

coincidentally that this interaction is only present in inactive states and absent in active receptors [5,7,8,11–13]. Formation of the “ionic lock” between the DRY motif (R131<sup>3.50</sup>) and TM6 (E268<sup>6.30</sup>) was also not observed, neither in the active nor in the inactive states. According to earlier proposals, this interaction acts as a constraint in the inactive state and gets disrupted upon receptor activation, followed by the release and disposition of TM6 [21]. Mutations affecting the participants of this ionic lock resulted in the elevated constitutional activity of the  $\beta_2$ AR [39], however, the presence of this ionic lock was not corroborated by the crystallographic structures of this receptor [5,12]. The absence of ionic lock interactions in these structures was attributed to residual basal activity present in the crystalline state [12]. Formation of this ionic lock was observed previously in MD simulations of receptor deactivation, but in the absence of intracellular signaling proteins and at significantly longer timescales [17].



**Figure 5.** Dynamic cross-correlation matrices of the  $G_s$  protein-bound  $\beta_2$ AR in active and inactive states. Panels (a–c) are magnified views of regions of amino acid residues of interest. Black and white panels show correlations above the threshold of 0.65 MI.





**Figure 6.** The polar signaling channel of the  $\beta_2$ AR identified by dynamic cross-correlation analysis. (a) Polar amino acids of which motions are correlated in the  $G_s$  protein-bound active state. (b) Polar amino acids of which motions are correlated and connecting the orthosteric binding pocket to the G protein-binding interface. Non-polar hydrogens are omitted for clarity. (c) Active  $\beta_2$ AR- $G_s$  protein-epinephrine complex, 1st replica; (d) inactive  $\beta_2$ AR- $G_s$  protein-epinephrine complex; (e) active  $\beta_2$ AR- $\beta$ -arrestin-2-epinephrine complex; (f) inactive  $\beta_2$ AR- $\beta$ -arrestin-2-epinephrine complex. Red arrows indicate correlated motions of the respective amino acids. (g) Degree of conservation of polar signaling channel residues of human class A GPCRs.

In agreement with the experimental structure [7], the systematic presence of H-bonds between D130<sup>3.49</sup> of the DRY motif and Y141 of ICL2 was indicated in simulations of the active state  $\beta_2$ AR. In inactive states, D130<sup>3.49</sup> of the DRY motif was found to interact with S143 and L144. This is in agreement with the results of the secondary structure analysis of ICL2 and supports the discussion above, regarding the role of this loop in the activation mechanism. No considerable trends were observed between the different receptor states and the frequencies of DRY-TM5 [40], CWxP-TM7 [41], and D113<sup>3.32</sup>-Y316<sup>7.43</sup> interactions [40] within the time frame of simulations.

**Table 2.** Frequency of intramolecular salt bridges and H-bonds expressed as percentages of the total conformational ensemble, generated by MD simulations.

Interactions	Residues Involved	Epinephrine-Bound					Ligand-Free								
		Active State			Inactive State		Active State		Inactive State						
		G <sub>s</sub> Protein Complex	β-Arrestin-2	G <sub>s</sub> Protein Complex	β-Arrestin-2	G <sub>s</sub> Protein Complex	β-Arrestin-2	G <sub>s</sub> Protein Complex	β-Arrestin-2	G <sub>s</sub> Protein Complex					
		1	2	3	Restrained										
Salt Bridge															
Intra-DRY	D130 <sup>3,49</sup> ; R131 <sup>3,50</sup>	26.5	5.7	6.0	0	65.1	40.4	35.5	40.5	10.7	51.2	41.3			
H-bonds															
DRY-H8	R131 <sup>3,50</sup> ; S329 <sup>8,47</sup>	0.1	0	0.3	0	0.1	0	0	0.1	0.1	0.0	0.0			
BP	D113 <sup>3,32</sup> ; Y316 <sup>7,43</sup>	14.0	80.6	82.1	95.3	28.3	65.9	69.2	96.1	9.9	0.1	14.6			
intra-DRY	D130 <sup>3,49</sup> ; R131 <sup>3,50</sup>	65.9	64.5	17.9	16.8	35.7	80.0	72.9	82.3	30.5	99.8	83.5			
DRY-ICL2	D130 <sup>3,49</sup> ; Y141 <sup>ICL2</sup>	99.3	97.5	92.4	98.0	99.8	0	0	98.1	86.2	0	0			
	D130 <sup>3,49</sup> ; S143 <sup>ICL2</sup> ; L144 <sup>ICL2</sup>	0	0	0	0	0	99.5	90.8	0	0	0	37.3			
DRY-TM5	R131 <sup>3,50</sup> ; Y219 <sup>5,58</sup>	3.9	10.5	4.6	45.2	30.9	0	0	5.7	3.3	0	0			
DRY-TM6	R131 <sup>3,50</sup> ; E268 <sup>6,30</sup>	0	0	0	0	0	0	0	0	0	0	0			
CWxP-TM7	C285 <sup>6,47</sup> ; W286 <sup>6,48</sup> ; N318 <sup>7,45</sup>	1.8	3.4	45.1	4.1	5.3	11.7	4.3	17.6	0.5	53.7	3.8			

BP = orthosteric binding pocket of the β<sub>2</sub>AR; Ball et al. numbering of residues is indicated in superscript

## 2.6. Intermolecular Interactions

The results of analyses of intermolecular interactions are summarized in Table 3. Similar to that observed previously for the MOP receptor [27], the interaction between the ligand and the anchor residue of the binding pocket (D113<sup>3,32</sup>) was strongest in the active, G<sub>s</sub> protein-bound  $\beta_2$ AR. The difference between active, inactive, G<sub>s</sub> protein, and  $\beta$ -arrestin-2 bound states was not as outstanding as in the case of MOP. This observed trend is also in slight contradiction with the fact that ligand disposition also occurred in G<sub>s</sub> protein-bound active states. Analysis of  $\beta_2$ AR–G<sub>s</sub> protein interactions demonstrated that contacts between ICL1 and helix 5 of the  $\alpha$  subunit of the G<sub>s</sub> protein (H5G $\alpha$ ) are negligible in ligand-bound receptors, regardless of the activation state. Medium frequency was, however, observed in ligand-free states. The ICL2–H5G $\alpha$  contact was expected to be the most specific among the contacts between the G<sub>s</sub> protein and the  $\beta_2$ AR [7], based on previous results [27] and the above presented secondary structure and intramolecular H-bond analysis. Even so, this contact was found to be very weak during simulations of the active state, epinephrine and G<sub>s</sub> protein-bound  $\beta_2$ AR. Higher frequencies of H-bonds between ICL2 and H5G $\alpha$  were observed in inactive states and the absence of epinephrine. This may suggest that the loss of interaction between ICL2 and H5G $\alpha$  in agonist-bound active states indicates the initiation of G<sub>s</sub> protein dissociation, although longer simulations would be needed to confirm this assumption. The frequency of ICL3–H5G $\alpha$  H-bonds was found to be high in all simulation setups, therefore, a specific role of this contact cannot be deduced from the results presented here. Differences were found in the interactions between  $\beta$ -arrestin-2 and  $\beta_2$ AR, depending on the presence of epinephrine. In the active ligand-bound state the finger loop (FL) of  $\beta$ -arrestin-2 was found to be in frequent contact with H8, ICL1, and ICL3, whereas in the active ligand-free state FL was in stronger contact with ICL2 at the expense of contacts with ICL3. In the inactive states, a higher preference of FL towards ICL1 was observed, regardless of the presence of epinephrine. No contacts were found between ICL3 of the agonist-bound  $\beta_1$  adrenergic receptor ( $\beta_1$ AR) and the finger loop of  $\beta$ -arrestin-1 in a recent cryo-electronmicroscopic (cryo-EM) structure of this molecular complex [42]. However, a parallel study of the neurotensin receptor 1 (NTSR1)– $\beta$ -arrestin-1 complex revealed, that the interface between  $\beta$ -arrestin-1 and NTSR1, including the finger loop, is highly dynamic and the relative orientations captured by the cryo-EM structure are likely to represent one of many conformational states [43]. A specific contact between the C-loop of  $\beta$ -arrestin-2 (CL) and the unfolded ICL2 was indicated in the inactive ligand-bound receptor. In the inactive ligand-free receptor ICL2 was rather in contact with the middle loop of  $\beta$ -arrestin-2 (ML), but with significantly lower frequency. The overall frequency of interactions was highest for inactive, epinephrine-bound  $\beta_2$ AR suggesting that it is the most preferred for  $\beta$ -arrestin-2 binding. However, taking into account that the ligand dissociated during the corresponding simulation, such an assumption cannot be taken. The second highest H-bond frequency between  $\beta_2$ AR and  $\beta$ -arrestin-2 was observed for the active, epinephrine-bound state. This latter apparent preference is corroborated by experimental data reporting the visual arrestin-bound [44] and active, G<sub>T</sub> protein-bound structures of rhodopsin [45], which were almost identical. On the other hand, the cryo-EM structure of  $\beta_1$ AR and  $\beta$ -arrestin-1 demonstrated that this receptor adopts an intermediate state with regard to the disposition of TM6 when bound by  $\beta$ -arrestin-1 [42].

**Table 3.** Frequency of intermolecular salt bridges and H-bonds expressed as percentages of the total conformational ensemble, generated by MD simulations.

Interactions	Residues Involved, Respectively	Epinephrine-Bound					Ligand-Free				
		Active State			Inactive State		Active State		Inactive State		
		G <sub>s</sub> Protein Complex	β-Arrestin-2	G <sub>s</sub> Protein Complex	β-Arrestin-2	G <sub>s</sub> Protein Complex	β-Arrestin-2	G <sub>s</sub> Protein Complex	β-Arrestin-2		
		1	2	3	Restained						
Salt Bridge											
BP-epi	D113 <sup>3,32</sup> ; epi	67.5	51.0	4.8	52.7	36.0	29.3	60.1	-	-	-
H-bonds											
BP-epi	D113 <sup>3,32</sup> ; epi	96.7	96.0	8.3	99.3	55.5	53.8	93.3	-	-	-
H8-H5Gα	S329-L340; T369-L394	0.0	7.3	4.0	0.0	-	0.0	-	0.0	0.0	-
IC11-H5Gα	F61-T66; T369-L394	0.2	0.1	0.0	0.2	-	0.6	-	30.3	40.0	-
IC12-H5Gα	S137-T146; T369-L394	0.1	0.1	17.9	0.5	-	9.9	-	5.1	31.5	-
IC13-H5Gα	E237-K267; T369-L394	90.5	92.5	97.5	96.3	-	77.1	-	76.2	76.7	-
H8-FL	S329-L340; G65-K78	-	-	-	-	33.4	-	0.0	-	-	0.0
IC11-FL	F61-T66; G65-K78	-	-	-	-	25.6	-	58.8	-	22.9	49.3
IC12-FL	S137-T146; G65-K78	-	-	-	-	8.7	-	30.4	-	77.8	35.6
IC13-FL	E237-K267; G65-K78	-	-	-	-	85.5	-	71.7	-	14.8	14.0
H8-ML	S329-L340; P132-A140	-	-	-	-	0.0	-	0.0	-	0.0	0.0
IC11-ML	F61-T66; P132-A140	-	-	-	-	0.9	-	0.1	-	0.0	0.0
IC12-ML	S137-T146; P132-A140	-	-	-	-	7.1	-	4.2	-	9.3	30.3
IC13-ML	E237-K267; P132-A140	-	-	-	-	0.0	-	0.0	-	0	0
H8-CL	S329-L340; V307-G317	-	-	-	-	0.0	-	0.0	-	0.0	0.0
IC11-CL	F61-T66; V307-G317	-	-	-	-	0.0	-	0.0	-	0.0	0.0
IC12-CL	S137-T146; V307-G317	-	-	-	-	4.6	-	58.4	-	0.2	8.9
IC13-CL	E237-K267; V307-G317	-	-	-	-	0.0	-	0.0	-	4.0	0

BP = orthosteric binding pocket of the β<sub>2</sub>AR; epi = epinephrine; H5Gα = helix 5 of the G<sub>s</sub> protein α subunit; FL/ML/CL = finger loop/middle loop/C loop of β-arrestin-2. Ballosteros-Weinstein numbering of residues is indicated in superscript.



### 3. Methods

#### 3.1. System Building

The sequence of the human  $\beta_2$ AR (UniProtKB-P07550-ADRB2) was obtained from the UniProt database (<http://www.uniprot.org>, accessed on 23 August 2021). All X-ray crystallographic structures used in this study were downloaded from the Brookhaven Protein Data Bank (<http://www.rcsb.org>, accessed on 23 August 2021). The active and inactive state  $\beta_2$ AR (pdb codes: 3SN6 and 2RH1, respectively) [5,7],  $\beta$ -arrestin-2 (pdb code: 3P2D) [44], the  $G_s\alpha$  protein (pdb code: 1AZT) [46] and the  $\beta_2$ AR-bound epinephrine (pdb code: 4LDO) [8] were used as starting structures for MD simulations in this study. All crystallographic chaperones and fusion proteins were removed from the corresponding structures. The  $\alpha$  subunit of the  $G_s$  protein complex was missing from the crystallographic structure of the active  $\beta_2$ AR (pdb code: 3SN6) [7], therefore, it was supplemented from an independent crystallographic structure of that subunit (pdb code: 1AZT) [46], together with the bound GDP. Epinephrine was inserted in the binding pocket of the receptors in a protonated form. The third intracellular loop (ICL3, E237-K267) of  $\beta_2$ AR, missing from the crystallographic structures was modeled using the Modeller ver. 9.20 software [47] and the missing residues in the second extracellular loop (ECL2, A176-H178) were retrieved using the Swiss-PdbViewer ver. 4.10 program [48]. The missing *N*- and C-terminal domains (M1-D29 and C341-L413, respectively) of  $\beta_2$ AR were modeled by performing 10 ns folding simulations using the GROMACS ver. 2018.3 program package [49], following a previously described protocol [27] and attached to the TM domain of the receptor manually.

The CHARMM-GUI [50] web-based platform was used to include post-translational modifications of  $\beta_2$ AR and to insert the receptor in a solvated membrane bilayer. Residues N6, N15, and N187 were glycosylated [51] and residue C341 was palmitoylated in both the active and inactive state receptors [52], whereas phosphorylations at the C-terminal domain (S355, S356, and S364) [53,54] were included only for  $\beta$ -arrestin-2 bound systems. Complex type glycans were used for glycosylation of the *N*-terminal domain, consisting of a common core (Man $\alpha$ 1-3 (Man $\alpha$ 1-6) Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc $\beta$ 1-N) and sialic acid (*N*-acetylneuraminic acid).

The receptor complexes were inserted in a previously introduced and examined explicit caveolar membrane bilayer [27] using the membrane builder tool of CHARMM-GUI [50]. The bilayer consisted of the following components: cholesterol (CHL-32.8%), 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC-14.9%), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE-27.8%), 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine (POPS-3.6%), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoinositol (POPI2-6.0%), palmitoyl-sphingomyelin (PSM-9.9%), and monosialodihexosylganglioside (GM3-5.0%) [55]. The asymmetric lower and upper leaflet compositions of the membrane were set according to the recent literature data [56]. Membrane-inserted  $\beta_2$ AR complexes were then solvated with explicit TIP3P water molecules in a hexagonal-shaped periodic box, and  $\text{Na}^+$  and  $\text{Cl}^-$  ions (0.15 M) were added to neutralize the system and to attain physiological ionic strength. System coordinates and topologies were generated in GROMACS format and CHARMM36 all-atom force field parameters were assigned to all system components [57].

#### 3.2. MD Simulations

All energy minimizations and MD simulations were performed using the GROMACS 2018.3 program package [49]. Initially, all complex systems were subjected to 5000 steps steepest descent, and then 5000 steps conjugate gradient energy minimization. The convergence criteria were set to  $1000 \text{ kJ mol}^{-1} \text{ nm}^{-1}$  for both minimization steps. Minimized systems were then thoroughly equilibrated following a six-step protocol, supplied by CHARMM-GUI. According to the protocol, two consecutive MD simulations were executed at 303.15 K temperature in the canonical (NVT) ensemble, then four further simulations in the isobaric (NPT) ensemble at 303.15 K temperature and 1 bar pressure. Positional restraints were applied on the heavy atoms of the proteins and membrane constituents which were decreased gradually throughout the steps of the equilibration protocol. The

first three equilibration MD runs were 25 ps long and were performed in 1 fs time steps. The following two were continued for 100 ps in 2 fs time steps. The last equilibration step was extended to 50 ns and was executed in 2 fs time steps. The chemical bonds were constrained to their correct lengths using the LINCS algorithm. The v-rescale algorithm [58] with a coupling constant of 1 ps was applied for temperature control. The pressure was regulated using Berendsen (semi-isotropic) pressure coupling [59] with a 5 ps coupling constant and  $4.5 \times 10^{-5} \text{ bar}^{-1}$  isothermal compressibility. The Particle Mesh Ewald (PME) method was used to calculate energy contributions from electrostatic interactions. Van der Waals interactions were calculated using a twin-range cutoff. All cut-off values were set to 1.2 nm.

Eleven independent production simulations were performed at 310 K in the NPT ensemble, with other parameters similar as above. Each simulation was 1  $\mu\text{s}$  long and included the active and inactive state  $\beta_2\text{AR}$ , complexed either with the heterotrimeric  $G_s$  protein or  $\beta$ -arrestin-2, in the presence or the absence of orthosterically-bound epinephrine. The simulation of the active  $\beta_2\text{AR}$  bound to the  $G_s$  protein and epinephrine was performed in three replicates. An additional reference simulation was performed for this system in which a mild positional restraint ( $200 \text{ kJ mol}^{-1}$  on heavy atoms) was applied for epinephrine to prevent spontaneous ejection from the binding pocket. None of the other production simulations have included any restraints. System coordinates were saved after every 5000 steps providing trajectories with 100,000 snapshots.

### 3.3. MD Trajectory Analysis

MD trajectories were examined using the analysis suite of the GROMACS 2018.3 package [49] to check the integrity of simulation systems and to observe protein conformational changes and minute details of different activation states of the  $\beta_2\text{AR}$ . Specific analyses were performed to compare results to those obtained for the MOP in a recent study [27].

Root mean square deviation (RMSD) of protein backbone atoms were calculated and compared between different systems to assess the structural stability of the macromolecular complexes and to identify significant displacements of key structural components as a function of time. The *gmx rms* program was used for RMSD calculation. The *gmx helix* utility was used to calculate properties of TM helices and to measure their deviation from the ideal  $\alpha$ -helical structure. Secondary structures of intracellular loops (ICL1: F61-T66, ICL2: S137-T146, ICL3: E237-K267), and the cytosolic helix (H8: S329-L340) was assigned using the DSSP method [60]. The frequency of intra- and intermolecular H-bonds was calculated using the *gmx hbond* program. H-bonds were assigned within 0.35 nm donor-acceptor distance and below 30.0 degrees of donor-hydrogen-acceptor angle. The presence of salt bridges between acidic and basic functional groups was assigned with 0.40 nm distance and 90.0 degrees angle cutoff values. Distances and angles between these groups were calculated with the *gmx distance* and *gmx gangle* programs, respectively. The *gmx mindist* program was used to observe  $\text{Na}^+$  penetration into the TM domain and to calculate the minimum distance between periodic replicas of *N*- and *C*-terminal domains. Dynamic cross-correlation matrix (DCCM) analysis, available in an earlier version of the GROMACS suite (*g\_correlation*, ver. 3.3) [61], was performed to examine the dynamic motions of amino acid side chains in the TM domain and connecting loops. DCCM matrices were converted to heat map images using the *gmx xpm2ps* utility and analyzed using the Gimp ver. 2.8 software. Color intensities corresponding to 0.65 MI (mutual information) and the participation of at least four atoms from each amino acid side chain were set as the threshold of correlation assignment. The Pymol ver. 2.1.0 and VMD 1.9.4a12 software were used for molecular visualization. The Xmgrace ver. 5.1.25. program was used to prepare graphs.

### 3.4. Sequence Alignment and Conservation Analysis

Downloaded from the UniProt database in FASTA format were 267 sequences of class A human GPCRs (without orphan and olfactory receptors). Multiple-sequence alignment was carried out using the Clustal Omega program [62] and analyzed with the Jalview 2.10.5 software [63]. The ADRB2\_HUMAN (P07550) was set as the reference sequence for conservation analysis.

## 4. Conclusions

Results presented here for the  $\beta_2$ AR provide support for our previously proposed hypothesis [27] and justify its extension to other class A GPCRs. The above results also suggest that the previously proposed potential contribution of the electrostatic balance in the TM domain is warranted for detailed, quantitative examination. While several previous assumptions drawn from results gathered for the MOP receptor were reinforced, some had to be adjusted to account for the differences observed in the case of  $\beta_2$ AR. The general features of GPCR activation proposed here and previously by others [16,20,25,26] and the receptor-specific, characteristic details may provide alternative opportunities for the discovery of a new class of GPCR drugs. The extended perspective of the activation mechanism, if further pursued, may provide a more in-depth explanation for ligand-induced effects in multiple functional states and could help to identify and quantitatively assess specific physico-chemical properties of GPCR ligands that furnish different functional properties.

**Supplementary Materials:** The Supplementary Materials are available online at <https://www.mdpi.com/article/10.3390/ijms221910423/s1>.

**Author Contributions:** Conceptualization, A.B.; methodology, A.M. and A.B.; formal analysis, A.M., A.S.; investigation, A.M. and A.S.; writing—original draft preparation, A.M. and A.B.; writing—review and editing, A.M., A.S. and A.B.; visualization, A.M. and A.B.; supervision, A.B.; project administration, A.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** Scholarship for A.S. was provided by the ‘Stipendium Hungaricum’ program of the Hungarian Ministry of Foreign Affairs and Trade and the Tempus Public Foundation.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All data contained within the article or Supplementary Materials are available upon request.

**Acknowledgments:** Computing resources were provided by the Government Agency of Information Technology Development, Hungary.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Venter, J.C.; Adams, M.D.; Myers, E.W.; Li, P.W.; Mural, R.J.; Sutton, G.G.; Smith, H.O.; Yandell, M.; Evans, C.A.; Holt, R.A.; et al. The sequence of the human genome. *Science* **2001**, *291*, 1304–1351. [CrossRef] [PubMed]
- Hauser, A.; Attwood, M.M.; Rask-Andersen, M.; Schiöth, H.B.; Gloriam, D. Trends in GPCR drug discovery: New agents, targets and indications. *Nat. Rev. Drug Discov.* **2017**, *16*, 829–842. [CrossRef] [PubMed]
- Congreve, M.; de Graaf, C.; Swain, N.A.; Tate, C.G. Impact of GPCR Structures on Drug Discovery. *Cell* **2020**, *181*, 81–91. [CrossRef] [PubMed]
- Brandt, D.R.; Asano, T.; Pedersen, S.E.; Ross, E.M. Reconstitution of catecholamine-stimulated guanosinetriphosphatase activity. *Biochemistry* **1983**, *22*, 4357–4362. [CrossRef]
- Cherezov, V.; Rosenbaum, D.M.; Hanson, M.A.; Rasmussen, S.G.; Thian, F.S.; Kobilka, T.S.; Choi, H.J.; Kuhn, P.; Weis, W.I.; Kobilka, B.K.; et al. High-resolution crystal structure of an engineered human  $\beta_2$ -adrenergic G protein-coupled receptor. *Science* **2007**, *318*, 1258–1265. [CrossRef] [PubMed]
- Hanson, M.A.; Cherezov, V.; Griffith, M.T.; Roth, C.B.; Jaakola, V.P.; Chien, E.Y.; Velasquez, J.; Kuhn, P.; Stevens, R.C. A specific cholesterol binding site is established by the 2.8 Å structure of the human  $\beta_2$ -adrenergic receptor. *Structure* **2008**, *16*, 897–905. [CrossRef]



7. Rasmussen, S.G.F.; DeVree, B.T.; Zou, Y.; Kruse, A.C.; Chung, K.Y.; Kobilka, T.S.; Thian, F.S.; Chae, P.S.; Pardon, E.; Calinski, D.; et al. Crystal structure of the beta2-adrenergic receptor-Gs protein complex. *Nature* **2011**, *477*, 549–555. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Ring, A.; Manglik, A.; Kruse, A.C.; Enos, M.D.; Weis, W.; Garcia, K.C.; Kobilka, B.K. Adrenaline-activated structure of  $\beta$ 2-adrenoceptor stabilized by an engineered nanobody. *Nature* **2013**, *502*, 575–579. [\[CrossRef\]](#)
9. Weichert, D.; Kruse, A.C.; Manglik, A.; Hiller, C.; Zhang, C.; Hübner, H.; Kobilka, B.K.; Gmeiner, P. Covalent agonists for studying G protein-coupled receptor activation. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 10744–10748. [\[CrossRef\]](#)
10. Zou, Y.; Weis, W.I.; Kobilka, B.K. N-Terminal T4 Lysozyme Fusion Facilitates Crystallization of a G Protein Coupled Receptor. *PLoS ONE* **2012**, *7*, e46039. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Wacker, D.; Fenalti, G.; Brown, M.A.; Katritch, V.; Abagyan, R.; Cherezov, V.; Stevens, R.C. Conserved binding mode of human beta2-adrenergic receptor inverse agonists and antagonist revealed by X-ray crystallography. *J. Am. Chem. Soc.* **2010**, *132*, 11443–11445. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Rasmussen, S.G.F.; Choi, H.J.; Rosenbaum, D.M.; Kobilka, T.S.; Thian, F.S.; Edwards, P.C.; Burghammer, M.; Ratnala, V.R.; Sanishvili, R.; Fischetti, R.F.; et al. Crystal structure of the human beta2-adrenergic G-protein-coupled receptor. *Nature* **2007**, *450*, 383–387. [\[CrossRef\]](#)
13. Rasmussen, S.; Choi, H.-J.; Fung, J.J.; Pardon, E.; Casarosa, P.; Chae, P.S.; DeVree, B.; Rosenbaum, D.M.; Thian, F.S.; Kobilka, T.S.; et al. Structure of a nanobody-stabilized active state of the  $\beta$ 2 adrenoceptor. *Nature* **2011**, *469*, 175–180. [\[CrossRef\]](#)
14. Rosenbaum, D.M.; Zhang, C.; Lyons, J.; Holl, R.; Aragao, D.; Arlow, D.H.; Rasmussen, S.; Choi, H.-J.; DeVree, B.; Sunahara, R.K.; et al. Structure and function of an irreversible agonist- $\beta$ 2 adrenoceptor complex. *Nature* **2011**, *469*, 236–240. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Munk, C.; Mutt, E.; Isberg, V.; Nikolajsen, L.F.; Bibbe, J.M.; Flock, T.; Hanson, M.A.; Stevens, R.C.; Deupi, X.; Gloriam, D.E. An online resource for GPCR structure determination and analysis. *Nat. Methods* **2019**, *16*, 151–162. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Latorraca, N.R.; Venkatakrishnan, A.J.; Dror, R.O. GPCR Dynamics: Structures in Motion. *Chem. Rev.* **2016**, *117*, 139–155. [\[CrossRef\]](#)
17. Dror, R.O.; Arlow, D.H.; Maragakis, P.; Mildorf, T.J.; Pan, A.C.; Xu, H.; Borhani, D.W.; Shaw, D.E. Activation mechanism of the beta2-adrenergic receptor. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 18684–18689. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Nygaard, R.; Zou, Y.; Dror, R.O.; Mildorf, T.J.; Arlow, D.H.; Manglik, A.; Pan, A.C.; Liu, C.W.; Fung, J.J.; Bokoch, M.P.; et al. The dynamic process of  $\beta$ (2)-adrenergic receptor activation. *Cell* **2013**, *152*, 532–542. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Dror, R.O.; Arlow, D.H.; Borhani, D.W.; Jensen, M.Ø.; Piana, S.; Shaw, D.E. Identification of two distinct inactive conformations of the beta2-adrenergic receptor reconciles structural and biochemical observations. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 4689–4694. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Marino, K.A.; Shang, Y.; Filizola, M. Insights into the function of opioid receptors from molecular dynamics simulations of available crystal structures. *Br. J. Pharmacol.* **2017**, *175*, 2834–2845. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C.; Motoshima, H.; Fox, B.; Le Trong, I.; Teller, D.C.; Okada, T.; Stenkamp, R.E.; et al. Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. *Science* **2000**, *289*, 739–745. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Stoddart, L.A.; Kellam, B.; Briddon, S.J.; Hill, S.J. Effect of a toggle switch mutation in TM6 of the human adenosine A3 receptor on G protein-dependent signalling and Gi-independent receptor internalization. *Br. J. Pharmacol.* **2014**, *171*, 3827–3844. [\[CrossRef\]](#)
23. Pert, C.B.; Pasternak, G.; Snyder, S.H. Opiate Agonists and Antagonists Discriminated by Receptor Binding in Brain. *Science* **1973**, *182*, 1359–1361. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Liu, W.; Chun, E.; Thompson, A.A.; Chubukov, P.; Xu, F.; Katritch, V.; Han, G.W.; Roth, C.B.; Heitman, L.; Ijzerman, A.; et al. Structural Basis for Allosteric Regulation of GPCRs by Sodium Ions. *Science* **2012**, *337*, 232–236. [\[CrossRef\]](#)
25. Zhou, Q.; Yang, D.; Wu, M.; Guo, Y.; Guo, W.; Zhong, L.; Cai, X.; Dai, A.; Jang, W.; I Shakhnovich, E.; et al. Common activation mechanism of class A GPCRs. *eLife* **2019**, *8*, e50279. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Filipek, S. Molecular switches in GPCRs. *Curr. Opin. Struct. Biol.* **2019**, *55*, 114–120. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Mitra, A.; Sarkar, A.; Szabó, M.; Borics, A. Correlated Motions of Conserved Polar Motifs Lay out a Plausible Mechanism of G Protein-Coupled Receptor Activation. *Biomolecules* **2021**, *11*, 670. [\[CrossRef\]](#)
28. Gregorio, G.G.; Masureel, M.; Hilger, D.; Terry, D.S.; Juetten, M.; Zhao, H.; Zhou, Z.; Perez-Aguilar, J.M.; Hauge, M.; Mathiasen, S.; et al. Single-molecule analysis of ligand efficacy in  $\beta$ 2AR-G-protein activation. *Nature* **2017**, *547*, 68–73. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Fleetwood, O.; Matricon, P.; Carlsson, J.; Delemotte, L. Energy Landscapes Reveal Agonist Control of G Protein-Coupled Receptor Activation via Microswitches. *Biochemistry* **2020**, *59*, 880–891. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Yuan, S.; Vogel, H.; Filipek, S. The Role of Water and Sodium Ions in the Activation of the  $\mu$ -Opioid Receptor. *Angew. Chem. Int. Ed.* **2013**, *52*, 10112–10115. [\[CrossRef\]](#)
31. Shang, Y.; Lerouzic, V.; Schneider, S.; Bisignano, P.; Pasternak, G.W.; Filizola, M. Mechanistic Insights into the Allosteric Modulation of Opioid Receptors by Sodium Ions. *Biochemistry* **2014**, *53*, 5140–5149. [\[CrossRef\]](#)
32. Selent, J.; Sanz, F.; Pastor, M.; De Fabritiis, G. Induced Effects of Sodium Ions on Dopaminergic G-Protein Coupled Receptors. *PLoS Comput. Biol.* **2010**, *6*, e1000884. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Hu, X.; Wang, Y.; Hunkele, A.; Provasi, D.; Pasternak, G.W.; Filizola, M. Kinetic and thermodynamic insights into sodium ion translocation through the  $\mu$ -opioid receptor from molecular dynamics and machine learning analysis. *PLoS Comput. Biol.* **2019**, *15*, e1006689. [\[CrossRef\]](#) [\[PubMed\]](#)

34. Hol, W.G. Effects of the  $\alpha$ -helix dipole upon the functioning and structure of proteins and peptides. *Adv. Biophys.* **1985**, *19*, 133–165. [\[CrossRef\]](#)
35. Ma, X.; Hu, Y.; Batebi, H.; Heng, J.; Xu, J.; Liu, X.; Niu, X.; Li, H.; Hildebrand, P.W.; Jin, C.; et al. Analysis of  $\beta$ 2AR-Gs and  $\beta$ 2AR-Gi complex formation by NMR spectroscopy. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 23096–23105. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Zhang, Y.; Yang, F.; Ling, S.; Lv, P.; Zhou, Y.; Fang, W.; Sun, W.; Zhang, L.; Shi, P.; Tian, C. Single-particle cryo-EM structural studies of the  $\beta$ 2AR-Gs complex bound with a full agonist formoterol. *Cell Discov.* **2020**, *6*, 1–5. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Yang, F.; Ling, S.; Zhou, Y.; Zhang, Y.; Lv, P.; Liu, S.; Fang, W.; Sun, W.; Hu, L.A.; Zhang, L.; et al. Different conformational responses of the  $\beta$ 2-adrenergic receptor-Gs complex upon binding of the partial agonist salbutamol or the full agonist isoprenaline. *Natl. Sci. Rev.* **2020**, *8*, nwaa284. [\[CrossRef\]](#)
38. Liu, R.; Nahon, D.; le Roy, B.; Lenselink, E.B.; Ijzerman, A.P. Scanning Mutagenesis in a Yeast System Delineates the Role of the NPxxY(x)5,6 F Motif and Helix 8 of the Adenosine A2B Receptor in G Protein Coupling. *Biochem. Pharmacol.* **2015**, *95*, 290–300. [\[CrossRef\]](#)
39. Ballesteros, J.A.; Jensen, A.D.; Liapakis, G.; Rasmussen, S.; Shi, L.; Gether, U.; Javitch, J. Activation of the  $\beta$ 2-Adrenergic Receptor Involves Disruption of an Ionic Lock between the Cytoplasmic Ends of Transmembrane Segments 3 and 6. *J. Biol. Chem.* **2001**, *276*, 29171–29177. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Huang, W.; Manglik, A.; Venkatakrisnan, A.J.; Laeremans, T.; Feinberg, E.; Sanborn, A.L.; Kato, H.; Livingston, K.E.; Thorsen, T.S.; Kling, R.C.; et al. Structural insights into  $\mu$ -opioid receptor activation. *Nature* **2015**, *524*, 315–321. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Jongejan, A.; Bruysters, M.; Ballesteros, J.A.; Haaksma, E.; Bakker, R.A.; Pardo, L.; Leurs, R. Linking Agonist Binding to Histamine H1 Receptor Activation. *Nat. Chem. Biol.* **2005**, *1*, 98–103. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Lee, Y.; Wame, T.; Nehmé, R.; Pandey, S.; Dwivedi-Agnihotri, H.; Chaturvedi, M.; Edwards, P.C.; García-Nafria, J.; Leslie, A.G.W.; Shukla, A.K.; et al. Molecular basis of  $\beta$ -arrestin coupling to formoterol-bound  $\beta$ 1-adrenoceptor. *Nature* **2020**, *583*, 862–866. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Yin, W.; Li, Z.; Jin, M.; Yin, Y.-L.; de Waal, P.W.; Pal, K.; Yin, Y.; Gao, X.; He, Y.; Gao, J.; et al. A complex structure of arrestin-2 bound to a G protein-coupled receptor. *Cell Res.* **2019**, *29*, 971–983. [\[CrossRef\]](#)
44. Zhan, X.; Gimenez, L.E.D.; Gurevich, V.V.; Spiller, B.W. Crystal Structure of Arrestin-3 Reveals the Basis of the Difference in Receptor Binding Between Two Non-visual Subtypes. *J. Mol. Biol.* **2011**, *406*, 467–478. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Gao, Y.; Hu, H.; Ramachandran, S.; Erickson, J.W.; Cerione, R.A.; Skiniotis, G. Structures of the Rhodopsin-Transducin Complex: Insights into G-Protein Activation. *Mol. Cell* **2019**, *75*, 781–790. [\[CrossRef\]](#)
46. Sunahara, R.K.; Tesmer, J.J.G.; Gilman, A.G.; Sprang, S.R. Crystal Structure of the Adenylyl Cyclase Activator Gs. *Science* **1997**, *278*, 1943–1947. [\[CrossRef\]](#)
47. Fiser, A.; Do, R.K.G.; Sali, A. Modeling of loops in protein structures. *Protein Sci.* **2000**, *9*, 1753–1773. [\[CrossRef\]](#)
48. Johansson, M.U.; Zoete, V.; Michielin, O.; Guex, N. Defining and searching for structural motifs using DeepView/Swiss-PdbViewer. *BMC Bioinform.* **2012**, *13*, 173. [\[CrossRef\]](#)
49. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. [\[CrossRef\]](#)
50. Jo, S.; Kim, T.; Iyer, V.G.; Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859–1865. [\[CrossRef\]](#)
51. Miale-Perez, J.; Green, S.A.; Miller, W.E.; Liggett, S.B. A primate-dominant third glycosylation site of the  $\beta$ 2-adrenergic receptor routes receptors to degradation during agonist regulation. *J. Biol. Chem.* **2004**, *279*, 38603–38607. [\[CrossRef\]](#)
52. O'Dowd, B.F.; Hnatowich, M.; Caron, M.G.; Lefkowitz, R.J.; Bouvier, M. Palmitoylation of the human beta 2-adrenergic receptor. Mutation of Cys341 in the carboxyl tail leads to an uncoupled nonpalmitoylated form of the receptor. *J. Biol. Chem.* **1989**, *264*, 7564–7569. [\[CrossRef\]](#)
53. Zamah, A.M.; Delahunty, M.; Luttrell, L.; Lefkowitz, R.J. Protein Kinase A-mediated Phosphorylation of the  $\beta$ 2-Adrenergic Receptor Regulates Its Coupling to Gs and Gi. *J. Biol. Chem.* **2002**, *277*, 31249–31256. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Hausdorff, W.P.; Bouvier, M.; O'Dowd, B.F.; Irons, G.P.; Caron, M.G.; Lefkowitz, R.J. Phosphorylation Sites on Two Domains of the  $\beta$ 2-Adrenergic Receptor Are Involved in Distinct Pathways of Receptor Desensitization. *J. Biol. Chem.* **1989**, *264*, 12657–12665. [\[CrossRef\]](#)
55. Pike, L.J.; Han, X.; Chung, K.N.; Gross, R.W. Lipid rafts are enriched in arachidonic acid and plasmalogen ethanolamine and their composition is independent of caveolin-1 expression: A quantitative electrospray ionization/mass spectrometric analysis. *Biochemistry* **2002**, *41*, 2075–2088. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Ingólfsson, H.I.; Melo, M.N.; van Eerden, F.J.; Arnarez, C.; Lopez, C.A.; Wassenaar, T.A.; Periole, X.; de Vries, A.H.; Tieleman, D.P.; Marrink, S.J. Lipid Organization of the Plasma Membrane. *J. Am. Chem. Soc.* **2014**, *136*, 14554–14559. [\[CrossRef\]](#)
57. Huang, J.; MacKerell, A.D., Jr. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **2013**, *34*, 2135–2145. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101. [\[CrossRef\]](#)
59. Berendsen, H.J.C.; Postma, J.P.M.; Van Gunsteren, W.F.; DiNola, A.; Haak, J.R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690. [\[CrossRef\]](#)
60. Kabsch, W.; Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577–2637. [\[CrossRef\]](#)

61. Lange, O.F.; Grubmüller, H. Generalized correlation for biomolecular dynamics. *Proteins Struct. Funct. Bioinform.* **2005**, *62*, 1053–1061. [[CrossRef](#)] [[PubMed](#)]
62. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; López, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [[CrossRef](#)] [[PubMed](#)]
63. Waterhouse, A.M.; Procter, J.; Martin, D.; Clamp, M.; Barton, G.J. Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189–1191. [[CrossRef](#)] [[PubMed](#)]