Mechanistic exploration of ligand binding to ET_B receptor embedded in different membrane environments

Paulina Frolovaitė
Master's Thesis
Master's Degree Programme in Digital Health and Life Sciences
Department of Computing
Faculty of Technology
University of Turku
May 2021

Supervisors:
Dr. Parthiban Marimuthu
Dr. Outi Salo-Ahen
Dr. Martti Tolvanen

ABSTRACT

UNIVERSITY OF TURKU
Department of Computing/Faculty of Technology
FROLOVAITĖ, PAULINA: Mechanistic exploration of ligand binding to ET_B receptor embedded in different membrane environments
Master's thesis, 77 p., 2 appendix pages
Bioinformatics
May 2021

Endothelin B (ET_B) receptor belongs to a GPCR family, and it consists of seven transmembrane helices connected by three extracellular and intracellular loops. ET_B receptor's main function is vasodilation, thus the receptor is involved in diseases affecting the vasculature. Therefore, ET_B targeting drugs have been a major topic in drug discovery and design. Several drugs have been developed, with two of them being bosentan and K-8794. Bosentan is an antagonist targeting both ET_A and ET_B receptors, while K-8794 is bosentan's high affinity analog (150 nM) targeting selectively to ET_B receptor. In this current investigation, the main goal was to explore and understand the mechanistic basis of bosentan and K-8794 binding to ET_B receptor embedded in POPC, DPPC, POPE and DMPC membrane environments. To achieve this, a series of long-range molecular dynamics (MD) simulations were carried out. The trajectories from the simulation outputs were analysed for the receptor's structural stability (RMSD, RMSF, Rg, H-bond). Moreover, essential dynamics analysis based on principal component analysis (PCA), binding free energy (BFE) estimation based on MM/GBSA approach, and dihedral angle analysis for the hotspot residues were carried out. According to the RMSD analysis revealed that all the systems have reached the equilibrium state during the last 200 ns of the MD simulations. The RMSF investigation reported that the most fluctuating regions of the receptor were the extracellular and intracellular loops. The Rg analysis showed that the ligand-receptor complexes maintained their overall compactness in most of the membrane environments. The H-bond analysis suggested that the ET_B receptor residues Lys182, Lys273, and Arg343 formed relevant H-bond interactions with the two ligands and the ET_B receptor. The PCA analysis showed that the ligand-ET_B complexes were highly flexible by forming several conformations; however, the two ligands were restraining the receptor by disallowing any drastic helical movement. The BFE estimation revealed the most energy contributing (hotspot) residues – Lys182 and Trp336 for both bosentan and K-8794-ET_B complexes. The dihedral angle analysis showed that both hotspot residues maintained similar conformation for all of the ligand-bound ET_B complexes embedded in the four membrane environments.

Keywords
Endothelin B receptor, bosentan, K-8794, Molecular Dynamics Simulation, lipid membrane environment, mechanistic exploration of ligand binding, PCA, binding free energy estimation, dihedral angle analysis.

ACKNOWLEDGMENT

First and foremost, I would like to express my deepest appreciation to my supervisors Dr. Parthiban Marimuthu and Dr. Outi Salo-Ahen for an opportunity to work at their laboratory. Also, for their valuable advice given, patients that cannot be underestimated, profound belief in my abilities, and of course guidance during this project.

I would like to extend my sincere thanks to Dr. Martti Tolvanen for his immense support, patience and guidance during this project and Master studies, and for being a source of inspiration in the classes he taught.

I am also grateful to Prof. Juho Heimonen for his invaluable patience and guidance during my Master studies.

Finally, special thanks to my dear family and friends for their understanding, support and love given throughout this journey.

Thank you!

Paulina Frolovaitė,
Turku, Finland, May, 2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

A_1AR – Adenosine A_1 receptor

AH – Arterial hypertension

BFE – Binding free energy

BM – Bone marrow

CHF – Chronic heart failure

C_{ij} – Covariance matrix

CPU – Central processing units

DMPC – 1,2-dimyristoyl-sn-glycero-3-phosphocholine

DOPC – 1,2-dioleoyl-sn-glycero-3-phosphocholine

DOPE – 1,2-dioleoyl-sn-glycero-3-phosphocholine

DOPE – Discrete Optimized Protein Energy

DOPS – 1,2-dioleoyl-sn-glycero-3-phospho-L-serine

DPPC – 1,2-dipalmitoyl-sn-glycero-3-phosphocholine

E/DRY – glutamic acid/aspartic acid-arginine-tyrosine

EARTH – Endothelin A Receptor Antagonist Trial in Heart Failure

ECE – Endothelin-converting enzymes

ECL – Extracellular loop

ECL1 – Extracellular loop 1

ECL2 – Extracellular loops 2

ECL3 – Extracellular loop 3

ECs – Endothelial cells

ED – Essential Dynamics

E_{ele} – Sum of electrostatic energies

E_{lipo} – Lipophilic energy

EMA – European Medicines' Agency

EMM – OPLS molecular mechanics energies

ENABLE – Endothelin Antagonist Bosentan for Lowering Cardiac Events

eNOS – Nitric oxide synthase 3

ET – Endothelin

ET-1 – Endothelin-1

ET-2 – Endothelin-2

ET-3 – Endothelin-3

ET_A – Endothelin A

ET_B – Endothelin B

ETR – Endothelin receptor

E_{VdW} – Van der Waals energy

FDA – U.S. Food and Drug Administration

FEL – Free Energy Landscape

GaMD – Gaussian accelerated molecular dynamics

GB – Generalized Born

GNP – Nonpolar solvation term

GPCR – G protein-coupled receptors

GPU – Graphic processing unit

GSGB – VSGB solvation model for polar solvation

H-bonds – Hydrogen bonds

HEAT – Heart Failure ET(A) Receptor Blockade Trial

HPS – Hepatopulmonary syndrome

IC – Intracellular

ICL – Intracellular loop

ICL1 – Intracellular loop 1

ICL2 – Intracellular loop 2

ICL3 – Intracellular loop 3

LBM – Lipid bilayer membrane

MD – Molecular Dynamics

MDS – Molecular Dynamics Simulation

MM/GBSA – Molecular Mechanics with Generalized Born Surface Area

mT4L – T4-lysosyme

NYHA – New York Heart Association

PAH – Pulmonary arterial hypertension

PC – Phosphatidylcholine

PC1 – First principal component

PC2 – Second principal component

PCA – Principal Component Analysis

PCs – Principal components

PDB – Protein Data Bank

PE – Phosphatidylethanolamine

PI – Phosphatidylinositol

PM – Plasma membrane

POPC – 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine

POPE – 1-palmatoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine

POPS – 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine

PRD – per-residue decomposition

PS – Phosphatidylserine

REACH – Research on Endothelin Antagonism in Chronic Heart Failure

Rg – Radius of Gyration

RITZ – The Randomized Intravenous TeZosentan

RMSD – Root Mean Square Deviation

RMSF – Root Mean Square Fluctuation

RTK – Receptor tyrosine kinase

SASA – Solvent-accessible surface area

SCD – Sickle cell disease

TM – Transmembrane

TM1 – Transmembrane helix 1

TM2 – Transmembrane helix 2

TM3 – Transmembrane helix 3

TM4 – Transmembrane helix 4

TM5 – Transmembrane helix 5

TM6 – Transmembrane helix 6

TM7 – Transmembrane helix 7

VSMC – Vascular smooth muscle cell

WHO – World Health Organisation

β_2AR – β_2-adrenergic receptor

ΔG_{bind} – Binding free energy

$\Delta G_{coulomb}$ – Coulomb energy

$\Delta G_{covalent}$ – Covalent binding energy

ΔG_{GB} – Generalized Born electrostatic solvation energy

ΔG_{solv} – solvation free energy

ΔG_{VdW} – Van der Waals energy

μOR – μ opioid receptor

φ – phi

ψ – psi

1. INTRODUCTION

1.1. Cell membrane

Membranes are a necessary part of a living cell. Plasma membrane (PM) (also known as the cell membrane) creates a cover surrounding the inner part of the cell, by separating it and its organelles, from the outside environment. PM is also involved in selective nutrients and waste transfer, and signal transmission because of the embedded receptors, channels, and carriers. The cell membrane is mostly built out of lipids and proteins, and both have different functions (*Lipid Bilayer Membranes - Chemistry LibreTexts*, n.d.; Tien & Ottova, 2003).

Lipids act as the main structural building blocks of the cell membrane, creating a lipid bilayer membrane (LBM). The bilayer is a fluid-like substance, and consists of two layers of lipid molecules, making it easy for proteins and other polymers to arrange themselves around the membrane (Fig 1.1A). The main functions of the lipids are to protect, retain cell structure, and act as enzyme cofactors and electron carriers. One of the most important groups of lipids is phospholipids. They are biological amphiphilic chemical compounds, constructed out of two parts, with one being hydrophobic and the other one hydrophilic, resembling a tadpole (Fig 1.1B) (*Lipid Bilayer Membranes - Chemistry LibreTexts*, n.d.; Tien & Ottova, 2003).

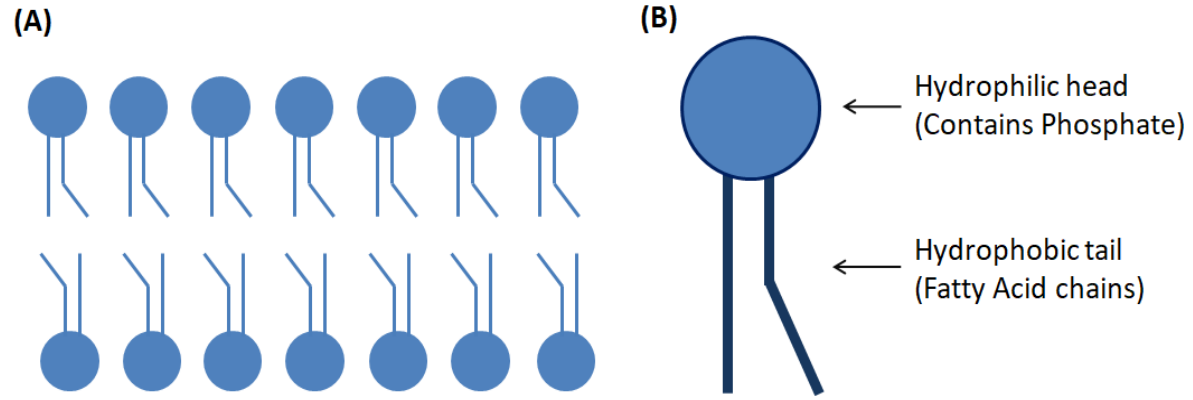


Figure 1.1. Phospholipid bilayer. (A) Representation of a bilayer consisting of two lipid layers, where the hydrophilic part of the lipid (sphere) is facing the outside of the bilayer, and the hydrophobic (line) – the inside. (B) Phospholipid representation, where the sphere is the hydrophilic head and the two lines are the hydrophobic tail (*Lipid Bilayer Membranes - Chemistry LibreTexts*, n.d.; Tien & Ottova, 2003)

Phospholipids can have different behaviours based on their polar (hydrophilic) head, and the two fatty acid chain (hydrophobic) tail. The hydrophobic part can be approximately 16 to 20 carbon atoms long. The fatty acid chains can be saturated (have no double bonds) and unsaturated (double-bonded). Most of the fatty acids in the bilayer are unsaturated. As

seen in Figure 1.1B, one of the chains has a slight kink. The particular reason for that are the double bonds that are mostly in *cis* form. This kink blocks the tight packing of the fatty acids and creates the liquid-like state of the membrane (*Lipid Bilayer Membranes - Chemistry LibreTexts*, n.d.; Tien & Ottova, 2003).

Based on the hydrophilic moiety's functional group phospholipids are classified as: phosphatidylcholines (PCs), phosphatidylethanolamines (PEs), phosphatidylserines (PSs), and phosphatidylinositols (PIs). In the mammalian cell, the PM is made out of 65 percent phosphoglycerides, 10 percent sphingolipids, and 25 percent sterols. Even though the polar heads in phospholipids are placed in different classes, the heads can still easily pack together. This creates asymmetry in the cell (Shahane et al., 2019; Tien & Ottova, 2003) Asymmetry is an important factor for cell viability, and the loss of it can induce apoptosis (cell death) (Tien & Ottova, 2003). For example, a plasma membrane of the red blood cell is constructed out of phospholipids consisting of lipid polar heads from several classes. The extracellular side of the bilayer (outer leaflet) is of PCs, while the intracellular (inner leaflet) is of PSs and PIs (Tien & Ottova, 2003). In other cells, the inner part of the membrane can also be of PEs. However, PS phospholipid mostly occupies the intracellular leaflet of the cell (Shahane et al., 2019). The PCs are 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), sphingomyelin and glycosphingolipids. PEs are 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) and 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPE). Further, PSs can be 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine (POPS) and 1,2-dioleoyl-sn-glycero-3-phospho-L-serine (DOPS) (Shahane et al., 2019).

The following text discusses some of the most used membranes in computational and experimental models, consisting of the previously mentioned phospholipids. Overall, PC is the most abundant phospholipid class in the membrane of an animal cell. As mentioned before, PCs are mostly found in the extracellular leaflet of the membrane (Shahane et al., 2019). Its member POPC is a neutral monounsaturated phospholipid with one fatty acid chain being saturated and the other unsaturated. The lipid can form membranes that are used for biologically mimetic model systems, and is found naturally and can be produced synthetically (Wanderlingh et al., 2017). 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) phospholipid is also part of the PC class. However, DMPC is a saturated phospholipid, meaning it does not have double bonds in one of the side chains. The following, however, is far less expensive (1 g costs approximately 186 EUR) when used in experimental models, compared to POPC (1 g costs approximately 310 EUR) (Marquardt et al., 2014). DMPC is

commonly used in computational models because it can be easily parameterized for most biomolecular force fields (Vila-Viçosa et al., 2014). This lipid is also very stable in experimental settings and easy to obtain commercially (Marquardt et al., 2014). 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC) is another member of the PC phospholipid class. DPPC is naturally present in lung surfactants, representing approximately 50 percent of all PCs. Studies show that DPPC and POPC are related to respiratory distress syndrome in infants. Ashton et al. (1992) analysed the content of PCs in infants with respiratory distress syndrome. They observed that infants with such disease exhibit lower DPPC percentage and DPPC:POPC ratio compared with the healthy subjects. Another interesting feature of DPPC is that it is a saturated lipid, similarly to DMPC, whereby it exhibits less fluidity and higher constraining ability for a signalling protein (Angladon et al., 2019). Also, DPPC has longer side chains than DMPC (Angladon et al., 2019). POPE is a member of the PE phospholipid class. This phospholipid is the most abundant of PEs. POPE is mostly found in the intracellular leaflet of the cell membrane (Shahane et al., 2019). Like the other phospholipids mentioned before, POPE is used for many different experimental membrane systems, such as micelles, liposomes, and more (Fattal et al., 2004; Simões et al., 2004).

1.1.1. Cell membrane receptors

For the cell to retain and perform its designated functions, as well as communicate to the external environment, some very important interactions are necessary. One of them is signal transduction. It is a process where an outer cellular signal is transferred through a membrane into the intracellular part, where an alteration happens. This process is mediated through the lipid bilayer and ligand-binding receptors' (also known as cell surface receptors) interactions. The receptors are transmembrane, meaning that they are placed into the membrane and extending across it. There are many receptors in a cell and they have very high specificity regarding certain cells (Yeagle, 2016; Tien & Ottova, 2003). All of the receptors are placed in three major receptor classes: ion channel-linked receptors, enzyme-linked receptors, and G protein-coupled receptors (*9.1C: Types of Receptors - Biology LibreTexts*, n.d.).

Ion channel-linked receptors are large receptors with several subunits, and with ligand bound opens up a channel allowing ions such as Na^+, K^+ Ca^{2+}, and Cl^- to pass in and out of the cell (Bylund, 2014). For example, acetylcholine, gamma-aminobutyric acid, and glycine

are ligands binding to the channel-linked receptors. These ligands are highly present in neurons and are mostly called neurotransmitters (Alexander Yi & Jan, 2002).

The second class is the enzyme-linked receptors. As the name suggests these receptors are associated with a specific enzyme. These receptors have an extracellular binding domain for ligand binding purposes and an intracellular part for enzyme binding. Unlike other receptor classes, they have only one transmembrane helical region (Litalien & Beaulieu, 2011; Waller & Sampson, 2018). Most of the associated enzymes are protein kinases, wherefore the receptors are placed in five groups: receptor tyrosine kinase (RTK), tyrosine phosphatase, tyrosine kinase-associated receptors, receptor serine-threonine kinase, and receptor guanylyl cyclase (Waller & Sampson, 2018). These receptors act upon transferring signals when a ligand binds to the extracellular N-terminal domain and actives the intracellular enzyme bound (Bylund, 2014). Enzyme-linked receptors are involved in cell apoptosis, differentiation, division and growth, as well as immune response, inflammation, and tissue repair (Alberts et al., 2002).

Lastly, the third class of membrane receptors is G protein-coupled receptors (GPCRs). This receptor class is one of the most diverse receptor classes, including such receptors as secretin, rhodopsin, dopamine, serotonin, and many more. GPCRs participate in a large number of vital signalling functions. Because of such variety, they are also the most studied type or receptors for drug research and design (Yeagle, 2016; Tien & Ottova, 2003). In the following chapter, the structure, function, and clinical relevance of GPCRs are described more in-depth.

1.2. G protein-coupled receptors

Since long ago the fascination surrounding membrane receptors has been remarkable in the scientific community. However, despite the fact that nowadays these receptors are one of the key players in the medical field, the initial idea that there may be receptors that are embedded into a membrane and can both locate and transmit signals, was hugely frowned upon. Nonetheless, later on, it was proven that such a receptor does exist. They were called G protein-coupled receptors (GPCRs) also known as seven transmembrane receptors (Lefkowitz, 2013).

GPCRs consist of seven α-helices that are looping partly in and out of the cellular membrane. These are called transmembrane (TM) helices, while the “middle” part is submerged in between the membrane, the N-terminus positioned on the outside of the

membrane and the C-terminus – in the cell. The helices are connected by three extracellular and three intracellular loops (ECL and ICL respectively), thus creating a tube-shaped tertiary structure, with the middle part being hollow for effector binding purposes (Katritch et al., 2013; Latorraca et al., 2017; Rosenbaum et al., 2009). Overall, most of the GPCRs consist of 310 to 470 residues. Some GPCRs have a very long N-terminus region, which places them in certain classes. G protein (guanine nucleotide regulatory protein) binds towards the C-terminal side of the intracellular region of GPCRs creating a ternary complex, hence the name (Mirzadegan et al., 2003; Palczewski, 2006).

Approximately 3 % of the human genome consists of GPCRs, which are one of the biggest and most diverse superfamilies of receptors in eukaryotes. These receptors contribute to a large number of different intracellular responses, such as molecular transport, light conversion and detection, and several more, by binding many different molecules and effectors, including lipids, protons, drugs, hormones, and neurotransmitters (Hanlon & Andrew, 2015; Satagopam et al., 2010). Most of the structural deviations between the GPCR family members are related to the extracellular and intracellular loops, and ligand-binding regions. From the intracellular perspective, the major structural differences between inactive and active receptor states are observed in the interaction between the glutamic acid/aspartic acid-arginine-tyrosine (E/DRY) motif of transmembrane helix 3 (TM3) and glutamate of transmembrane helix 6 (TM6), which stabilizes the inactive state (Rosenbaum et al., 2009). GPCRs can change its structure to become active in both ligand-bound and unbound states. However, it is more likely that with agonistic effector bound, the structural changes will become more drastic. Upon activation, a large outward movement away from the transmembrane bundle is observed in the intracellular end of TM6. This part of the helix also moves closer to the transmembrane helix 5 (TM5) of the receptor, creating a cavity. That is when the G protein C-terminus binds to the receptor and also stabilizes it, continuing by the closure of the extracellular loops over the effector binding site (Congreve et al., 2020; Rosenbaum et al., 2009). While G protein binds to the cytoplasmic part of the receptor, ligands bind, for example to the extracellular part through the N-terminus or the loops of the receptor in a hydrophobic pocket, called the orthosteric binding site, in the top middle of the GPCRs as shown in Figure 1.2 (Hanlon & Andrew 2015; Katritch et al., 2013; Latorraca et al., 2017; Rosenbaum et al., 2009).

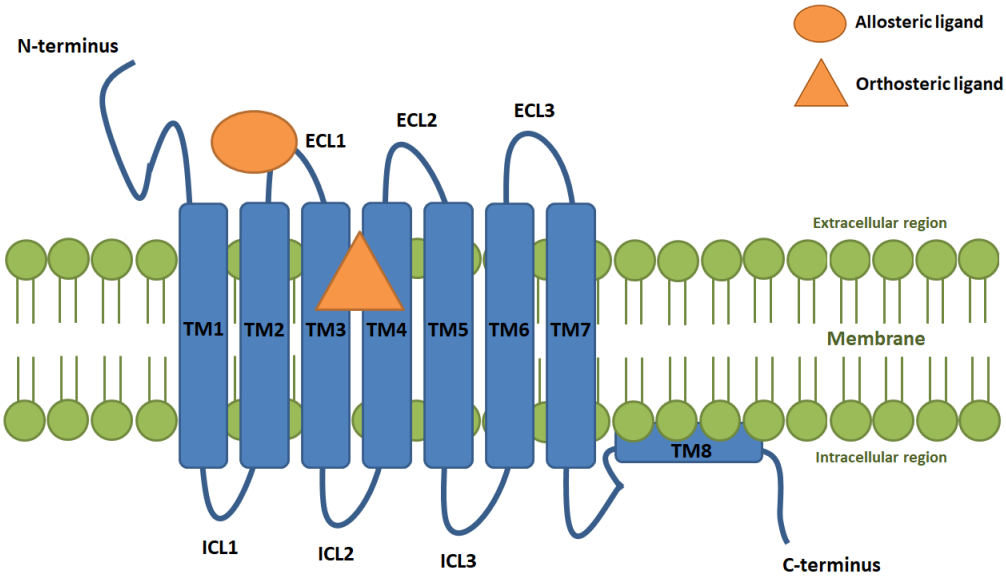


Figure 1.2. Schematic GPCR representation. Blue rectangles (helices) and strings (loops) represent the GPCR protein, while the green horizontal spheres and lines represent the cellular membrane. The GPCR is embedded partly into the membrane while its loop region extends towards both intra- and extracellular regions. An effector, in this case, an allosteric ligand (orange sphere) binds to the outside of the receptor (allosteric binding site), while the orthosteric ligand (orange triangle) binds to the middle top (orthosteric binding site) of the GPCR (Latorraca et al., 2017; Rosenbaum et al., 2009).

Generally, G protein-coupled receptors have two effector binding sites that are called orthosteric and allosteric. The former is surrounded by the highly conserved tryptophan of the TM6 and the extracellular loop 2 (ECL2). The residues that interact with the orthosteric ligand are mostly in TM3, TM5, TM6, and transmembrane helix 7 (TM7) (Chan et al., 2019, Rosenbaum et al., 2009). It is also stated by Chan et al. (2019) that the residues involved in the binding of ligands are also seen in transmembrane helix 2 (TM2) and the extracellular loops. Because it is the traditional binding site, most of the GPCR targeting drugs bind to the orthosteric binding site. It is so because the ligand does not need to completely penetrate the membrane since the binding pocket is wide open and easier to locate (Chan et al., 2019)

Apart from the orthosteric binding site, there is a different way for an effector to bind to the GPCR. It is called an allosteric binding site (Figure 1.2). There are several such sites around the protein. These sites are further away from the binding pocket, thus mostly binding to the outside of the receptor. When an effector binds to an allosteric site, it can regulate the protein and help to modulate the orthosteric site, by either negative-allosteric modulation (stopping signalling by an agonist) or positive-allosteric modulation (promoting signalling by an agonist) (Chan et al., 2019; Congreve et al., 2020). Chan and their team (2019) have

observed that effectors may bind allosterically to the surface of the ELC2, TM helices, and the TM helices in the centre of the bilayer, and the intracellular region. However, they have also stated that some of the allosteric binding sites may not be suitable for, for example – drug discovery, unlike the orthosteric site. The effector that binds to the GPCRs creates different structural changes. Upon which the activated G protein sends signals activating several biochemical reactions within the cell. Thus, it is very important to know whether the effectors target the allosteric or orthosteric binding site of the GPCRs.

1.2.1. Classes of GPCRs

GPCRs are found in not only humans but also bacteria, fungi and animals. It is stated that over 800 GPCRs are found in humans, 700 in zebrafish, 1000 in *Caenorhabditis elegans*, 1300 in mice, 200 in *Drosophila*, 50 in *Dictyostelium* and around 12 in a few species of yeast (Munk et al., 2019).

The receptors are classified based on their sequence, structural similarities, evolutionary homology, common ligands and function. Overall GPCRs are divided into six classes: Class A, Class B, Class C, Class D, Class E and Class F (Satagopam et al., 2010). However, in mammals GPCRs are divided into classes of four: rhodopsin (class A), secretin/adhesion (class B/B2), glutamate (class C) and Frizzled/Taste (class F). In some cases class B is sub-divided into two classes as class B and B2, making the whole mammalian GPCRs consisting of five classes instead of four (Krishnan et al., 2012).

Class A or rhodopsin class is the biggest class of GPCRs, and it is believed to be the most likely candidate for the “mother model” of all GPCRs. The class consists of 55 families and 640 subfamilies (Satagopam et al., 2010). Class A includes receptors for a wide variety of peptides, hormones, neurotransmitters, lipid-like substances and light. Most of the Class A receptors are involved in smell and vision. The highest amount of rhodopsin protein found is in the bovine retina (Palczewski, 2006). Taken up 80 % of all of the GPCRs, it is the most researched family regarding therapeutics (Davies et al., 2007; Satagopam et al., 2010).

Class B/B2 or secretin/adhesion class consists of 6 families and 16 subfamilies (Satagopam et al., 2010). The Class B receptors bind large lipids such as secretin, which is from the glucagon hormone family and is involved in the cAMP-mediated signalling pathways (Davies et al., 2007). Both Class B and Class B2 have very long N-terminus domain, with many so called adhesion domains (Watkins et al., 2012)

Class C or glutamate class consists of 4 families and 41 subfamilies (Satagopam et al., 2010). It is also called the metabotropic glutamate receptor class or mGluRs. Similarly to Class B/B2 it also contains a large extracellular N-terminus domain. This class of receptors are activated through an in the indirect metabotropic process (Davies et al., 2007).

Class F or Frizzled/Smoothened class has 2 families and 11 subfamilies (Satagopam et al., 2010). This class is the smallest one, but is very important in the hedgehog signalling pathway, and is a key regulator of development of an animal organism (Davies et al., 2007).

1.2.2. GPCR behaviour in the membrane environment

Since GPCRs are proteins inserted in the cell membrane, they have important and necessary interactions with the lipid bilayer. Those interactions are needed to perform protein designated functions. One of the protein-lipid interactions is called *protein lipidation*. It is a process where hydrophobic lipids bind to the receptor. Another one is how lipids place themselves in the cell membrane, for the purposes of affinity and localization of signalling proteins (Escribá et al., 2007). It is reported that bilayer organization is related to several diseases affecting GPCRs (Manna et al., 2019). Therefore, the GPCR behaviour in the phospholipid bilayer has been studied quite thoroughly, using synthetic lipid membrane systems. Two related studies will be discussed next.

Firstly, a computational study surrounding the adenosine A_1 receptor (A_1AR), which is part of the GPCR family, and membrane interactions was carried out by Bhattarai et al. (2020). The A_1AR is related to such diseases as cardiomyopathy, neuropathic pain and renal disease, and several more (Bhattarai et al., 2020; *WikiGenes - Adora1 - adenosine A1 receptor*, n.d.). Therefore, it is an important target for the drugs treating these diseases. The study was carried out by inserting the signalling protein in its two states (active and inactive) with and agonist or an antagonist, respectively, into the POPC membrane and simulating the ligand-protein-lipid interaction using Gaussian accelerated molecular dynamics (GaMD). Bhattarai and the team revealed that in both active and inactive state the ligand interaction with the protein remained stable. However, the fluctuations of the TM6 and TM5 intracellular region seemed to be much higher in the active state than inactive, explaining the membrane and lipid interaction importance for the fluidity. They also reported that the inactive state of the receptor interacted more with the POPC lipids in the outer part of the bilayer. However, the phospholipid head group heavily coordinated towards four lysine residues of the intracellular part of the TM6 in the A_1AR active state to stabilize the receptor. Hence, their

study revealed that GPCRs have different interactions with phospholipid bilayer depending on the activation state.

Secondly, a study by Angladon et al. (2019) was conducted regarding a different GPCR protein called μ opioid receptor (μOR). This receptor is related to several mood disorders and inflammatory bowel diseases, and mediates the effects of opioids (Angladon et al., 2019; Lutz & Kieffer, 2013; Philippe et al., 2006). In their Molecular Dynamics (MD) simulation study, they used DPPC and POPE lipid membranes, with POPC being used as a reference model. They reported that the biggest change in protein conformation was at TM5, where it exhibited large values of the tilt angles. Even though they reported similar changes in all three lipid systems, it seemed that the largest TM5 tilt value was observed for the POPC, followed by the DPPC and POPE lipid systems. They have also reported how differently the lipid membrane interacts and changes the conformation of the μOR: for all three membranes, the most interactive helices were transmembrane helix 1 (TM1), transmembrane helix 4 (TM4) and TM5. However, some helices preferred to interact differently among the membranes: TM1, TM5 and TM6 especially with POPC, TM5, TM6 and TM7 with POPE, TM4 and TM6 with DPPC. In summary, they concluded that DPPC, being saturated unlike POPC and POPE, can create more contacts with receptor helices. While POPC and POPE that are unsaturated, can introduce more flexibility to the GPCR's structure.

1.2.3. Clinical relevance of GPCRs

As mentioned previously GPCRs are abundant, and these receptors are involved in every step of the way of the development of an organism (Hanlon & Andrew, 2015). GPCRs are very sensitive to responding to sensory stimuli, such as light, odour and taste. They are also very important in the mediating hormones' effects (Satagopam et al., 2010). Since GPCRs are involved in a huge variety of intracellular signalling cascades, mutations of these receptors in humans are frequent and can lead to several diseases. For example, hormonal mutations or blockage of signalling caused by mutations can be the root cause of onset deaf, blindness, vascular diseases, endocrine disorders and many more, which are related to hormonal mutations or blockage of signalling (Hanlon & Andrew, 2015). This further supports the GPCRs' relevance in medicine and drug development.

As of now about 34 % of drugs bind to GPCRs and account for about 180 billion US dollars (149 billion euros) profit for pharmaceutical companies annually (Congreve et al.,

2020; Hauser et al., 2018). Most of the drugs created today bind to the orthosteric binding site and act as inhibitors (antagonists) or activators (agonists), which will be discussed in the later chapters. For the early development of most of the GPCR drugs, the high-resolution structural information of the receptors was not available. However, over time some important GPCR structures have been successfully determined and more unique receptors and ligands have been found, which has created even more opportunities for the design and creation of new medications (Hauser et al., 2017).

1.3. Endothelin

After the discovery of endothelium-derived relaxing factor back in 1980, five years later the existence of vasoconstricting factor was reported. However, the research for the vasoconstricting factor began only in 1987 by Masaki and the team, which also led to believe that the factor is actually a peptide. Thus the research had led to the discovery of endothelin (Masaki, 1998).

Endothelin (ET) is a polypeptide consisting of 21 amino acids with four cysteine residues that are interlinked to form two disulfide bridges (Figure 1.3 A) ET functions as a blood vessel tightening agent and also as a blood pressure raiser. However, it can also act as a vasodilator. The polypeptide is encoded by the human ET gene. Structurally the hydrophobic C-terminus and N-terminus of the peptide have slightly different functions: amino acids in the C-terminus bind to the receptor, while the N-terminus acts as a determinant for the binding affinity (Khimji & Rockey, 2010).

As stated by Masaki et al., (1998), endothelin was found in cultured bovine endothelial cells. However, after further analysis of the human ET gene, it was discovered that there are two other ET-like peptide genes. These peptides are called endothelin-2 (ET-2) (Figure 1.3 B) and endothelin-3 (ET-3) (Figure 1.3 C), thus renaming the first discovered ET as endothelin-1 (ET-1) (Masaki, 1998). This actively demonstrates that endothelin in total has three isoforms and they differ from each other by two (Trp and Leu) to six amino acids (Leu, Phe, two Tyr and two Thr) in their peptide chain (Davenport et al., 2016; Khimji & Rockey, 2010).

(A) H-Cys(1)-Ser-Cys(2)-Ser-Ser-Leu-Met-Asp-Lys-Glu-Cys(2)-Val-Tyr-Phe-Cys(1)-His-Leu-Asp-Ile-Ile-Trp-OH
(B) H-Cys(1)-Ser-Cys(2)-Ser-Ser-Trp-Leu-Asp-Lys-Glu-Cys(2)-Val-Tyr-Phe-Cys(1)-His-Leu-Asp-Ile-Ile-Trp-OH
(C) H-Cys(1)-Thr-Cys(2)-Phe-Thr-Tyr-Lys-Asp-Lys-Glu-Cys(2)-Val-Tyr-Tyr-Cys(1)-His-Leu-Asp-Ile-Ile-Trp-OH

Figure 1.3. Endothelin structure in IUPAC representation. (A) endothelin-1, (B) endothelin-2, (C) endothelin-3 (*Endothelin-1 | C109H159N25O32S5 - PubChem*, n.d.; *Endothelin 2, human | C115H160N26O32S4 - PubChem*, n.d.; *Human endothelin-3 | C121H168N26O33S4 - PubChem*, n.d.).

The three isoforms are produced from different human organs and cells. Endothelin-1 is mainly sourced from endothelial cells, but can also be found in the cardiomyocytes, vascular smooth muscle cells as well as macrophages and monocytes, airway epithelial cells, fibroblasts, brain neurons, enteric glial cells and many other cells. Endothelin-2 is mostly found in the intestine, kidney, ovary, vasculature, heart and lung cells. Unlike ET-1 and ET-2, endothelin-3 is not synthesised by human endothelial cells, but is expressed in the adrenal gland, glial cells, and neuronal stromal cells. It also plays an important role in cellular proliferation and development, as well as mediates release of vasodilators (Barton & Yanagisawa, 2008; Davenport et al., 2016; Masaki, 1998).

The synthetic pathway of endothelin is very complex as seen in Figure 1.4. For the synthesis of ET-1, the pathway starts by the transcription and translation of the preproendothelin-1 mRNA where the preproendothelin-1 is synthesized. Preproendothelin-1 is a 203 amino acid peptide. This peptide undergoes proteolysis, a breakdown into shorter peptides, with the help of furin-like protease. This process creates big endothelin-1, which is a chain of 37 to 41 amino acids; however, this type of ET is biologically inactive. Big endothelin-1 then undergoes cleavage with the aid of membrane-bound zinc metalloprotease called endothelin-converting enzymes (ECEs) and also the non-ECE metalloproteinase, chymase and vascular smooth muscle cell (VSMC) chymase. This produces the biologically active endothelin-1. The same pathway is also assumed to be followed by both ET-2 and ET-3. However, for both ET-2 and ET-3, only the ECEs are active during the big endothelin cleavage (Barton & Yanagisawa, 2008; Khimji & Rockey, 2010).

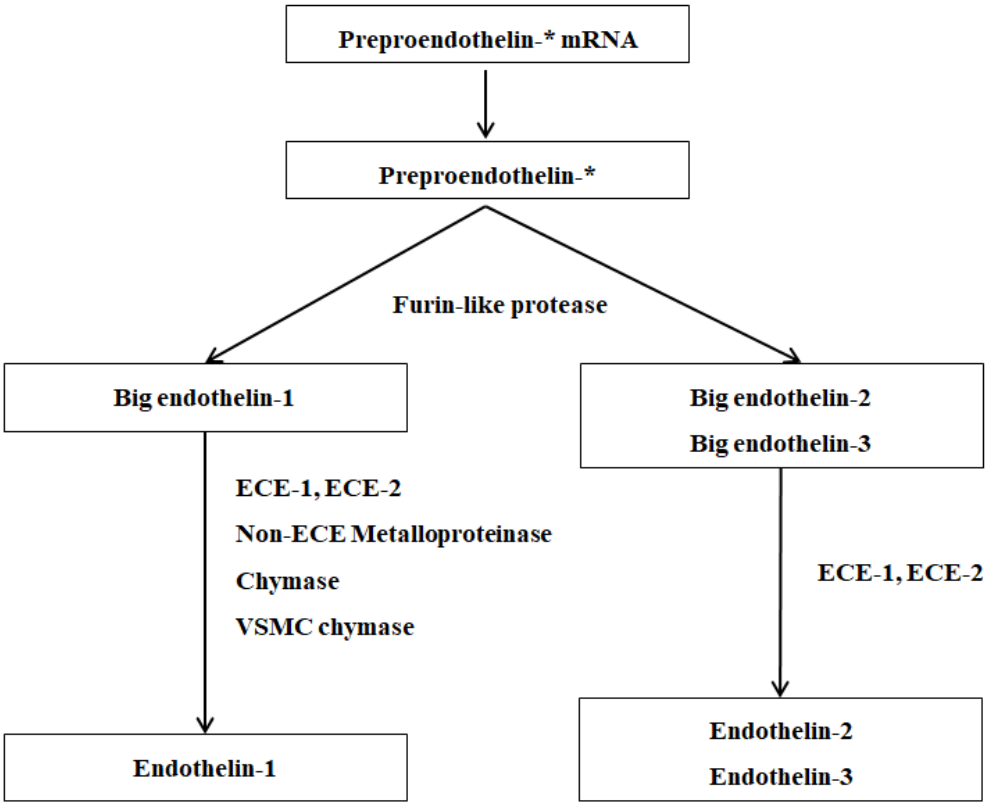


Figure 1.4. Endothelin synthesis pathway. The asterisk (*) represents number 1, 2 or 3 for either ET-1, ET-2 or ET-3 (Barton & Yanagisawa, 2008).

1.3.1. Endothelin receptors

Endothelins participate in a wide range of physiological functions in the human body. Therefore these peptides function through binding to their highly selective receptors in an extremely complicated process. One of the steps in the process is binding its C-terminal tail deeply into the receptors' orthosteric binding pocket (Houde et al., 2016).

The receptors are called endothelin A (ET_A) and endothelin B (ET_B). ET_A is a non-peptide selective receptor, while ET_B is peptide-selective. Both of these receptors belong to the GPCR family of Class A or rhodopsin class, whose structure includes seven transmembrane helices. The receptors are constructed out of approximately 400 amino acids in total, with the transmembrane domains made out of 22 to 26 hydrophobic amino acids (Khimji & Rockey, 2010). Interestingly enough, in the paper of Davenport et al. (2016), it is stated that the genes encoding peptides are an innovation that has evolved strictly in the vertebrates such as jawless vertebrates, cartilaginous vertebrates and bony vertebrates.

Endothelin A and endothelin B in human are expressed slightly differently between the organs. ET_A is comprised of 427 amino acids as described in the UniProt database under entry code P25101 and it is 63 percent similar to the endothelin B in human (Davenport et al.,

2016). Generally, endothelin A receptor is less abundant than ET_B and can be mostly sourced from the smooth muscle cells, myofibroblasts, cardiomyocytes, hepatocytes, neurons, osteoblasts, keratinocytes and adipocytes (Khimji & Rockey, 2010). Endothelin B unlike ET_A has a slightly longer peptide sequence comprised of 442 amino acids. Compared to ET_A the receptor has a far longer N-terminus. Regarding the binding of an effector, TMs that contribute are TM1 to TM3 and TM7, as well as the related extracellular loops. Transmembrane helices TM4 to TM6 and their ECLs are involved in the selectivity of the receptor binding, and C-terminal participates in the signalling (Mazzuca & Khalil, 2012). These actions are important for activation of the receptor. ET_B can be located in the same cells as the ET_A mentioned before, however ET_B can also be spotted in the endothelial cell and renal collecting-duct cells (Khimji & Rockey, 2010). With respect to the organ distribution, ET_B in the human brain represents up to 90 percent of total endothelin receptors (ETRs) (Davenport et al., 2016). In heart tissue, ET_A comprises 60 percent of endothelin, while in the kidney 70 percent of ETRs are reported to be ET_B. In addition, in the smooth muscle layer and human vasculature ET_A is more abundant compared to ET_B. Thus taking all into account, human kidney, lung, brain and liver are mostly ET_B rich, whereas heart is ET_A rich (Maguire & Davenport, 2015).

Both ET receptors can signal via calcium-dependent and calcium-independent pathways (Houde et al., 2016). However, their ligand affinity differs when it comes to ET-1, ET-2 and ET-3. Only ET-1 and ET-2 bind to ET_A and signal via it, while all three ET isoforms bind to ET_B (Davenport et al., 2016). Additionally, as reported by Shihoya et al. (2017), when ET-1 activates the ET_B receptor, it induces conformational changes to the TM6 and TM7, causing their extracellular regions to move inwards by 4 Å, where after the following conformational changes in the core and the TM7 cytoplasmic part cause the TM6 intracellular region to move outwards. The receptors also express different types of activities from each other. When ET_A is activated, it functions as a narrower of blood vessels. On the contrary, ET_B receptor upon activation releases vasodilators and expands the blood vessels, and the receptor may even act as inhibitor for both cell growth, and vascular constriction in the blood vessel system (Kawanabe & Nauli, 2011). Because of these opposite effects to those of the ET_A receptor, ET_B-selective agonists have been studied as vasodilator drugs for the improvement of tumour drug delivery, as well as for the treatment of hypertension, chronic heart failure and more (Davenport et al., 2016; Kawanabe & Nauli, 2011). Targeting the ET_B signalling system may be an important step towards the treatment of related human

diseases. For example, as stated by M. Houde et al. (2016) about the relevance of ET-1 and its receptor interactions:

> A better understanding of how ET-1 is synthesized and how ET_A and ET_B receptors interact could help design better pharmacological agents in the treatment of cardiovascular diseases where targeting the ET-1 system is indicated (p. 144).

1.3.2. Endothelin related diseases and treatment

1.3.2.1. General receptor-targeting drug types

Generally, there are three types of receptor-targeting drugs: agonists, antagonists and allosteric modulators. The agonist and antagonist types are characterized based on their biological response (%) for a given tissue. Between the two, the drug type with the highest response is an agonist; while the lowest response is observed in the antagonist binding, where it shows no response at all and blocks the effect of an agonist (Kenakin, 1987). The two drug types bind to the receptor mimetically to the natural ligand (Kenakin, 1987).

Receptor agonists are classified into three categories: full, partial and inverse. A full agonist when bound activates the receptor to its fullest. A partial agonist acts similarly but with lower efficacy, producing a non-maximal response. An inverse agonist binds like a full agonist but expresses inhibitory effects and exhibits a complete negative response (Kenakin, 1987; *Pharmacological Glossary | Tocris Bioscience*, n.d.).

Receptor antagonists, on the contrary, do not activate the receptors when bound. Hence and antagonist inhibits the receptor's activity and prevents other molecules from binding and activating the receptor. Antagonists are classified into two categories: competitive and non-competitive. Competitive is the most well-known type of antagonists; it acts as an inhibitor. On the other hand, a non-competitive antagonist binds to the allosteric binding pocket instead of the usual orthosteric site and prevents the activation of the receptor (*Pharmacological Glossary | Tocris Bioscience*, n.d.).

Allosteric modulators, unlike agonists and antagonist, bind to the allosteric site, which is distinct from the orthosteric site. This type of drug can regulate the protein and help to modulate the orthosteric site. Additionally, when bound it induces conformational changes and increases affinity of a ligand bound to the orthosteric pocket. Allosteric modulators are classified into positive and negative. The former class of modulators promotes signalling by an agonist and increases its affinity, while latter stops signalling by an agonist and decreases its affinity (*Pharmacological Glossary | Tocris Bioscience*, n.d.).

1.3.2.2. Diseases and treatment

Out of the three ET isoforms, ET-1 is the most abundant one, and thus the most researched isoform. Since ET-1 is both ET_A and ET_B selective, most of the endothelin-related diseases and their treatment are associated with the ET-1 system (Davenport et al., 2016). Several clinical studies have demonstrated that endothelin is involved in diseases such as pulmonary arterial hypertension (PAH), arterial hypertension (AH), renal diseases, allograft rejection, diabetes and insulin resistance, cancer, heart failure, and atherosclerosis (Barton & Yanagisawa, 2008).

Most endothelin-related diseases are affecting the vasculature (Barton & Yanagisawa, 2008; Kawanabe & Nauli, 2011). The most well-known of such diseases is called pulmonary arterial hypertension (PAH). PAH is a rare disorder that is progressive and shows signs of elevated blood pressure (hypertension), in the arteries of the lungs (pulmonary artery) (Kawanabe & Nauli, 2011; Schneider et al., 2007). The cause of the disease is unknown. People mostly affected by this disorder are women of ages 30 to 60, and people with disorders such as connective tissue disease, heart failure and HIV infection (Barton & Yanagisawa, 2008; Schneider et al., 2007). The late prognosis of patients with PAH leads to poor treatment outcomes, and most patients die within two to three years after the diagnosis was made. The first successful treatment was vasodilator therapy using calcium channel blockers or prostaglandin administration. However, it proved to cause side effects and the efficacy was very limited, especially for patients with non-vasoreactive PAH (Schneider et al., 2007). Therefore, other pharmacological options were being considered, one of them being ET receptor antagonists (Iglarz et al., 2015). Regarding ET and the disease in general, ET-1 is very important for normal blood circulation in the lung artery. In people with normal blood pressure, the levels of ET-1 are stable. However, patients affected with PAH exhibit high levels of ET-1 in the pulmonary artery (Kawanabe & Nauli, 2011; Schneider et al., 2007). It has also been observed that high levels of ET-1 cause vascular remodelling (Kawanabe & Nauli, 2011). This discovery led to the belief that there might be an overproduction or decreased clearance of the circulating ET-1. It was shown then that the PAH patients exhibited regular ET-1 clearance but showed high production of ET-1. Therefore, this proved that one of the main causes for PAH is the overproduction of ET-1, due to the increased ET-1 mRNA expression in the pulmonary vascular endothelial cells (Kawanabe & Nauli, 2011; Schneider et al., 2007). Overproduction of ET-1 leads to overexpression of ET_A and ET_B. Both receptors are very important for retaining normal pulmonary vascular tone in a hypoxic environment, which means that any changes in ET_A or

ET_B receptor expression can cause PAH (Mazzuca & Khalil, 2012). For example, Ivy et al. (2005) in their study discovered that ET_B receptor downregulation in rats can cause PAH. On the contrary, ET_A receptor increase can also lead to the development of PAH as reported by Takahashi et al. (2001) who mention that ET_A expression can become upregulated two-fold. This suggests that ET-1 and endothelin-receptor targeting drugs could be good alternatives for PAH treatment.

As mentioned before, the previously tested drugs displayed low efficacy and related side effects. Therefore, there was a need for improved medication options. Several ET receptor antagonists have been developed that can help to slow down the further development of PAH (Schneider et al., 2007). The first of its kind ET receptor antagonist was called bosentan (Figure 1.5).

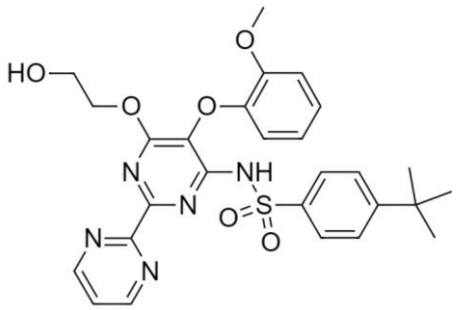


Figure 1.5. 2D chemical structure of bosentan (*Bosentan | C27H29N5O6S | ChemSpider*, n.d.).

It became the first to be approved as a PAH targeting medication by U.S. Food and Drug Administration (FDA) back in 2002, specifically for the patients that exhibit symptoms in the New York Heart Association (NYHA) and World Health Organisation (WHO) functional class II and IV. It is also the most utilized drug to date (Barton & Yanagisawa, 2008; Kawanabe & Nauli, 2011; Mazzuca & Khalil, 2012). Bosentan is a non-peptide sulphonamide with mixed ET_A and ET_B inhibitory activity. It acts by binding to both endothelin receptor types and blocking them, thus inhibiting the effects of endothelin-1. The drug has been shown to be effective in two trials with a survival rate of between 96 and 89 percent after one and two years respectively, with the orally administered dose of 125 mg twice a day (Kawanabe & Nauli, 2011; Mazzuca & Khalil, 2012; Schneider et al., 2007). Bosentan has a vasodilatory effect and can inhibit pulmonary arterial smooth muscle cell proliferation and prevent pulmonary vasculature remodelling, as well as help to treat digital ulcers of systematic sclerosis (Kato et al., 2020; Kawanabe & Nauli, 2011). In a very recent study by Kato et al. (2020) regarding bosentan, it was discovered that besides the involvement of smooth muscle cells, endothelial cells (ECs), macrophages and fibroblasts in

pulmonary vasculature, and bone marrow (BM) stem or progenitor cells can contribute to the development of PAH. They reported that the regulation of BM-derived ECs and BM-derived macrophages by bosentan in the lungs can help to control PAH in mice. More specifically BM-derived cells might have been promoted under upregulation of nitric oxide synthase 3 (eNOS) when ET_B blockage was active. Also through blocking the ET_A receptor, macrophage expression was downregulated.

Other medications approved by the FDA for the PAH patients exhibiting the WHO class II and IV symptoms were ambrisentan, enrasentan and sitaxentan (Figure 1.6).

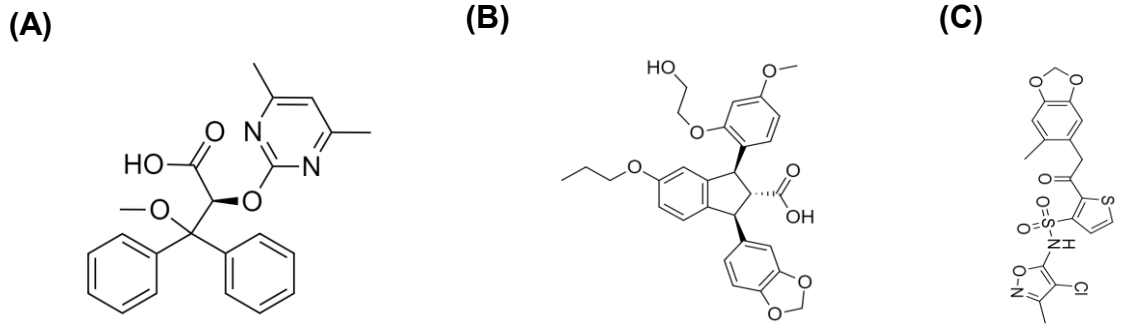


Figure 1.6. 2D chemical structures of (A) ambrisentan, (B) enrasentan and (C) sitaxentan (*ambrisentan | C22H22N2O4 | ChemSpider*, n.d.; *Enrasentan | C29H30O8 | ChemSpider*, n.d.; *sitaxentan | C18H15ClN2O6S2 | ChemSpider*, n.d.).

However, unlike bosentan that is binding non-selectively to the ETRs, all three drugs are ET_A selective antagonists. For example, ambrisentan seems to be effective at the dosage of 5 to 10 mg when taken daily, showing improved one-year survivability. All three drugs reduce pulmonary artery pressure and help to control vascular remodelling (Kawanabe & Nauli, 2011; Mazzuca & Khalil, 2012; Schneider et al., 2007).

On the other hand, medications are usually paired with several unwanted adverse effects. Bosentan is no exception. Its adverse effects include hepatotoxicity, edema, teratogenicity in pregnant women and several more (Belge & Delcroix, 2019). The latter is a common adverse effect among all ETR targeting drugs. Therefore a new ET receptor antagonistic drug was developed called macitentan (Figure 1.7).

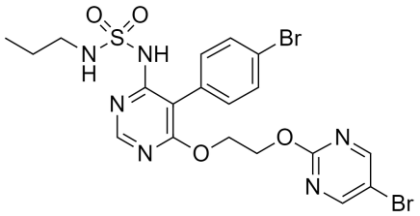


Figure 1.7. 2D chemical structure of macitentan (*Macitentan | C19H20Br2N6O4S | ChemSpider*, n.d.).

Macitentan is a non-ETR selective antagonist. This drug was approved as a viable option for PAH treatment back in 2013 by FDA and European Medicines' Agency (EMA) at the dose of 10 mg taken orally daily (Bedan et al., 2018). In the studies by Bedan et al. (2018) and Iglarz et al. (2015), macitentan is illustrated as having a far greater and more prolonged ET receptor binding occupancy compared with bosentan. Therefore, macitentan can better block the ETRs despite the high prevailing ET-1 concentration. Regarding the binding of macitentan, Gatfield et al. (2012) suggested that since the drug differs greatly from the structure of bosentan, it most likely has a different binding mode at the ET receptors. Moreover, studies by both Bedan et al. (2018) and Iglarz et al. (2015) revealed that macitentan, unlike bosentan, has better tissue penetration and distribution properties. Treatment by macitentan results also to lower morbidity, and the mortality risk reduced to 50 percent. Regarding some of the adverse effects of bosentan, macitentan greatly reduced the chance of developing edema (Maguire & Davenport, 2015). However, regarding the risk of teratogenicity in pregnant women, macitentan showed no difference compared to bosentan (Bedan et al., 2018).

Apart from the PAH disease, ET-1 and ETRs are involved in other disorders as well. One such reported disorder is sickle cell disease (SCD). SCD is characterized by an inherited unique mutation in the β-globin gene that is related to haemoglobin. This leads to normal blood flow blockage by red blood cells related to the modification of their shape. The reported ET-1 involvement is illustrated by Koehl and their team (2017). They mention that in patients with SDC ET-1 system is being activated and the concentration of ET-1 is high compared to healthy subjects. This leads to a prominent adhesion of neutrophils to endothelium by the functioning ET_B receptors. The high concentration of active neutrophils may interact with several vascular cells, and cause decreased blood flow and endothelial injury.

Another disease that is related to ET system is hepatopulmonary syndrome (HPS). HPS is described by dilatation of blood vessels in the lungs and as a complication of a liver disease (scarring of the liver or cirrhosis). This disorder occurs in 5-32 percent of patients that have such a liver disease (*Hepatopulmonary Syndrome - NORD (National Organization for Rare Disorders)*, n.d.). Ling et al. (2004), has reported that ET-1 and ET_B receptor has a relevant role in the experimental HPS (a model where the physiological abnormalities are reproduced in animals similarly to human disease). Ling and the team reported that ET-1 concentration is enhanced by ET_B receptor activation in the pulmonary microvascular endothelium while HPS is still developing. They also discovered that selective ET_B

antagonism helped to improve HPS. Therefore, these reported findings may have relevance to the human disease, because in human patients with pulmonary microvascular dilation ET-1 levels were also elevated compared to healthy subjects as reported by Koch et al. (2012).

Chronic heart failure (CHF) or also known as congestive heart failure is another disease that is highly researched. Similarly to PAH, in patients that exhibit CHF, ET-1 levels are highly increased, two to three times higher than normal, due to the minimized clearance of ET-1 (Khimji & Rockey, 2010; Schneider et al., 2007). When normal cardiac function is impaired, the human heart starts to fail showing unusually high levels of endothelin tissue, while in the healthy human heart normal levels of ET contribute to inotropy, chronotropy, and arrhythmogenesis, and contractile function of cardiac myocytes (Barton & Yanagisawa, 2008, Schneider et al., 2007). ET-1 also plays a big role in the ventricular remodelling of the heart. Interestingly enough, ET-1 overexpression mostly happens in the left ventricular, alongside with the overexpression of the ET_A receptor. On the contrary, the ET_B receptor is downregulated (Rich & McLaughlin, 2003).

Many clinical trials have been carried out to study the endothelin receptor blockers as a medication for CHF (Barton & Yanagisawa, 2008). For example, RITZ (The Randomized Intravenous TeZosentan) trial underwent examination of the usage of tezosentan (Figure 1.8).

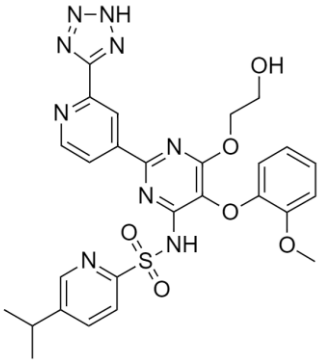


Figure 1.8. 2D chemical structure of tezosentan (*tezosentan | C27H27N9O6S | ChemSpider*, n.d.).

However, the outcome of the trial was mostly negative as it showed no significant difference in the primary endpoint. Another such trial was REACH (Research on Endothelin Antagonism in Chronic Heart Failure) where the examined ET receptor antagonist was bosentan. However, the trial was prematurely terminated because of a risen concern over the elevated concentration of liver enzymes (up to 15.6 %) (Rich & McLaughlin, 2003; Schneider et al., 2007). Another trial that also examined bosentan's effect on CHF was called ENABLE (Endothelin Antagonist Bosentan for Lowering Cardiac Events). The trial seemed to be a failure as well, as it showed no improvement. Instead, it was reported that the trial

subjects exhibited elevated fluid retention, edema and transaminase concentration (Rich & McLaughlin, 2003; Schneider et al., 2007).

On the other hand, trials HEAT (Heart Failure ET(A) Receptor Blockade Trial) and EARTH (Endothelin A Receptor Antagonist Trial in Heart Failure) were demonstrating some promising results. The both trials examined ET_A receptor antagonist darusentan (Figure 1.9).

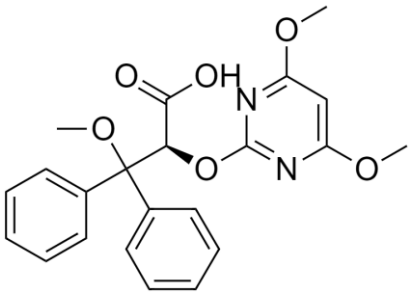


Figure 1.9. 2D chemical structure of darusentan (*Darusentan | C22H22N2O6 | ChemSpider*, n.d.).

This HEAT trial showed significant improvement in cardiac index. Yet increased ET-1 levels were reported even though the ET_A receptor was blocked. Other than that, major concerns were raised as 4 deaths were reported. Even though the EARTH trial was initially giving some promising results as darusentan was well tolerated, unfortunately after 6 months no beneficial effects were observed, as CHF worsened and 4.7 percent of patients died (Rich & McLaughlin, 2003; Schneider et al., 2007).

Additionally, Sawaki et al. (2000) researched the effects of K-8794 (Figure 1.10), an orally taken ET_B selective antagonist, for CHF treatment administered in a CHF experimental model (dogs with heart failure).

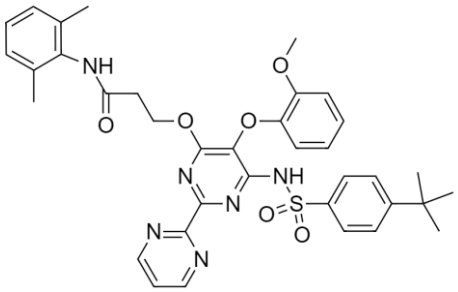


Figure 1.10. 2D chemical structure of K-8794 (*K-8794 | Ligand page | IUPHAR/BPS Guide to PHARMACOLOGY*, n.d.).

Two groups were investigated: placebo (control) and K-8794 group. They reported that when K-8794 was taken, blood pressure decreased. However when compared to the placebo group the values of K-8794 group were higher. Increase in levels of systematic vascular resistance was also demonstrated in both groups, with K-8794 group displaying

higher values. They concluded that K-8794 may still help with tackling CHF by prevention of body fluid retention, even though the antagonist expressed disadvantages to the flow of blood in the experimental model.

In sum, ET-1 and ETRs are involved in many different diseases. Apart from the PAH, HPS, SCD and CHF disorders, the ET-1 system is also involved in renal disease, diabetes and obesity (D'Orléans-Juste et al., 2019). Endothelin is also reported to have an important role in essential hypertension, atherosclerosis, central nervous system diseases, carcinogenesis, Alzheimer's disease, and Hirschsprung's disease as presented by Kusafuka & Puri (1997), and many more (D'Orléans-Juste et al., 2019). In some of the diseases, several ET receptor targeting blockers have shown good results. However, in some of the trials, the results demonstrated a failure, as ETR antagonists display unwanted adverse effects (Barton & Yanagisawa, 2008). Nevertheless, new and improved drugs are being continuously developed, such as macitentan, showing lesser adverse effects. Furthermore, well-designed clinical studies may present a far better therapeutic potential for the new generation of drugs.

1.4. Molecular Dynamics Simulation

In the previous chapters, Molecular Dynamics (MD) simulation was mentioned as the method used in the discussed lipid-receptor interaction studies. Because of its effortless approachability, MD simulation has become a popular approach for various studies aiming to understand molecular motions and interactions. MD simulation is a computer simulation method that incorporates Newtonian laws of motion to investigate the physical movements of atoms and molecules such as water, macromolecules, lipid membranes, nucleic acids and more complex systems of biomolecules, to reproduce the behaviour of the biological environment and understand the biological aspects of such systems (Miller, 1995). The approach was developed in the late 70s and has advanced into a far more complex method leveraging the super-computational power (Hospital et al., 2015).

MD simulation can be applied for example for studying allostery, structure prediction, molecular docking and drug design (Hospital et al., 2015). The latter is one of the most interesting topics nowadays. For example, MD simulations can simulate and reveal information about a biomolecule's conformational changes that take place during ligand binding. By the method it is possible to spot changes in a protein, induced by the outer environment such as membrane, and reveal direction of movements related to folding and much more. The method can furthermore help to effectively understand structural stability

and changes in free energy (Hernández-Rodríguez et al., 2016; Hollingsworth & Dror, 2018; Hospital et al., 2015).

The actual algorithm is quite complex, but the idea is easy to comprehend. In the beginning, the whole simulation time must be separated into several short time steps – around 10^{-15} seconds per step. Fortunately, biological events happen in short time scales (nanoseconds, microseconds), making it possible to simulate them with the available computational power. A molecular mechanics force field must be selected that calculates forces acting on each atom. Later the position and velocity of an atom must be calculated and updated every step using Newtonian laws of motion (Hollingsworth & Dror, 2018). Force field includes a combined calculation of bond angles, van der Waals and electrostatic interactions for parameterized molecular systems. The most popular biomolecular force fields used nowadays are AMBER, CHARM, GROMACS and OPLS (Hernández-Rodríguez et al., 2016). The choice of the force field depends on the scientists themselves, and the research carried out. Apart from the force field, the importance of the solvent rises. The biological solvent is water. Therefore, before running the experiment the solvent molecules are necessary to add because they instigate important changes to the biomolecule. Salt ions are added to neutralize the system and often also to simulate the salt concentration at physiological conditions, whereas lipids are used to imitate cellular membranes. Thus, the simulation box that is used to perform the MD simulation is a model of a biomolecular system that consists of the biomolecule(s) in the solvent (Figure 1.11) (Hernández-Rodríguez et al., 2016; Hollingsworth & Dror, 2018; Hospital et al., 2015).

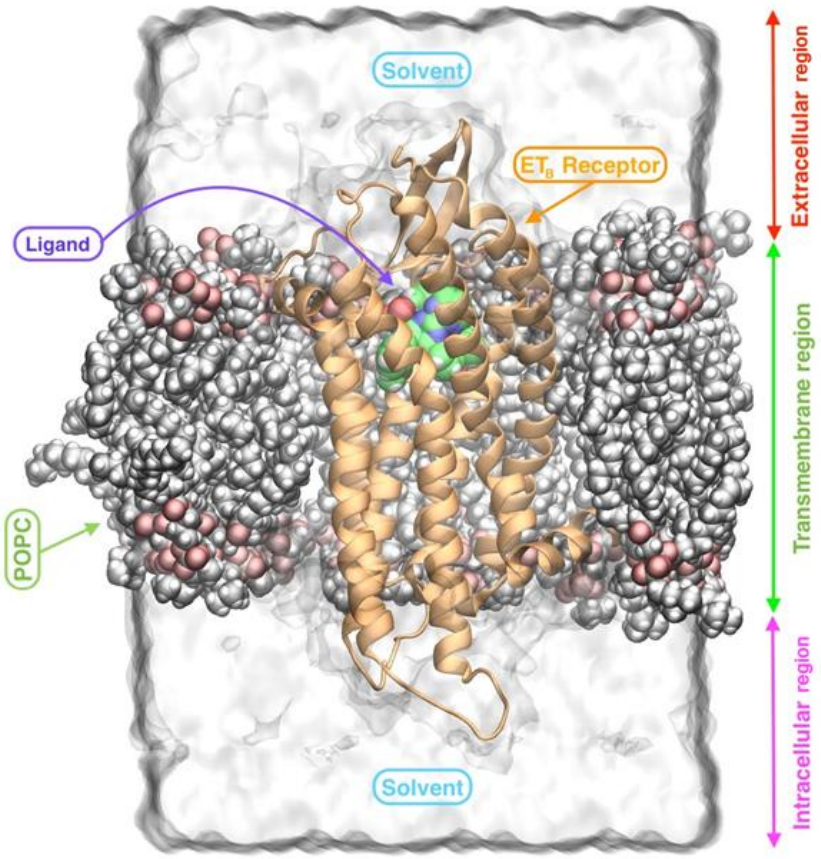


Figure 1.11. The simulation box of a system comprising of the ET_B receptor in complex with K-8794 embedded in POPC membrane used for MD simulations.

The 3D structure of biomolecular targets for the simulations are usually retrieved from the Protein Data Bank (PDB), which is an archive for storing the information about different types of biological structures (biomolecules), and their experimentally determined 3D shapes (Berman et al., 2000). On average, MD simulations can routinely be carried out for systems containing ca. 100,000 atoms. However, depending on what type of computer hardware or MD algorithms are available, much bigger systems are also becoming feasible to simulate (Hollingsworth & Dror, 2018; Hospital et al., 2015). Besides that, there are many MD simulation (MDS) softwares to choose from such as AMBER, NAMD, CHARMM, GROMACS, and DESMOND to name a few. In current days, most of the mentioned software runs on graphic processing unit (GPU) utilization, which promises accelerated performance when compared to central processing units (CPU) (Hernández-Rodríguez et al., 2016). The simulation must have a defined simulation time, which varies from picoseconds to milliseconds, and a frame count (how many snapshots are saved along the simulation). This at the end creates a trajectory. The trajectory is assembled of the system snapshots created

throughout the simulation run that can be superimposed over the first initial frame (initial system), creating a movie like an animation (Hollingsworth & Dror, 2018).

MD simulation can reveal different types of biological and structural details. For the analysis of such, many different approaches can be used. The most basic method is the visual analysis of the trajectory, whereas more informative tools include, to name a few: Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Radius of Gyration (Rg), Clustering Analysis, Free Energy Calculations, Principal Component Analysis (PCA) and more (Hernández-Rodríguez et al., 2016).

Like every method, MD simulation has its limitations. The method is very time consuming (Biagini et al., 2019). Especially when the simulation is set to run for 500 nanoseconds or even more. The MD simulation can take up to days or weeks of simulation job time (Hollingsworth & Dror, 2018; Hospital et al., 2015). Apart from the actual simulation, the analyses of the trajectory can also take several hours. Importantly, one run of MD simulation is usually not enough to gather accurate results, because for example, MD simulation did not reach equilibrium. This means that several simulations of one system must be carried out and analysed, which can take up to weeks (Hollingsworth & Dror, 2018; Hospital et al., 2015). Therefore software and hardware improvements, such as GPU utilization, have been a major advantage regarding MD simulation (Biagini et al., 2019). Furthermore, free energy calculations are merely approximations and offered methods may be more or less superior to each other. However, at the end of the day, the most important aspect of MD simulation precision is the accuracy of the 3D structures analysed (Hernández-Rodríguez et al., 2016; Hollingsworth & Dror, 2018; Hospital et al., 2015).

1.4.1. Molecular Dynamics Simulation and GPCRs

Several studies have incorporated MDS for the purpose of analysing how effectors interact, affect and induce structural changes to the signalling protein. The changes induced, however, depend on the effector structure and type (Ferruz et al., 2018; Renault et al., 2019; Singh et al., 2019; Yuan & Xu, 2018) In the following text several studies, of which MDS was used for analyses of GPCRs, will be discussed.

Histamine H3 receptor of the GPCR family activity has been studied by Herrera-Zúñiga and the team (2019) using MD simulation. The active state, agonist-bound and the inactive state, antagonist-bound to H3 receptor was simulated in the DPPC phospholipid bilayer system, together with the apo receptor in the active state for a comparative analysis.

The study revealed that throughout the simulation all ligand-bound active and inactive states of the receptor go through conformational changes resulting in four clusters each. While the apo exhibits only three. Regarding the clusters, the scientists presented that for all the structures the extra and intra cellular loop motion is highly flexible, especially the ECL2, extracellular loop 3 (ECL3), and the intracellular loop 2 (ICL2), intracellular loop 3 (ICL3). On the other hand, the antagonist-bound structure revealed a tilt like movement of TM1, and a pendulum-like rigid-body motion of TM5-TM7 and their loops, while the helices of the agonist-bound structure (which represents the active state) did not exhibit important movements. Regarding the apo structure, the flexibility of the loops is even more enhanced. These findings show how the absence or the presence of a ligand can affect the receptor structure. They also observed that the ligands interact with the orthosteric binding pocket residues differently through the simulation by breaking and establishing new hydrogen bonds.

Another study has also reported MDS being a good method for understanding GPCR structural changes and signalling. Ichikawa et al. (2016) investigated how 14 different ligands (mostly agonists) of β_2-adrenergic receptor (β_2AR) affect the G-protein and β-arrestin bound active and inactive states, while embedded in the POPC membrane. During the study they observed that the well-known kink in helix 6 is present in the G-protein-active state. Similarly to the previously discussed study, high fluctuations were present in the TM3, TM5, TM6, TM7 and the extracellular loop 1 (ECL1), ECL2 for G protein-bound receptor. For β-arrestin-bound receptor, structural changes were mostly visible in TM1 and TM7. When comparing ligand biases between G-protein and β-arrestin-bound receptor states, it appeared to be that the ligands chosen showed higher bias towards G-protein-bound state.

In essence, the studies discussed here show how the chemical structure of a ligand can induce changes in the GPCR structure, and how throughout the time when the receptor is active or inactive, the non-covalent interactions between the binding pocket residues and the ligand changes. It is worth mentioning that ECL2 acts as a key player in stabilizing the GPCR receptor conformation as reported by Seibt et al. (2013), which the activity of ECL2 is also very prominent in the previously discussed studies.

2. AIM OF THE STUDY

As mentioned before, ET_B receptor is widely expressed in human body. Upon activation it releases vasodilators and expands the blood vessels, and the receptor may even act as inhibitor for both cell growth, and vascular constriction in the blood vessel system. The receptor is involved in several human diseases, such as pulmonary arterial hypertension (PAH), arterial hypertension (AH), renal diseases and more. Hence, ET_B targeting drugs were developed. However, some paired with several unwanted adverse effects (Davenport et al., 2016; Houde et al., 2016; Kawanabe & Nauli, 2011). Additionally, it is reported that some receptors can behave differently in different membrane environments (Angladon et al., 2019). Therefore, targeting the ET_B signalling system by means of different types of relevant computational methods can provide knowledge about the receptor and ligand interactions, behavioural insights when exposed to different biological environments, and may help with further drug development and discovery for the ET_B receptor-related medical disorders. However, there are far too few studies exploring the interaction between ET_B receptor and its ligands. Therefore, it may be interesting to explore the behaviour of ET_B-ligand complexes, when embedded in the well-known membrane environments: POPC (1-palmitoyl-2-oleoylphosphatidylcholine), POPE (1-palmitoyl-2-oleoylphosphatidylethanolamine), DPPC (1,2-dipalmitoyl-sn-glycero-3-phosphocholine), DMPC (1,2-dimyristoyl-sn-glycero-3-phosphocholine).

The aim of the study is to understand and explore the mechanistic behaviour of ET_B receptor in response to two antagonistic ligands: bosentan and K-8794, when embedded in POPC, DPPC, POPE and DMPC membrane environments, by means of Molecular Dynamics simulations.

The objectives of the study:

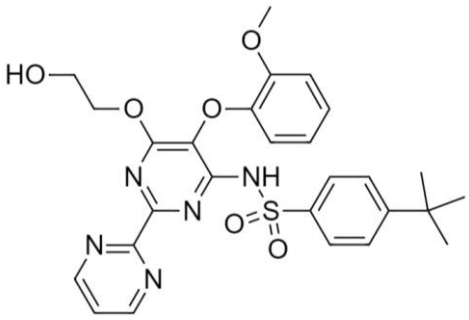
1. To analyse the scientific literature (previous chapter) about the various cell membranes, G-protein coupled receptors, ET_B receptor, the ET_B receptor targeting drugs and Molecular Dynamics simulations.
2. To analyse the mechanistic basis of bosentan binding to ET_B receptor embedded in POPC, DPPC, POPE and DMPC membranes.
3. To analyse the mechanistic basis of K-8794 binding to ET_B receptor embedded in POPC, DPPC, POPE and DMPC membranes.

3. MATERIALS AND METHODS

3.1. Overview of the structures

For the current project two ET_B structures were obtained from the Protein Data Bank (PDB) (Berman et al., 2000) in the inactive conformation bound to an antagonist bosentan (Figure 3.1A) at the resolution of 3.6 Å (ID: 5XPR), and bosentan's high-affinity ET_B selective analog K-8794 (Figure 3.1B) at the resolution of 2.2 Å (ID: 5X93) (Shihoya et al., 2017). Both ligands occupy the traditional orthosteric binding site of the ET_B. The residues involved in bosentan binding are Asp154, Gln181, Lys182, Lys273, Trp336 and Arg343 as reported by Shihoya et al. (2017). While K-8794 showed essentially the same interactions as bosentan, the analog interactions with His150 and Ser379 residues were also described. In addition, O=S=O of the sulfonamide group of both ligands interacts with the Arg343 that induces a slight inward movement of the extracellular region of TM6. Unlike ET-1, both bosentan and K-8794 interact poorly with TM7 and thus, do not induce any inward movement of that helix (Shihoya et al., 2017).

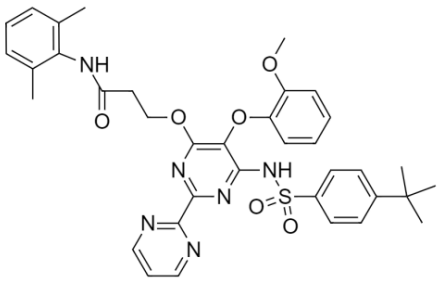
(A)



Bosentan

4-tert-butyl-N-[6-(2-hydroxyethyloxy)-5-(2-methoxyphenoxy)-2-pyrimidin-2-yl-pyrimidin-4-yl]benzenesulfonamide

(B)



K-8794

3-[6-[(4-tert-butylphenyl)sulfonylamino]-5-(2-methoxyphenoxy)-2-pyrimidin-2-yl-pyrimidin-4-yl]oxy-N-(2,6-dimethylphenyl)propanamide

Figure 3.1. 2D chemical structures of bosentan (A) and K-8794 (B) (Shihoya et al., 2017).

3.2. Preparation of the structures

In order to use the selected protein-ligand complexes for modelling and simulation purposes with the Maestro v2020.3 by Schrödinger LLC, the structures needed to be free from missing atoms, loops and any interatomic clashes. The initial preparation step before the usage of Maestro included the removal of the unnecessary protein regions that had been used for the crystallization process. Therefore, the T4-lysosyme (mT4L) segment present in the third intracellular loop (ICL3) of the receptor was removed from both receptor structures.

Apart from the mT4L, the K-8794 bound structure contained one cholesterol molecule, and other non-protein atoms such as five sulphate ions, eleven (2R)-2,3-dihydroxypropyl (9Z)-octadec-9-enoate (monoolein), and 106 water molecules, and the bosentan bound structure contained four sulphate ions.

Figure 3.2 shows the original protein sequences in FASTA format, taken from the PDB, with the highlighted residues of mT4L that was removed. Structure visualizations and deletion of unwanted residues was carried out using PyMOL Molecular Graphics System from Schrödinger LLC (*PyMOL | pymol.org*, n.d.).

(A)

```
>5XPR_1|Chain A|Endothelin B receptor,Endolysin,Endothelin B
receptor|Homo sapiens (9606)
GGGLAPAEVPKGDRTAGSPPRTISPPPCQGPIEIKETFKYINTVVSCLVFVLGIIGNSTLLYIIYKNK
CMRNGPNILIASLALGDLLHIVIDIPINVYKLLAEDWPFGAEMCKLVPFIQKASVGITVLSLCALSID
RYRAVASWSRIKGIGVPKWTAVEIVLIWVVSVVLAVPEAIGFDIITMDYKGSYLRICLLHPVQKTAFM
QFYATAKDWWLFSFYFCLPLAITAFFYTLMTCEMLRKNIFEMLRIDEGGGSGGDEAEKLFNQDVDAAV
RGILRNAKLKPVYDSLDAVRRAALINMVFQMGETGVAGFTNSLRMLQQKRWDEAAVNLAKSRWYNQTP
NRAKRVITTFRTGTWDAYLNDHLKQRREVAKTVFCLVLVFALCWLPLHLARILKLTLYNQNDPNRCEL
LSFLLVLDYIGINMASLNSCANPIALYLVSKRFKNAFKSALCCWAQSPSSENLYFQ
```

(B)

```
>5X93_1|Chain A|Endothelin B receptor,Endolysin,Endothelin B
receptor|Homo sapiens (9606)
GGGLAPAEVPKGDRTAGSPPRTISPPPCQGPIEIKETFKYINTVVSCLVFVLGIIGNSTLLYIIYKNK
CMRNGPNILIASLALGDLLHIVIAIPINVYKLLAEDWPFGAEMCKLVPFIQKASVGITVLSLCALSID
RYRAVASWSRIKGIGVPKWTAVEIVLIWVVSVVLAVPEAIGFDIITMDYKGSYLRICLLHPVQKTAFM
QFYATAKDWWLFSFYFCLPLAITAFFYTLMTCEMLRKNIFEMLRIDEGGGSGGDEAEKLFNQDVDAAV
RGILRNAKLKPVYDSLDAVRRAALINMVFQMGETGVAGFTNSLRMLQQKRWDEAAVNLAKSRWYNQTP
NRAKRVITTFRTGTWDAYLNDHLKQRREVAKTVFCLVLVFALCWLPLHLARILKLTLYNQNDPNRCEL
LSFLLVLDYIGINMASLNSCANPIALYLVSKRFKNAFKSALCCWAQSPSSENLYFQ
```

Figure 3.2. FASTA sequences of PDB ID: 5XPR (A) and 5X93 (B) crystal structures of the ET_B receptor. The grey highlighted regions are the mT4L segments that were removed during the protein preparation

Upon removal of the mT4L, the unwanted gaps in the ET_B structure at ICL3 were modelled using the widely utilized homology modelling tool MODELLER v9.24 (Fiser et al., 2000; Martí-Renom et al., 2000; Šali & Blundell, 1993; Webb & Sali, 2016) based on user defined sequence alignment (Figure 3.3) between the template (X-ray structure sequence) and the target (N and C-terminal truncated ET_B sequence from the UniProt database (ID: P24530)). The alignment was performed using the Clustal Omega tool by European Molecular Biology Laboratory – European Bioinformatics Institute (EMBL-EBI) (*Clustal Omega < Multiple Sequence Alignment < EMBL-EBI*, n.d.; Madeira et al., 2019)

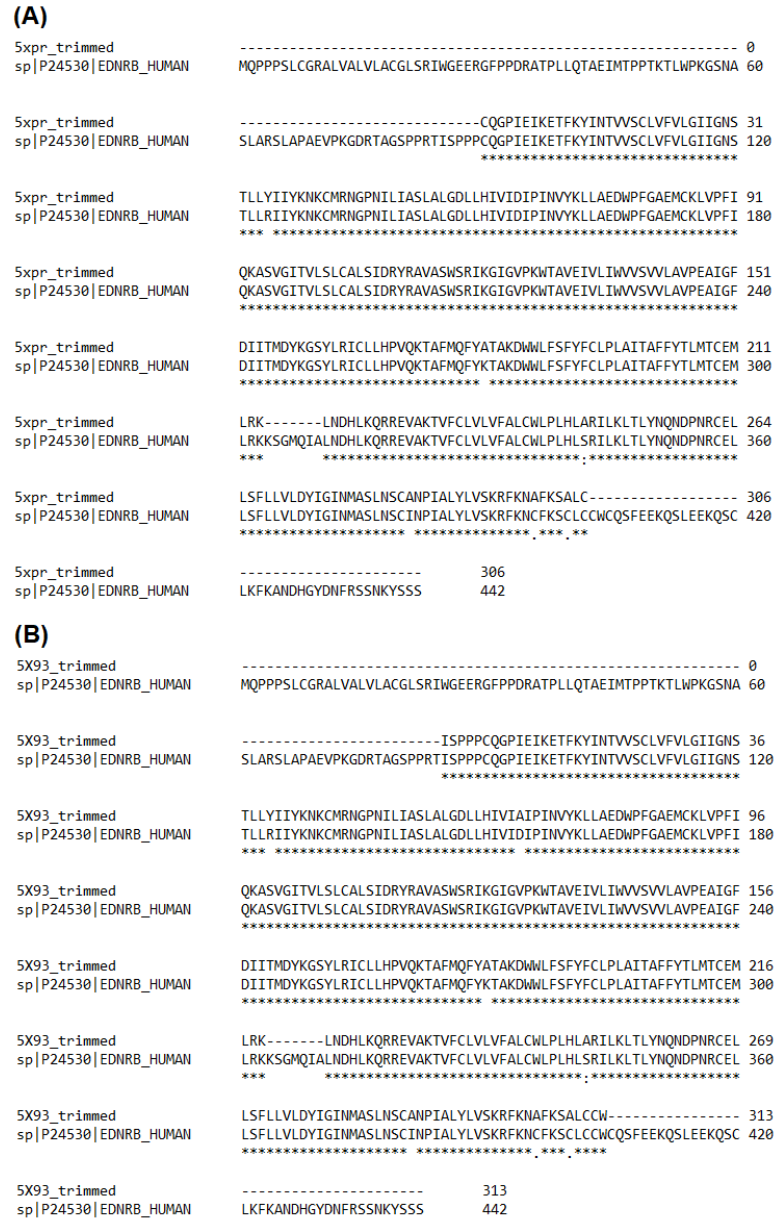


Figure 3.3. Sequence alignment of the crystal structure sequences (without the mT4L segment) (A) PDB ID: 5XPR and (B) 5X93 mT4L free and the ET_B UniProt sequence (ID: P24530). The crystal structure sequences served as a template, and the UniProt sequence as a target.

The MODELLER tool works based on a Python script (Figure A1) that generated 20 separate models. The best model was selected based on Discrete Optimized Protein Energy (DOPE) score (Shen & Sali, 2006). DOPE is a statistical potential optimized for model assessment, and is part of the MODELLER scoring module. Models with a low DOPE score are considered to be of higher quality (Shen & Sali, 2006). The stereo-chemical quality of the model was accessed using Maestro *Structural Quality Report* tool (Bowers et al., 2007) that calculates φ (phi) and ψ (psi) angles of each residue in the protein structure, and plots the residues on the allowed and disallowed regions on the Ramachandran map.

Out of all generated models one representative (one of bosentan-bound and one of K-8794-bound) model for each complex was selected based on the lowest DOPE score, and the least number of residues falling in the disallowed regions of the Ramachandran plot. Overall, the created models include only the ET_B receptor and bound ligands, and their conformation is close to X-ray structure when superimposed with PyMOL by aligning the Cα atoms of the receptor backbone. Table A1 shows all the generated models and their DOPE scores.

Prior to running the Molecular Dynamics (MD) Simulations, the 3D structure was prepared using *Protein Preparation Wizard* (Madhavi Sastry et al., 2013). During the process the missing hydrogen atoms were added, incomplete side chains and loops were modelled, ambiguous protonation states, and flipped residues were visually investigated and manually adjusted if necessary (Madhavi Sastry et al., 2013).

3.3. Molecular Dynamics Simulation Details

After the structure preparation, the protein-ligand complexes were inserted into the four different lipid bilayers (membranes) (Figure 3.4), covering the transmembrane region of the protein using the *System Builder* tool (Bowers et al., 2007).

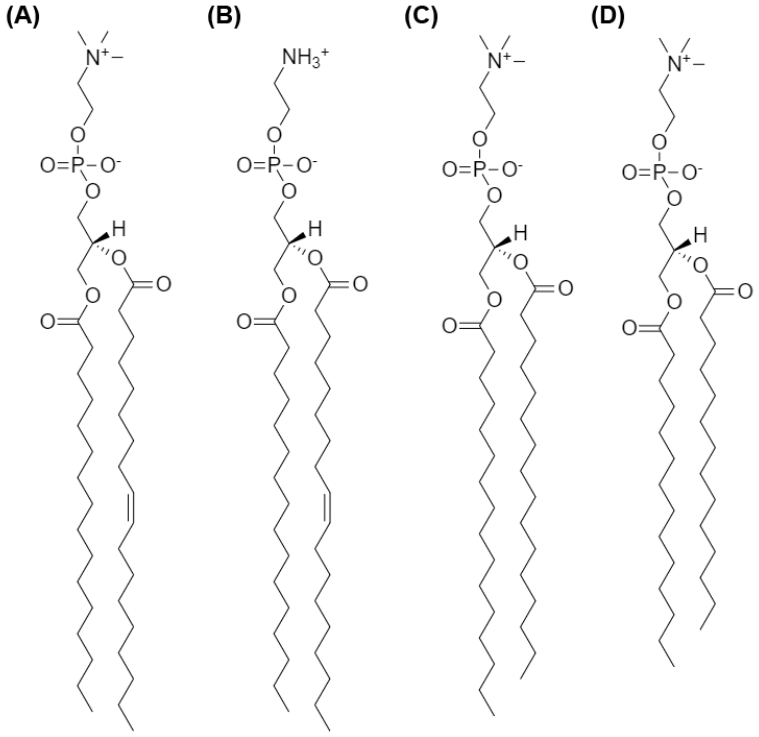


Figure 3.4. 2D chemical structures of phospholipids. (A) POPC (1-palmitoyl-2-oleoylphosphatidylcholine), (B) POPE (1-palmitoyl-2-oleoylphosphatidylethanolamine), (C) DPPC (1,2-dipalmitoyl-sn-glycero-3-phosphocholine), (D) DMPC (1,2-dimyristoyl-sn-glycero-3-phosphocholine) (Shahane et al., 2019).

The rest of the parameters were set to default as follows: the protein-ligand complex was placed inside the orthorhombic box measuring 10 x 10 x 10 Å in all directions. The system was solvated by the TIP3P solvent model, a simple 3-site water model (Jorgensen et al., 1983). The force field applied was OPLS3e (Roos et al., 2019) and the unbalanced charge in the system was neutralized by adding counter ions. The structures embedded in the different membranes produced 16 independent systems (8 apo – ligand-free and 8 holo – ligand-bound). Subsequently, all the systems were subjected to the *Molecular Dynamics Panel* by Desmond Molecular Dynamics System developed by D. E. Shaw Research (Bowers et al., 2007), included in Maestro v2020.3 to perform the simulations. The simulations were carried out in multiple stages using the parameters listed in Table 3.1.

Table 3.1. MD Simulation parameters and protocol.

Simulation time period (ns)	400	
Recording interval (ps)	Trajectory: 50	Energy: 1.2
Number of frames	10 000	
Ensemble class	NPγT	
Temperature (K)	300	
Pressure (bar)	1.01325	
Surface tensions (bar Å)	0	
System relaxation protocol	Default NPγT	
Hardware	Puhti supercomputer (Atos BullSequana X400 cluster based on Intel CPU) at CSC – IT Center for Science (Finland)	

Before the actual production run, the system was subjected to a series of step-by-step relaxation processes. Firstly, the system was simulated using Brownian Dynamics (BD) in NVT ensemble at 10 K temperature for 100 ps by restraining solute heavy atoms. This was followed by BD simulation in NPT ensemble at 100 K for 20 ps, with solute non-hydrogen atoms restrained. Afterwards, an MD simulation of 10 ps in NPγT ensemble was carried out, using a Nose–Hoover chain thermostat (Martyna et al., 1992) and Martyna–Tobias–Klein barostat (Martyna et al., 1994) set to 100 K and 1.01325 bar pressure, while restraining non-hydrogen solute atoms. Next, an MD simulation was run for 10 ps in NPγT ensemble, heating from 100 to 300 K by using Nose–Hoover chain thermostat and Martyna–Tobias–Klein barostat. Subsequently, while gradually releasing the restrains the system was subjected to an MD simulation in NVT ensemble for 24 ps, maintaining the temperature at 300 K with the Nose–Hoover thermostat and the pressure at 1.01325 bar with the Martyna–Tobias–Klein

barostat. Finally, the unrestrained system was simulated again in NVT ensemble (Bowers et al., 2007).

For the production simulation, the NPγT ensemble class was chosen, to maintain the surface tension of the system, and to ensure that the simulation box does not deform significantly in the plane of the membrane while the pressure is applied normal to the membrane surface (Bowers et al., 2007, *Molecular Dynamics Panel*, n.d.). All the MD simulations were carried out using the Finnish IT Center for Science (CSC) high performance computing facility.

3.4. Preliminary Trajectory Analysis

Several preliminary trajectory analyses were performed to acquire details of the structural stability of the complexes. The trajectory analysis for each system was carried out using the *Simulation Interactions Analysis* tool (Bowers et al., 2007), from which the Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF) and Radius of Gyration (Rg) values were estimated, which is a fully automated process. In addition, the total number of hydrogen bonds was calculated through the *Interaction Count* tool. All the analyses mentioned were performed on the entire length of the MD trajectories.

RMSD measures the average change in displacement of Cα atoms, for a particular frame compared to the reference structure (the first frame of the simulation) over a set of atoms. The RMSD is calculated based on the following equation for the entire trajectory (Bowers et al., 2007):

$$\mathrm{RMSD}_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(r_i'(t_x)\right) - r_i(t_{ref}))^2} \quad (1)$$

where, x is the frame to calculate the RMSD for; N is the number of atoms in the atom selection; t_{ref} is the reference time (frame); r' is the position of the selected atoms in the x frame after superimposing on the reference frame.

RMSF measures the structural flexibility of Cα atoms along a protein chain. The N-terminus and C-terminus ends, and loops of the protein show higher fluctuations, while rigid parts of the protein (helices and sheets) fluctuate less. RMSF is calculated by the following equation (Bowers et al., 2007):

$$\mathrm{RMSF}_i = \sqrt{\frac{1}{T}\sum_{t=1}^{T} < (r_i'(t)) - r_i(t_{ref}))^2 >} \quad (2)$$

where T is the trajectory time over which the *RMSF* is calculated; t_{ref} is the reference time (frame); r_i is the position of the residue i; r' is the position of atoms in residue i after

superimposition on the reference frame; angle brackets < and > indicate that the average of the square distance is taken over the selection of atoms in the residue.

Radius of Gyration (R_g) measures the compactness of a structure and the evolution of it throughout the simulation. It is calculated by using the following equation (Bowers et al., 2007):

$$R_g = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(r(i) - r_{center})^2} \quad (3)$$

where N is the number of protein atoms; r_i and r_{center} are the position of atom i with respect to the centre of mass of the molecule.

Hydrogen Bond Count measures the number of hydrogen bonds that are present between the protein and its binding partner over time. Hydrogen bonds contribute to the stability of the complex (longer ligand residence times) and binding site rigidity (Torrens-Fontanals et al., 2020). Parallel to the Hydrogen Bond Count measurements, hydrogen bonds of the last trajectory frame (at 400 ns) for every ligand-bound system were visually investigated.

3.5. Principal Component Analysis

An essential dynamics analysis based on the Principal Component Analysis (PCA) was carried on all the MD trajectories by running the *trj_essential_dynamics.py* (Amadei et al., 1993) program on all the MD trajectories available in Maestro v2020.3 by Schrödinger LLC. The analysis performed on all the snapshots obtained from the last 200 ns (from 200 ns to 400 ns) for 16 trajectories (both ligand-bound complexes and apo structures). The PCA analysis produced 10 principal components (PCs modes) for every trajectory.

In PCA, a covariance matrix (C_{ij}) is built, using the coordinates of Cα atoms. Then the eigenvalue decomposition of the C_{ij} matrix is used to obtain eigenvectors (principal components) of the MD trajectories. The following equation is used to create the C_{ij} matrix (Amusengeri et al., 2020; Chen et al., 2016):

$$C_{ij} = (X_i - \overline{X}_\iota)(X_j - \overline{X}_j) > (i, j = 1,2, \ldots., 3N) \quad (4)$$

where X_i is the Cartesian coordinate of the ith Cα atom, N is the Cα atom used in building the matrix.

The eigenvectors are obtained by calculating the diagonal matrix Q using the following equation (Amusengeri et al., 2020; Chen et al., 2016):

$$Q = T^T C_{ij} T \quad (5)$$

where T is the orthogonal coordinate transformation matrix, and C_{ij} is the covariance matrix.

Additionally, the first principal component (PC1) was presented as a porcupine plot by visualising the secondary structure and loops of the protein using PyMOL, to analyse structural changes and direction of motion captured by the PC of the ET_B receptor. In the porcupine plot, the black arrows represent the direction of motion of every Cα atom, while the length represents the strength of the movement (Chen et al., 2016).

3.6. Free Energy Landscape

The first and the second principal component (PC1 and PC2) were plotted as Free Energy Landscapes (FEL). FEL assists the visualization and effortless study of the structural clusters (conformers) formed throughout the simulation and/or for the purpose of understanding the structural stability of a protein. This method also facilitates the discovery of the accessible conformational space of a protein to demonstrate the flexibility of the complex systems (Amusengeri et al., 2020; Chen et al., 2016). For the purpose of the project, the first two PCs were chosen because they contain the largest percentage of variance of the MD simulations. The FEL was calculated using the following equation (Amusengeri et al., 2020; Chen et al., 2016):

$$\Delta G_i = -k_B T[lnP(i) - lnP_{max}] \quad (6)$$

where k_B is the Boltzmann constant, T is the temperature set to 300 K, i are principal components, P is the probability distribution of the molecular systems along some coordinate i, $Pmax$ ensures that the lowest free energy minima is set to 0.

3.7. Binding free energy calculations by MM/GBSA

For the binding free energy (BFE) calculations the Molecular Mechanics with Generalized Born Surface Area (MM/GBSA) method was utilized (Bowers et al., 2007). The method's purpose is to predict the strength of interactions between a protein and its ligand by calculating BFE. A total of 100 evenly spaced snapshots from the last 200 ns of trajectories of the ligand-bound simulations were considered.

The total binding free energy of a receptor-ligand complex, and the contribution of different interaction components to the total energy was also calculated. The latter was done to evaluate the variables that affect the total energy, and their contribution to the total binding free energy. Additionally, per-residue decomposition energies for all protein residues were calculated. It was done to detect the hotspot residues that have the highest contribution

(binding strength) to the total binding free energy of the ligand binding to the receptor. In addition, the residues with a BFE of -2 kcal/mol and lower were visualised as a 3D image taken from the snapshot structure at 300 ns of the simulation. However, the residues identified and visualised must be the same throughout all of the simulations.

The MM/GBSA method was applied by running the *thermal_mmgbsa.py* script available in the Schrödinger's Maestro suite, where it incorporates the Prime module (Bowers et al., 2007). This method estimates the free energy of binding of a receptor-ligand complex, separate ligand or just receptor structures (Ylilauri & Pentikäinen, 2013). As described by Peddi et al. (2018), Prime/MMGBSA is binding free energy (ΔG_{bind}), and can be viewed as an equation at Eq. 8 in more detail:

> A combination of OPLS molecular mechanics energies (EMM), a VSGB solvation model for polar solvation (GSGB), and a nonpolar solvation term (GNP) composed of the non-polar solvent-accessible surface area (SASA) and van der Waals interactions (p. 9).

The equation for calculating the binding free energy (ΔG_{bind}) is as follows (Schrödinger LLC, 2021; *What do all the Prime MM-GBSA energy properties mean? | Schrödinger*, n.d.):

$$\Delta G_{\text{bind}} = G_{complex} - (G_{ligand\ (minimized)} + G_{receptor\ (minimized)}) \qquad (7)$$

where G is the energy of a complex, ligand and receptor.

Eq. 7 can be rewritten based on Lyne et al. (2006) as:

$$\Delta G_{\text{bind}} = \Delta E_{MM} + \Delta G_{solv} + \Delta G_{SA} \qquad (8)$$

where ΔE_{MM} is dissimilarity in energies between the complex structure and the sum of the energies of the ligand and the apo structure (7) or sum of electrostatic (E_{ele}), lipophilic (E_{lipo}) and van der Waals (E_{VdW}) energies; ΔG_{solv} (solvation free energy) is the difference in the GBSA solvation energy of the receptor-ligand complex and the sum of the solvation energies for the ligand and apo receptor; ΔG_{SA} is the difference in the surface area energy for the receptor-ligand complex and the sum of the surface area energies for the ligand and apo receptor.

Additionally, the total energy with the contributions of the different components is calculated with the equation Eq. 9 presented below, as reported by Li et al. (2011):

$$\Delta G_{\text{total(bind)}} = E_{coulomb} + E_{covalent} + E_{Hbond} + E_{lipo} + E_{packing} + E_{GB} + E_{VdW} \qquad (9)$$

that generally includes Coulomb energy, covalent binding energy, hydrogen-bonding correction, lipophilic energy, π-π packing correction, generalized Born (GB) electrostatic solvation energy, and van der Waals energy, respectively (Li et al., 2011).

3.8. Dihedral Angle Analysis

Dihedral angle analysis can be used to identify the conformational change that occurred in a specific residue due to ligand binding over a time period (Chen et al., 2016). A dihedral angle can be described as the angle between two adjacent planes in the polypeptide backbone. The dihedral angle consists of φ (phi) and ψ (psi) angles. The φ angle involves the backbone atoms of C-N-Cα-C, while the ψ angle involves the N-Cα-C-N atoms (Echave, 2009).

The dihedral angle analysis was performed on the residues that have the highest contribution to the total BFE (hotspot residues). For that purpose, *Add Workspace Measurement* tool present in Maestro was utilized to measure the dihedral angle considering both apo and holo conformations, using the last 200 ns (200 ns – 400 ns range) of the trajectories. To visualise the change in dihedral angle, FELs were constructed by plotting the φ and ψ angles.

3.9. Figure plotting and graphical visualisation

Overall, the protein structure analysis and visual investigation was performed using PyMOL and Maestro v2020.3 by Schrödinger LLC graphical user interface (GUI).

All the graphs were plotted using Python libraries developed for data analysis, such as NumPy, Pandas and Matplotlib.

4. RESULTS AND DISCUSSION

Firstly, the MODELLER software created 20 structures per ligand, as seen in Table A1, and automatically calculated its model assessment scores based on DOPE method (Fiser et al., 2000; Martí-Renom et al., 2000; Šali & Blundell, 1993; Webb & Sali, 2016). The chosen bosentan-bound ET_B receptor model had a DOPE score of -40353.57031 and its residues were 97.6 % in the most favoured regions in the Ramachandran plot. For the K-8794-bound ET_B model, the chosen model had a DOPE score of -42179.47266 and residues were 98.6 % in the most favoured regions.

After the initial system preparations (apo and holo complexes), the systems were subjected to MD simulation, embedded in four different membranes (POPC, DPPC, POPE, DMPC). A total of 16 molecular dynamics simulation trajectories were obtained:

- 4 for bosentan-bound ET_B;
- 4 for bosentan-free ET_B (APO);
- 4 for K-8794-bound ET_B;
- 4 for K-8794-free ET_B (APO).

Each system was analysed to understand the structural stability, mechanistic behaviour, and contributions to the binding free energy due to ligand binding to ET_B receptor. The outputs from the final MD simulations (protein + membrane) will be referred to as LIGAND-RECEPTOR-MEMBRANE system, and for APO as LIGAND-free-RECEPTOR-MEMBRANE system.

4.1. Structural stability analysis

4.1.1. Root Mean Square Deviation (RMSD) analysis

RMSDs for all of the eight ligand-bound ET_B complexes and eight ligand-free ET_B (APO) structures were calculated. RMSD plots revealed structural deviations of Cα atoms in the four different types of membranes, for the previously mentioned complexes and structures.

As seen in Figure 4.1, all 16 systems exhibited fluctuations between 2 to 3 Å. Most APO systems displayed high fluctuations, especially the K-8794-free-ET_B-DPPC system. Some displayed stability throughout all of the MD simulations (POPE and DMPC). The deviations for APO were expected, because of an empty orthosteric binding pocket, since there is no ligand to induce or inhibit the receptor activity.

Bosentan-ET_B-POPC system showed medium fluctuation: it first reached equilibrium after 100 ns, and then later after 200 ns. While the K-8794-ET_B complex embedded in POPC showed a very high deviation for the first 150 ns and started to maintain stability at around 160 ns. The Bosentan-ET_B in the DPPC system maintained stability throughout all of the MD simulation time. Similar behaviour was observed in the K-8794-ET_B-DPPC system, as it also maintained stable equilibrium during the simulation. Next, a relatively high instability is seen in the Bosentan-ET_B-POPE system for the first 200 ns, but during the rest of the simulation time, the system has attained equilibrium by staying at around the 3.5 Å range. K-8794-ET_B-POPE system at the start of the simulation exhibited a slight deviation of about 80 ns and maintained stability for the rest of the MD simulation. Both Bosentan-ET_B-DMPC and K-8794-ET_B-DMPC retained a stable plateau for all 400 ns.

These findings suggest that the trajectory of the last 200 ns (range of 200 ns – 400 ns) is reliable and can be used for further analysis, such as free binding energy estimation, Principal Component Analysis and dihedral angle analysis.

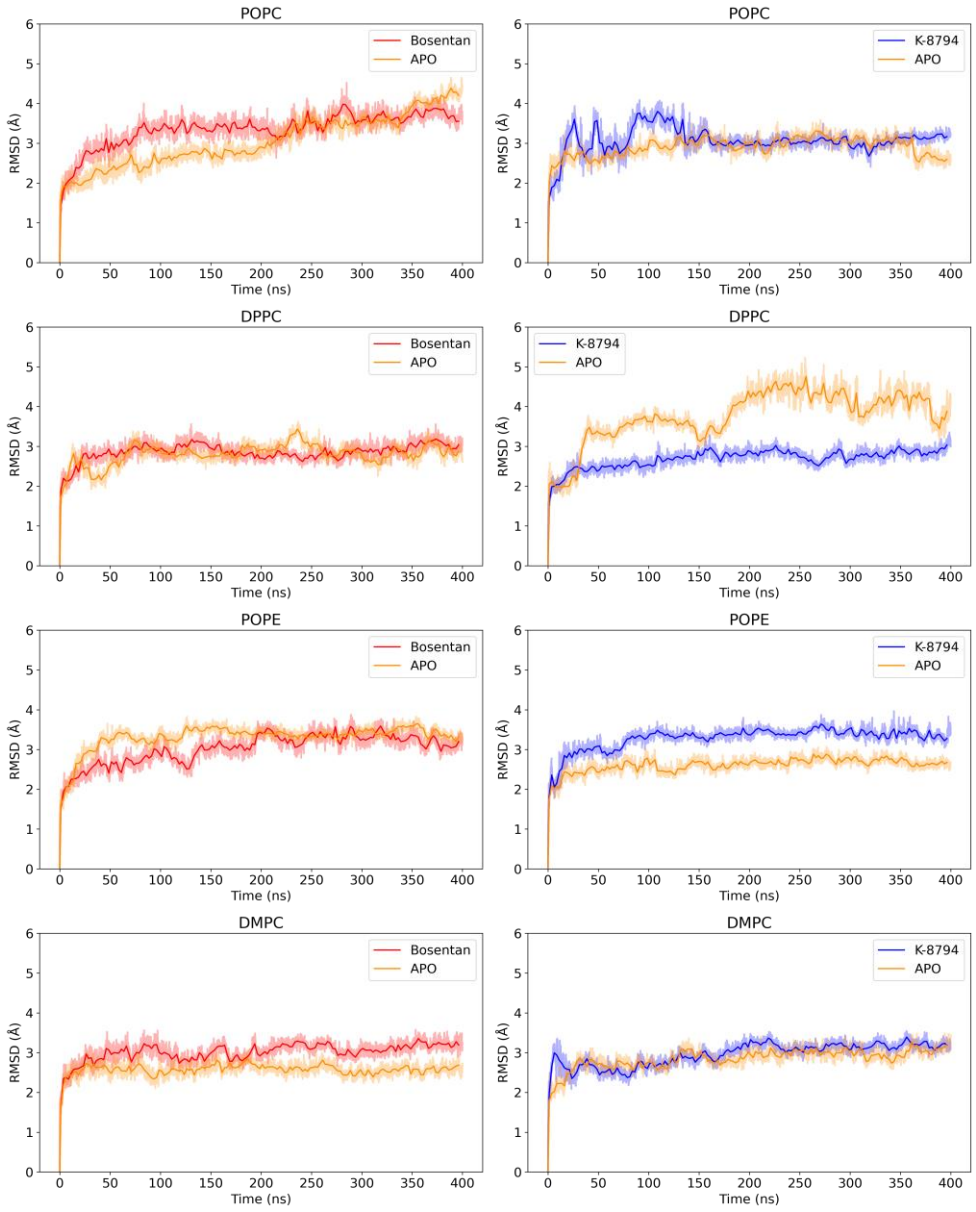


Figure 4.1. Root mean square deviation (RMSD) estimated for the Cα atoms throughout 400 ns for the bosetan (red) and K-8794-bound (blue) to the ET_B receptor, and ligand-free (yellow) states embedded into the POPC, DPPC, POPE, and DMPC membranes. Solid lines represent an average of every 50 frames (50 frames per 1 ns).

4.1.2. Root Mean Square Fluctuation (RMSF) analysis

RMSF for all of the eight ligand-bound ET_B complexes and eight ligand-free ET_B (APO) structures were calculated. RMSF has revealed structural flexibility of Cα atoms in the four different types of membranes, for the previously mentioned complexes and structures.

All 16 systems exhibited fluctuations as shown in the RMSF Figure 4.2. After a further 3D structure and RMSF analysis, it was revealed that the six peaks represent the most

fluctuating areas of the protein. These areas were the three extracellular and three intracellular loops, while the helices and sheets were more rigid and did not show high fluctuations. Additionally, the sudden slopes at the beginning and the end of the RMSF plots represent the N-terminus and the C-terminus of the receptor that were expected to fluctuate. Loops are more flexible because they are the region of the protein that is free from structural integrity. Comparing all of the ligand-bound ET_B systems with their APO equivalent, it is clearly seen that the APO structure deviated the most. This actively demonstrates that protein without inhibitor bound is more flexible, compared to the effector bound. However, when an effector binds to the binding pocket of the protein, this instigates the movement of intracellular (IC) loops trapping the effector in a cage-like conformation (Kempner, 1993).

Interestingly enough, in most of the systems (Figure 4.2), except in the K-8794-ET_B-POPE and K-8794-ET_B-DMPC systems, the highest fluctuations reaching 6 Å was of ICL3 (Lys303 – Asn312) that connects TM5 to TM6. In all systems, the least fluctuations starting at 1.5 Å were of intracellular loop 1 (ICL1) (Asn129 – Leu139) that connects TM1 to TM2. Except in the Bosentan-ET_B-POPC system, ECL1 (Leu162 – Ala171) had the least deviations being at 2.1 Å, compared to other ligand-bound ET_B systems.

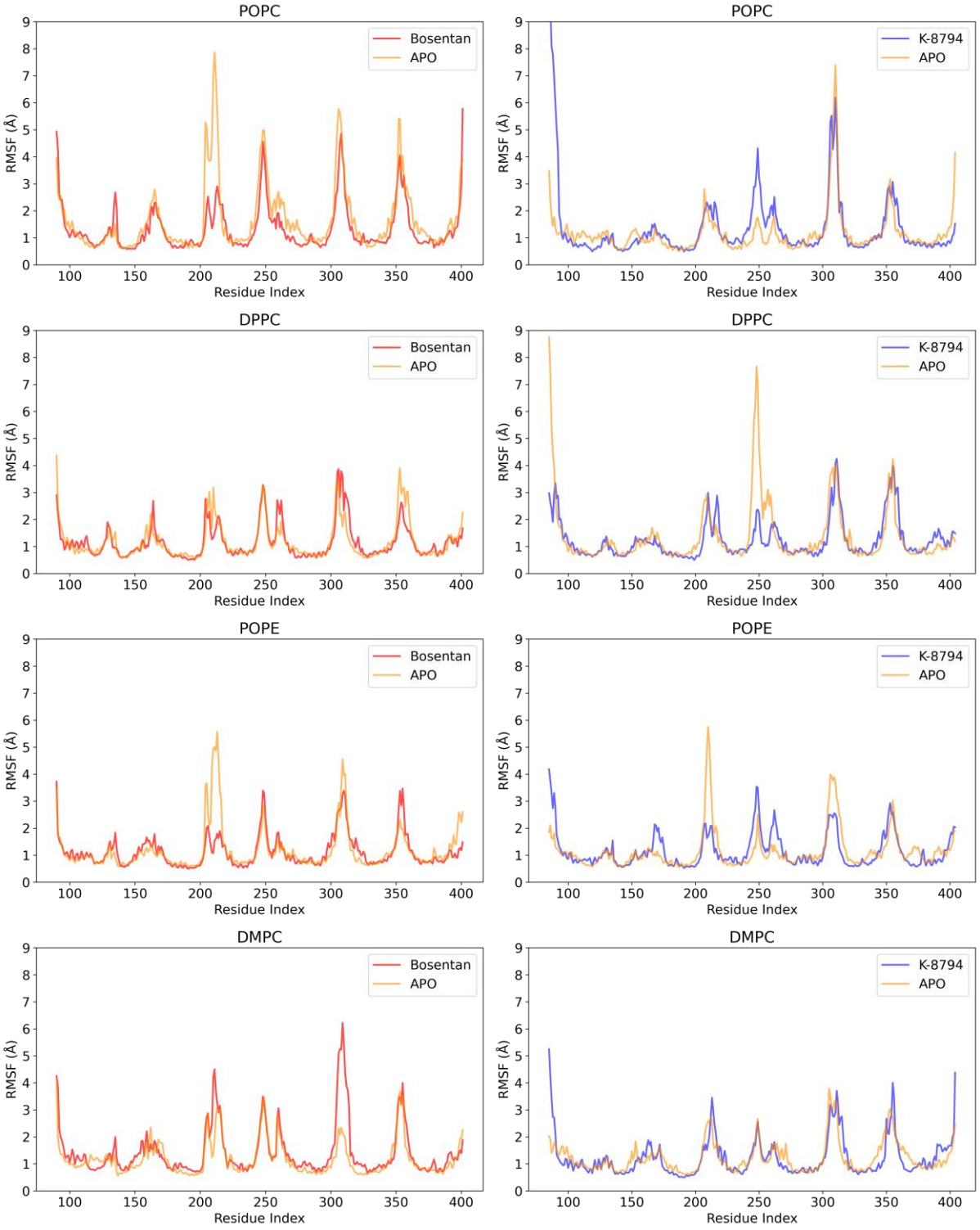


Figure 4.2. Root mean square fluctuations (RMSF) estimated for the Cα atoms throughout 400 ns for the bosetan (red) and K-8794-bound (blue) to the ET_B receptor, and ligand-free (yellow) states embedded into the POPC, DPPC, POPE, and DMPC membranes. Solid lines represent an average of every 50 frames (50 frames per 1 ns).

4.1.3. Radius of Gyration (R_g) analysis

The Radius of Gyration (R_g) for all the 16 complexes was calculated, to analyse the compactness and stability of the protein along the simulation time. More compact structures display smaller values (Ivankov et al., 2009).

All 16 systems experienced significant differences in the fluctuations as seen in Figure 4.3. Bosentan-ET_B-POPC exhibited fluctuations between 22.40-23.00 Å, and its APO showed similar compactness. On the other hand, K-8794-ET_B-POPC showed high fluctuations at the start of the MDS but later seemed to maintain compactness starting from ~150 ns between ~22.00-22.25 Å. Comparing the two systems mentioned previously, the K-8794-bound receptor maintained higher compactness. While its APO equivalent maintained a lower degree of compactness throughout most of the simulation. The Bosentan-ET_B-DPPC system maintained compactness between ~22.00-23.00 Å, similarly to its APO structure. K-8794-ET_B-DPPC exhibited compactness between 50 to 230 ns at ~22.25-22.50 Å and then later showed plateau from 230 to 400 ns at around 22.20 Å, which meant that the structure over the simulation time stayed relatively compact. However, between the two DPPC systems, K-8794-bound protein was more compact than bosentan-bound. Both bosentan-bound and K-8794 bound structures embedded in the POPE system showed similar levels of compactness to their APO equivalent. Comparing the Bosentan-ET_B-POPE system with the K-8794-ET_B-POPE, the former experienced a higher degree of compactness ranging at ~22.40 Å. While the latter was at ~22.50 Å. The Bosentan-ET_B-DMPC and K-8794-ET_B-DMPC systems expressed similar compactness levels, with bosentan-bound showing higher fluctuations at the beginning of the simulation. However, the following exhibited better compactness at around 22.30 Å from 230 to 240 ns, compared to K-8794-bound ET_B with lower compactness at around 22.40 Å throughout the MDS.

Most of the APO structures in all of the MD simulations showed similar deviations to their ligand-bound equivalent structures. Except for APO structure in POPC membrane as seen in Figure 4.3 experienced higher fluctuations later in the simulation, than the ligand-bound structure: APO form showed less compactness than ligand-bound form.

Out of all the 16 systems, the most compact were the K-8794-ET_B-POPC and K-8794-ET_B-DPPC systems. The least compact systems were Bosentan-ET_B-POPC and K-8794-ET_B-POPE. In addition, higher fluctuations showed bigger conformational changes.

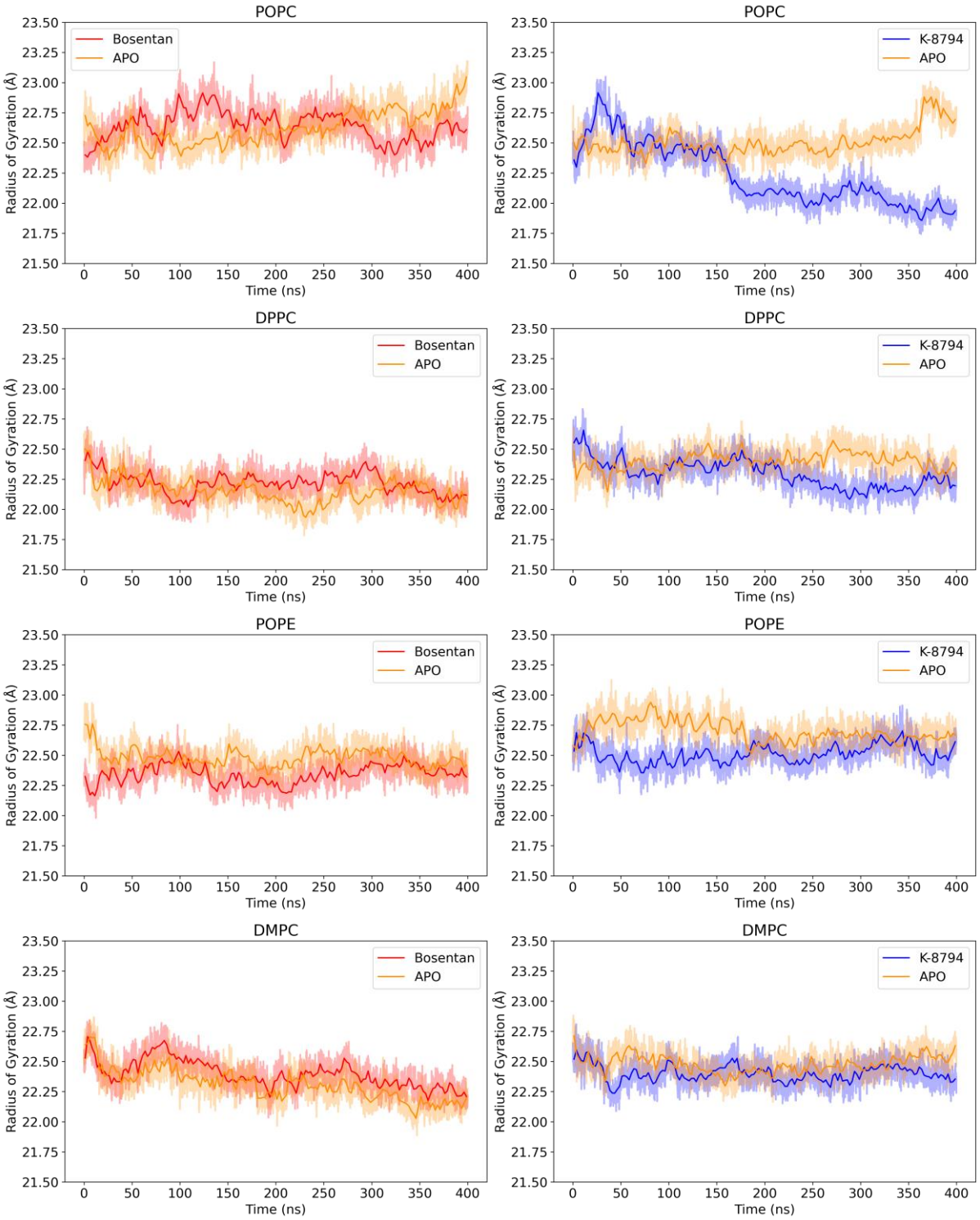


Figure 4.3. Radius of Gyration (R_g) plots of the protein structure compactness throughout 400 ns for bosetan-bound (red) and K-8794-bound (blue) ET_B receptor, and ligand-free (yellow) states embedded into the POPC, DPPC, POPE, and DMPC membranes. Solid lines represent an average of every 50 frames (50 frames per 1 ns).

4.1.4. Hydrogen Bond Count analysis

Ligands usually bind to the protein by non-covalent bonds such as hydrogen bonds. Hydrogen bonds (H-bonds) are important for ligand binding specificity, metabolization and adsorption. These bonds also contribute to the binding site rigidity, and stability of the complex (longer ligand residence times) (Torrens-Fontanals et al., 2020). Therefore, H-bond Count between binding residues and ligand was calculated for all of the 8 MD simulations. Additionally, the H-bonds between ligand and protein orthosteric binding pocket residues (Asp154, Ly161, Gln181, Lys182, Lys273, Arg343, Trp336, Ile372) at 4 Å distance of the last trajectory frame (400 ns) for every ligand-bound system were visualised.

The evolution of the number of hydrogen bonds in time is visualised in Figure 4.4. All the 8 MD simulation systems exhibit clear fluctuations in the number of H-bonds. Bosentan-ET_B-POPC systems experienced as high as 7 hydrogen bonds. However, it later exhibited a low in the number of H-bons between 100-150 ns averaging at ~2. Starting from ~150 ns the system reached a plateau averaging at around 4 hydrogen bonds, with the last frame having ~3 stable bonds. K-8794-ET_B-POPC experienced a very high number of H-bonds: rising at around 6 bonds (90-150 ns), then dropping to ~5 bonds lasting for 150 ns, and then dropping again to 2 but starting to rise at ~310 ns, with the final frame stabilizing at 4 bonds (Figure 4.4). Bosentan-ET_B and K-8794-ET_B structures embedded in DPPC membrane, displayed relatively stable bonds averaging at 4 for bosentan-bound ET_B, and after 50 ns for K-8794 system maintaining 2 bonds. The last frame had 4 stable bonds for bosentan-bound and 3 stable hydrogen bonds for K-8794-bound. The Bosentan-ET_B-POPE system also displayed low hydrogen bond deviations with an average of 3 hydrogen bonds. However, the K-8794-ET_B-POPE system experienced high fluctuations for ~150 ns, but reached equilibrium at ~160 ns and remained stable displaying 4 hydrogen bonds. The bosentan-ET_B in the DPPC membrane experienced uneven fluctuations with a sudden drop to ~1 hydrogen bond at 180-230 ns interval, then suddenly rising with visible fluctuations and stabilizing at 340 ns with 3 hydrogen bonds. However, K-8794-ET_B-DMPC exhibited high fluctuations for ~200 ns reaching 8 bonds but eventually stabilized during the last 200 ns having 3 hydrogen bonds.

Throughout all the systems, the hydrogen bond count change analysis revealed that the average stable hydrogen bond count was 3. Additionally, all the systems reached their equilibrium starting from 200 ns that indicates the 200-400 ns interval is a reliable range for further analysis.

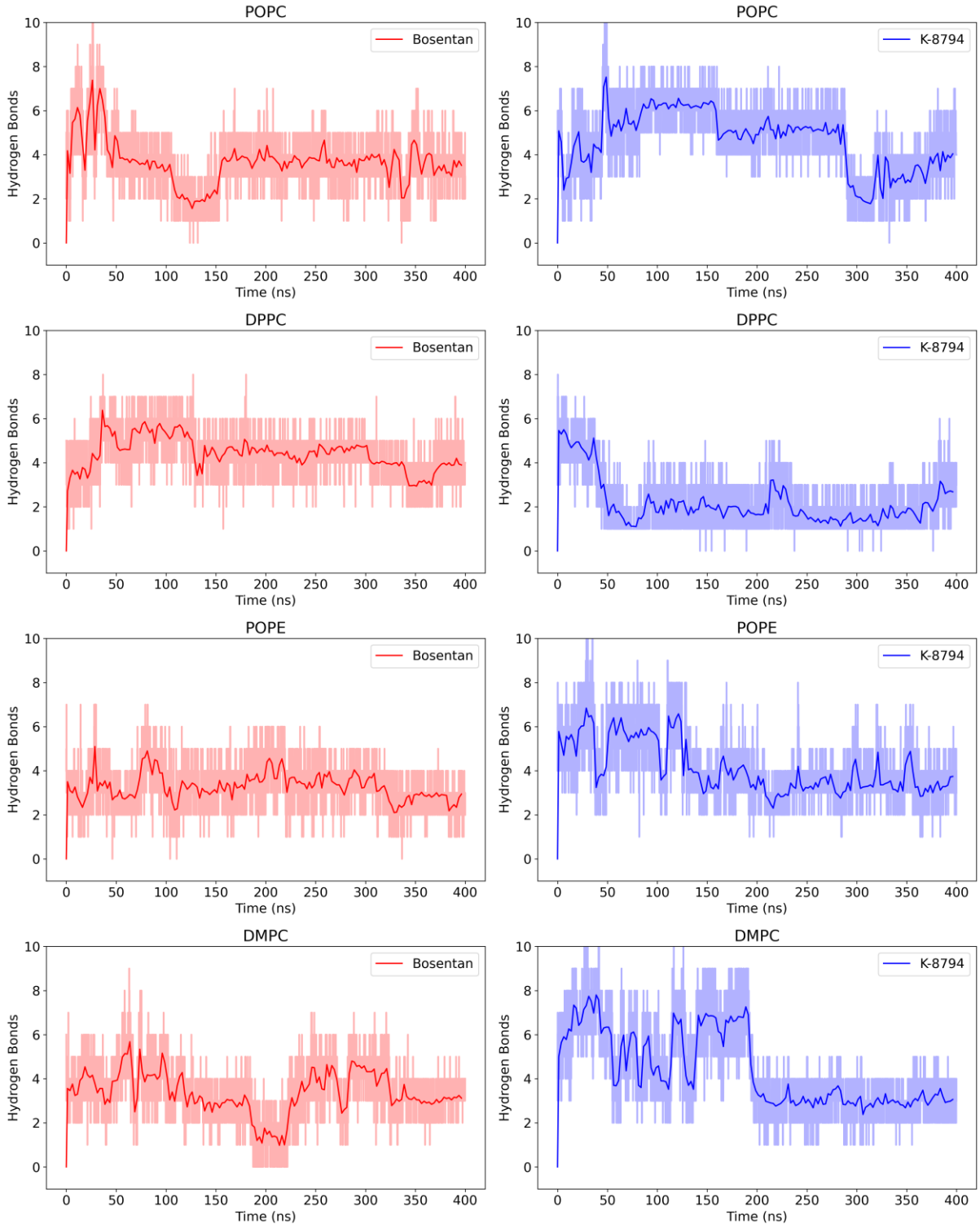


Figure 4.4. Hydrogen Bond plots indicating the change in number of Hbonds between protein and ligand throughout 400 ns for bosetan-bound (red) and K-8794-bound (blue) ET_B receptor embedded into the POPC, DPPC, POPE, and DMPC membranes. Solid lines represent an average of every 50 frames (50 frames per 1 ns).

Further analysis of the receptor-binding site at 4 Å distance from both ligands revealed the residues that contributed to the stable hydrogen bonds between ligand and ET_B receptor (Figure 4.5). For this, the last frame from the MD trajectory was extracted and used.

ET_B in complex with bosentan in POPC membrane (Figure 4.5 A) formed H-bonds between Gln181, Lys273, and Arg343. Similarly, hydrogen bonds are stabilized in ET_B in complex with K-8794 in POPC (Figure 4.5 B), however, instead of maintaining a bond with Gln181, a new hydrogen bond was established between K-8794 and Lys182, and an additional H-bond with Ile372. In the DPPC system bosentan maintained hydrogen bonds with Asp154, Lys182, Lys273 and, Arg343, while K-8794 with Lys161, Lys273 and Arg343 (Figure 4.5 C & D). ET_B in complex with bosentan in POPE membrane shows the hydrogens bonds formed by Lys273, Arg343, and Ile372 are stably maintained (Figure 4.5 E). While as shown in Figure 4.5 F, K-8794 formed hydrogen bonds with Lys182, Lys273, and Arg343 formed two bonds between one nitrogen atom and one oxygen atom. Both bosentan and K-8794 in the DMPC membrane system formed hydrogen bonds with the same ET_B binding pocket residues. The bosentan – with Lys182, Lys273, and Arg343 between the oxygen atoms of O=S=O group as shown in the Figure 4.5 G. K-8794 formed H-bonds between oxygen atoms of O=S=O group with Lys273 and Arg343, while Lys182 formed a hydrogen bond with the oxygen, which connects benzene ring to pyrimidine (Figure 4.5 H).

Ultimately, in all of the eight ligand-receptor-membrane systems, both ligands exhibited H-bonds forming between O=S=O i.e. the sulfonamide group of the ligand and protein residues Lys273 and Arg343. The residues Lys182, Lys273, Arg343 for both bosentan and K-8794 are in accordance with the fact that they form relevant H-bonds and contribute to the ligand binding as stated by Shihoya et al. (2017). More established H-bonds between ligand and surrounding residues may indicate stronger binding results. However, some hydrogen bonds can minimize the binding affinity of a ligand because of the unfavourable desolvation effect (Klebe, 2013). Therefore, further analysis such as free binding energy (FBE) estimations is necessary. Additionally, when several ligands are compared between each other, and they differ by a functional group that forms the H-bonds with the protein, the binding affinity can increase, remain the same, or even decrease (Klebe, 2013).

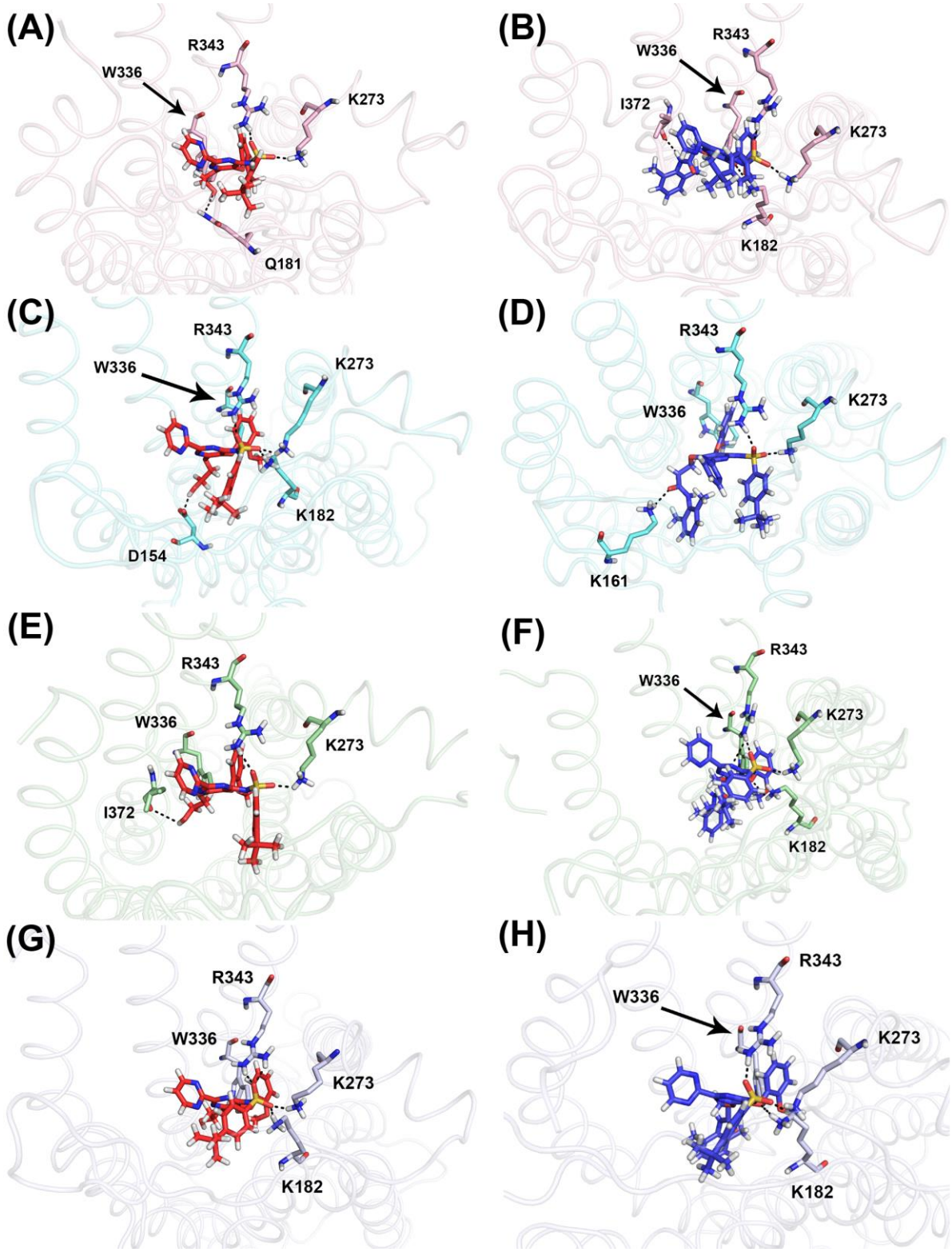


Figure 4.5. ET_B receptor orthosteric binding site showing stable hydrogen bonds at 4 Å distance between ligand and binding site residues that formed during the last MDS trajectory frame (400 ns) of all ligand-bound ET_B and membrane systems. The black dashed lines represent hydrogen bonds and sticks represent relevant residues. Bosentan – red sticks, K-8794 – blue sticks. (A) Bosentan-ET_B-POPC, (B) K-8794-ET_B-POPC, (C) Bosentan-ET_B-DPPC, (D) K-8794-ET_B-DPPC, (E) Bosentan-ET_B-POPE, (F) K-8794-ET_B-DPPC, (G) Bosentan-ET_B-DMPC, (H) K-8794-ET_B-DMPC.

4.2. Binding Free Energy estimation

By extracting 100 evenly spaced snapshots from the last 200 ns of the trajectory of the ligand-bound MD simulations, the overall binding free energy (BFE or ΔG_{bind}) of the complexes was estimated by MM/GBSA approach. The MM/GBSA predicts the binding strength between protein and its binding partner. In the current study, the residues that contribute to the ligand-binding through maximum BFEs (lowest minima) called hotspots have been identified. Lower BFE contributing residues are also favourable to the ligands binding.

In addition, the contribution to the total binding free energy from different components such as Coulomb energy, covalent binding energy, hydrogen bonding correction, π-π packing correction, generalized Born electrostatic solvation energy, and van der Waals energy was calculated for the 200 ns – 400 ns range trajectory of all ligand-bound system. The BFE estimation includes corrections of H-bonds and π-π packing. As stated by Li et al. (2011) these two values must be correct for an accurate protein structure prediction and stabilization of hydrophobic cores.

Performed analysis results (Table 4.1) revealed that for ET_B receptor in complex with bosentan the most favourable membrane was DPPC with BFE being -85.8762 kcal/mol, then followed by DMPC with BFE being -77.1620 kcal/mol, POPE with BFE being -73.7519 kcal/mol. The least favourable was POPC with BFE being -66.2156 kcal/mol. On the contrary, for ET_B receptor in complex with K-8794 the lowest minimum of -83.1504 kcal/mol was revealed to be when the receptor was embedded in the POPE membrane. Then favourability decreased following DMPC with BFE being -81.3768 kcal/mol and DPPC with -72.0154 kcal/mol. But similarly, as with bosentan, the least unfavourable membrane revealed to be POPC with a BFE value of -65.7402 kcal/mol.

The result revealed (Table 4.1) that for bosentan and K-8794 bound ET_B receptor in POPC, DPPC, POPE and DMPC membrane environment, the primary contributors in the BFE were Coulomb energy ($\Delta G_{coulomb)}$ and van der Waals energy (ΔG_{VdW}). The least favourable contributions were covalent binding energy ($\Delta G_{covalent}$) and generalized Born electrostatic solvation energy (ΔG_{GB}). The decrease of van der Waals energy is one of the characteristics that display weaker ligand binding (Chen et al., 2015). Even though the difference of van der Waals energies, when compared between membrane environments, is minor as presented in Table 3.1, it can still demonstrate ligand binding affinity and membrane environment favourability. Since ΔG_{VdW} is one of the biggest contributors to BFE. Therefore,

the lower the BFE is the stronger ligand binding and the more favourable membrane environment.

Table 4.1. Binding free energies (kcal/mol) and their separate components, calculated by MM/GBSA method for ET_B receptor in complex with bosentan and K-8794 embedded in different membrane environments.

	Component							
	Coulomb Energy $\Delta G_{coulomb}$ (kcal/mol)	**Covalent Binding Energy $\Delta G_{covalent}$ (kcal/mol)**	**Hydrogen Bonding Correction ΔG_{Hbond} (kcal/mol)**	**Lipophilic Energy ΔG_{lipo} (kcal/mol)**	**Pi-pi Packing Correction $\Delta G_{packing}$ (kcal/mol)**	**Generalized Born Electrostatic Solvation Energy ΔG_{GB} (kcal/mol)**	**Van der Waals Energy ΔG_{VdW} (kcal/mol)**	**Free Binding Energy ΔG_{bind} (kcal/mol)**
Membrane	**Bosentan**							
POPC	-48.9200	0.9443	-3.8652	-20.9129	-2.4458	60.7323	-51.7483	-66.2156
DPPC	-62.5312	2.3788	-3.8716	-27.0005	-4.1324	72.7557	-63.4750	-85.8762
POPE	-49.0362	3.0078	-3.4807	-26.3352	-3.8371	66.0648	-60.1353	-73.7519
DMPC	-42.9940	1.4615	-3.0339	-25.9265	-5.0364	58.1871	-59.8199	-77.1620
Membrane	**K-8794**							
POPC	-59.6918	3.1022	-3.7031	-20.2427	-3.9714	82.2483	-63.4817	-65.7402
DPPC	-58.4486	-0.2837	-1.9224	-25.4346	-3.6920	82.7090	-64.9431	-72.0154
POPE	-54.0754	4.3460	-3.2990	-29.8374	-4.1714	76.4793	-72.5925	-83.1504
DMPC	-54.4193	2.5841	-3.0706	-28.3599	-3.8806	76.0698	-70.3001	-81.3768

By analysing the per-residue decomposition (PRD) energies, two residues that contribute the most (hotspots) to the binding of both bosentan and K-8794 were identified (Figure 4.6). As shown in Figure 4.6 the hotspot residues, represented as the most negative energies, were Lys182 and Trp336. In accordance with Shihoya et al. (2017), these residues are important for ligand binding to the ET_B receptor. The other residues also exhibited relatively favourable BFEs as seen in Figure 4.6. However, this trend is not observed in all membrane systems.

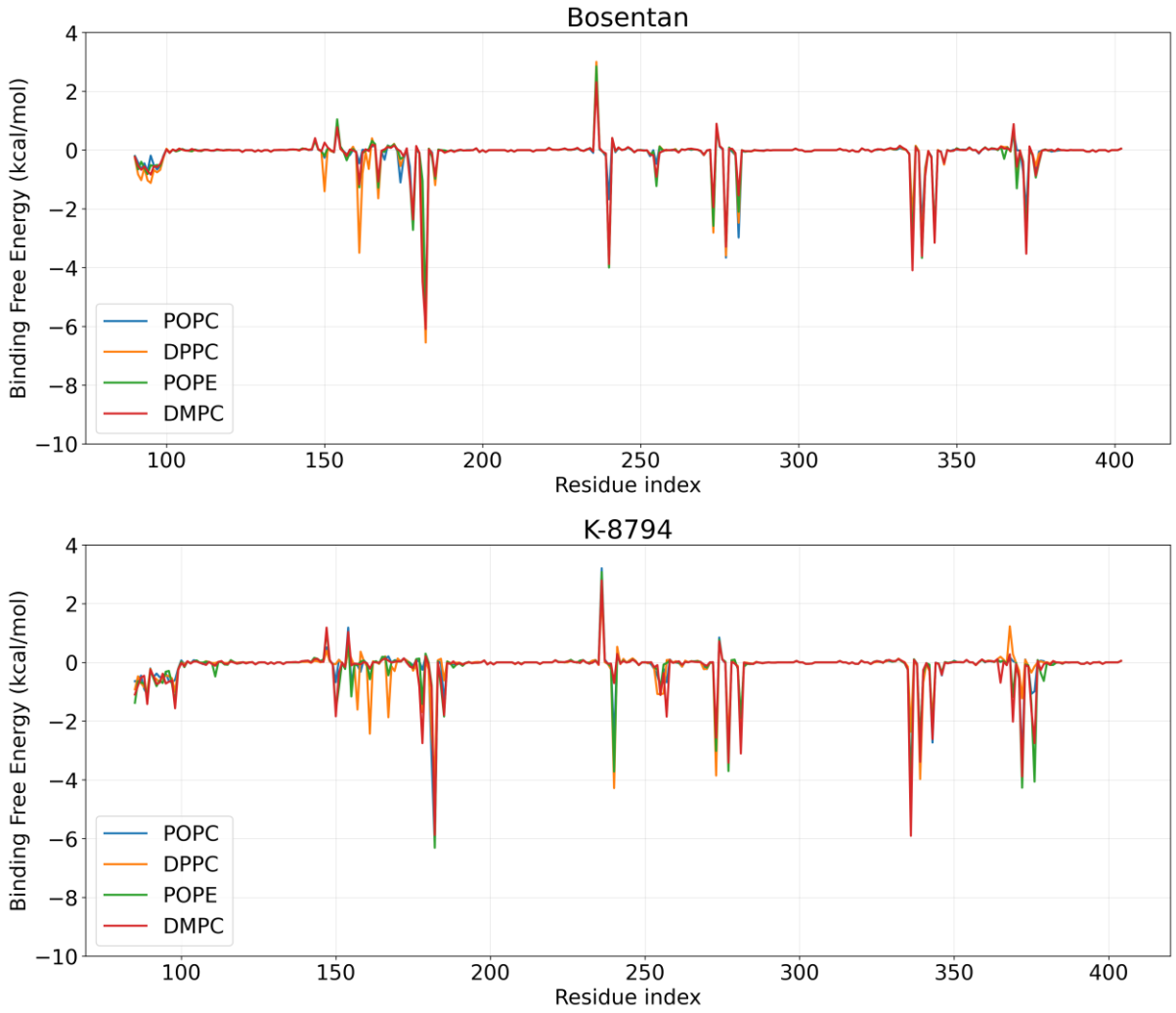


Figure 4.6. Per-residue BFEs (kcal/mol) estimated for bosentan and K-8794 in complex with ET_B receptor embedded in POPC (blue), DPPC (yellow), POPE (green) and DMPC (red) membranes. BFE was extracted from 100 evenly spaced snapshots from the last 200 ns trajectory of the ligand-bound ET_B receptor MD simulations.

The overall ET_B receptor residues that are of BFE -2 kcal/mol and lower, and appear in all of the systems, were identified and visualised in a 3D image for which the coordinates were extracted at 300 ns frame (Figure 4.7). For both bosentan and K-8794, as mentioned before, hotspot residues were Lys182 and Trp336.

For ET_B in complex with bosentan, Lys182 had far greater values than Trp336. In the Bosentan-ET_B-POPC system Lys182 BFE was -5.928 kcal/mol, and Trp336 – -3.428 kcal/mol (Figure 4.7 A). For the Bosentan-ET_B-DPPC system, Lys182 reached -6.5512 kcal/mol, while Trp336 was -2.428 kcal/mol (Figure 4.7 C). Lys182 exhibited a BFE value of -5.6572 kcal/mol, and Trp336 of -3.2918 kcal/mol in the Bosentan-ET_B-POPE system (Figure 4.7 E). Lastly, for the Bosentan-ET_B-DMPC system, Lys182 was -6.0874 kcal/mol, and Trp336 was -4.0919 kcal/mol (Figure 4.7 G).

For ET_B in complex with K-8794, similarly to bosentan Lys182 displayed more favourable BFE values than Trp336. In K-8794-ET_B-POPC system Lys182 exhibited -6.2063 kcal/mol BFE, while Trp336 – -4.4004 kcal/mol (Figure 4.7 B). For the K-8794-ET_B-DPPC system, Lys182 displayed a BFE value of -3.9672 kcal/mol, and Trp336 displayed -2.3608 kcal/mol (Figure 4.7 D). The Lys182 exhibited BFE at -6.3099 kcal/mol, while Trp336 exhibited -4.7622 kcal/mol in the K-8794-ET_B-POPE system (Figure 4.7 F). Finally, in the K-8794-ET_B-DMPC system, both hotspot residues exhibited similar BFE values, Lys182 BFE value was at -5.8749 kcal/mol, and Trp336 at -5.9025 kcal/mol (Figure 4.7 H).

On the whole, when comparing the two hotspot residue interactions with bosentan and K-8794 (Figure 4.6 – 4.7), the K-8794 had more favourable interactions with both Lys182 and Trp336 over bosentan. Since K-8794 is of higher affinity than bosentan, these findings are in accordance with the experimental results provided by Shihoya et al. (2017), where for K-8794 the K_D value is 150 nM and for bosentan – K_D value is 3,300 nM.

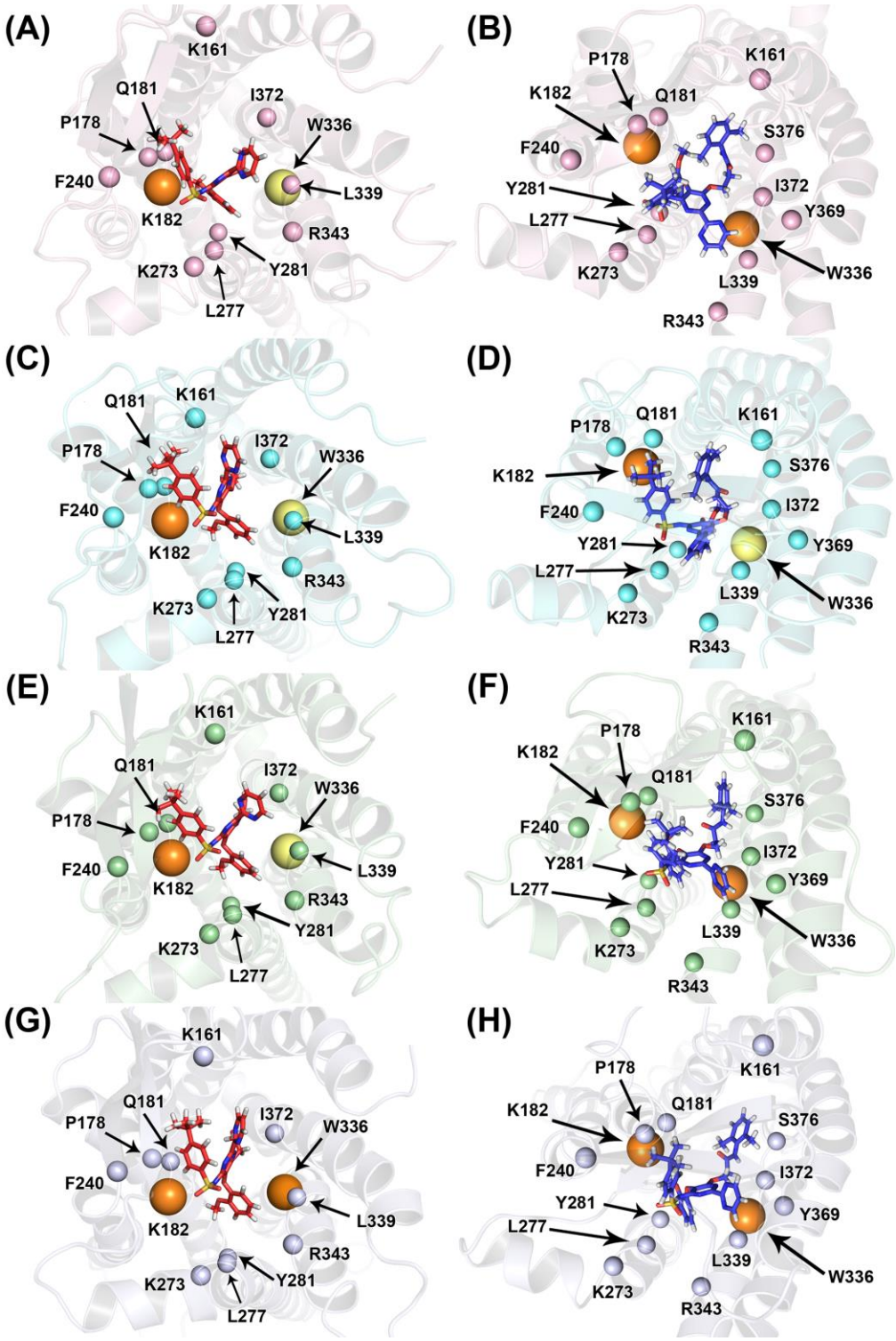


Figure 4.7. ET_B receptor (cartoon) orthosteric binding site displayed using the coordinates extracted from the 300 ns time frame. Bosentan is presented as red sticks, and K-8794 as blue sticks. Hotspot residues are highlighted as the enlarged spheres in orange having BFE of the lowest minima, and yellow having BFE above -4 kcal/mol. All of the other small spheres are residues of BFE higher minima than that of hotspot residues (lowest minima). (A) Bosentan-ET_B-POPC, (B) K-8794-ET_B-POPC, (C) Bosentan-ET_B-DPPC, (D) K-8794-ET_B-DPPC, (E) Bosentan-ET_B-POPE, (F) K-8794-ET_B-DPPC, (G) Bosentan-ET_B-DMPC, (H) K-8794-ET_B-DMPC.

4.3. Principal Component Analysis (PCA)

With the help of Principal Component Analysis (PCA) and Free Energy Landscape (FEL), it is possible to identify protein clusters (conformers) that formed during simulation time and study the protein flexibility by the wideness of the accessible conformational space visited. The PCA is often called Essential Dynamics (ED). It is an advanced MDS trajectory analysis method used for understanding and identifying the atomic fluctuations (conformational changes) of proteins. The changes are usually related to the function or biological properties. PCA works by reducing the dimensionality of the conformational space for the purpose of extracting the large scale collective motions of atoms. The large variance or total motion, which is coordinated to each other, of atoms is stored in the first principal component (PC1) and the rest of the variance is continued to be stored in the following principal components (PCs). Thus, having the largest variance, the first two principal components define the essential subspace of the system (Amusengeri et al., 2020; Chen et al., 2016).

FEL was built by plotting the first two PCs of PCA from the MD simulation data of ligand-bound complexes and APO structure. PC1 and PC2 were utilized, because of their high variance of MD simulation data. Low-lying energy basins (purple) represent visited clusters (conformers), and broad areas of FEL represent accessible conformational space (transition subspace). The targeted trajectory range for the analysis is 200 ns – 400 ns, as agreed by RMSD analysis results.

As seen in Figure 4.8, when comparing the Bosentan-free-ET_B-POPC (APO Bosentan) system with the Bosentan-ET_B-POPC (Bosentan) system, the APO structure formed one conformer, as presented by a single well energy landscape. While the Bosentan system formed two clusters (two low energy basins). However, both the complex and the structure spanned broad conformational spaces, with the APO structure exhibiting more flexibility. Similar events are seen when comparing K-8794-free-ET_B-POPC (APO K-8794) and K-8794-ET_B-POPC (K-8794) systems: APO structure demonstrated one cluster formed, while K-8794 complex formed two. The APO system displayed higher flexibility, because of broader conformational space, unlike the K-8794 system.

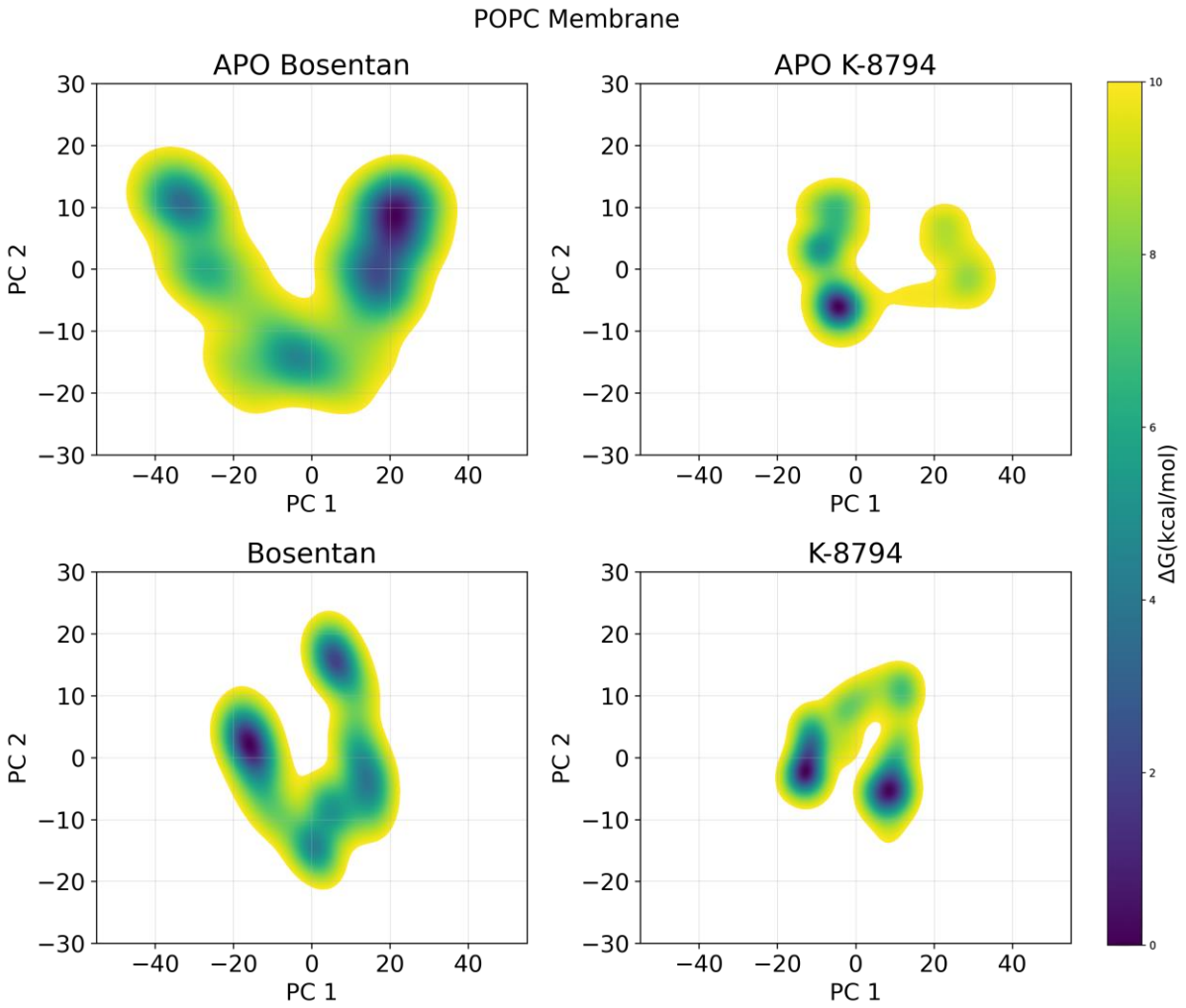


Figure 4.8. Free energy landscape (FEL) plots of bosentan and K-8794 bound and APO ET_B receptor embedded in POPC membrane, built by two principal components (PC1 and PC2) of MD trajectory. The landscapes are color-coded from yellow (energy maxima) to purple (energy minima).

When comparing bosentan-bound ET_B (Bosentan) structure and its equivalent APO (APO Bosentan) structure embedded in DMPC membrane (Figure 4.9), the APO structure obtained three conformations, while bosentan-bound formed two. The ligand-free structure also spanned a wider conformational space, unlike ligand-bound. K-8794-bound and unbound systems experienced different events. Ligand-free structure (APO K-8794) obtained one cluster, similarly to ligand-bound (K-8794). However, the APO structure also spanned wider conformational space. Therefore, the exhibited conformational space broadness of the unbound ET_B receptor indicates a highly flexible system, before transitioning to the clusters.

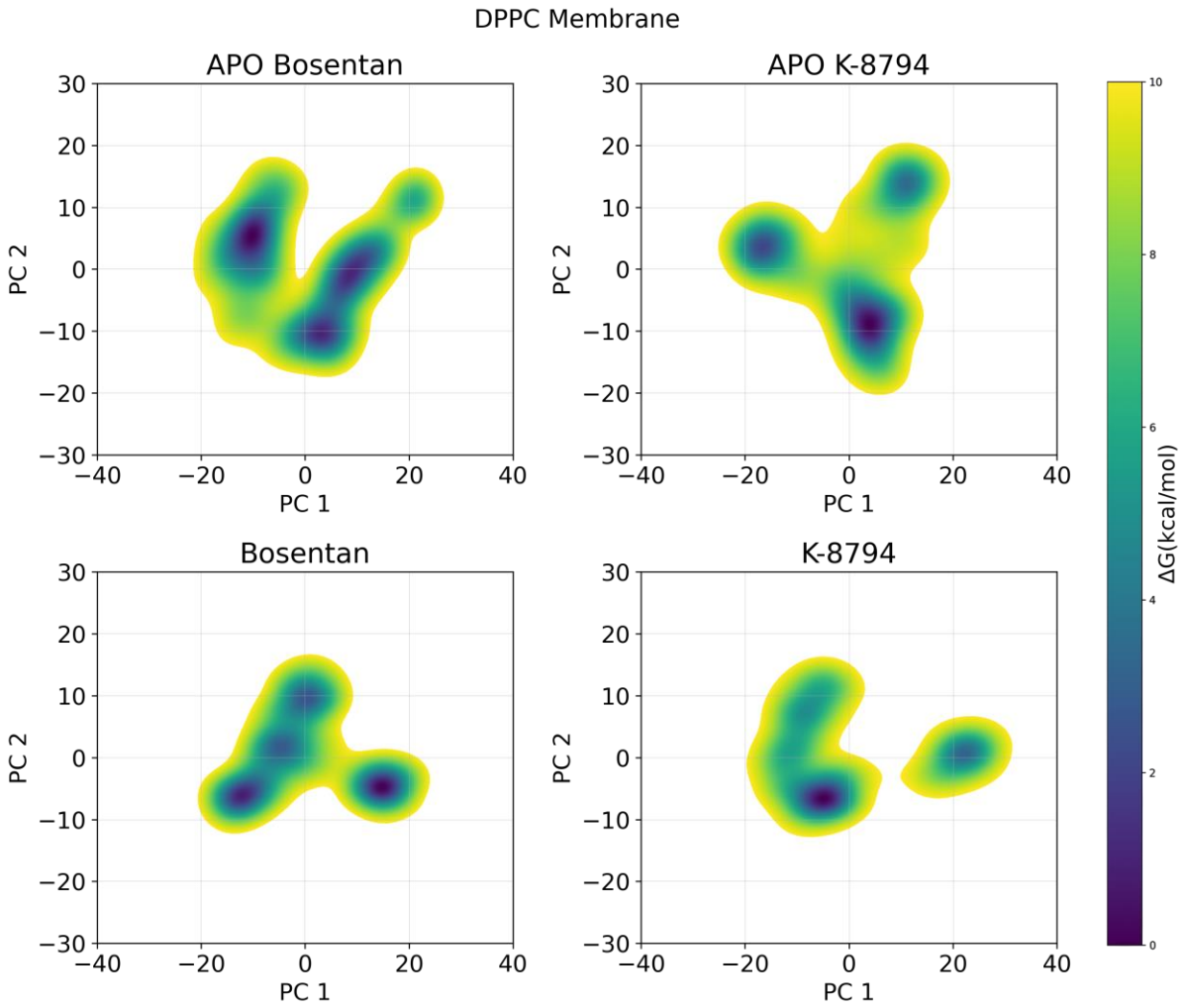


Figure 4.9. Free energy landscape (FEL) plots of bosentan and K-8794 bound and APO ET_B receptor embedded in DPPC membrane, built by two principal components (PC1 and PC2) of MD trajectory. The landscapes are color-coded from yellow (energy maxima) to purple (energy minima).

Systems of ligand-bound and ligand-free ET_B receptor embedded in the POPE membrane exhibited different conformational events (Figure 4.10). Comparing bosentan in complex with ET_B receptor (Bosentan) with its ligand-free structure (APO Bosentan), the APO system visited two metastable conformers, while ligand-bound visited one. However, the bosentan-bound system displayed a broader conformational space, unlike the bosentan-unbound. The K-8794-free (APO K-8794) and bound structure (K-8794) FEL revealed only one formed conformer for both systems. In addition, both systems explored wide conformational space, indicating high flexibility, before obtaining the clusters formed.

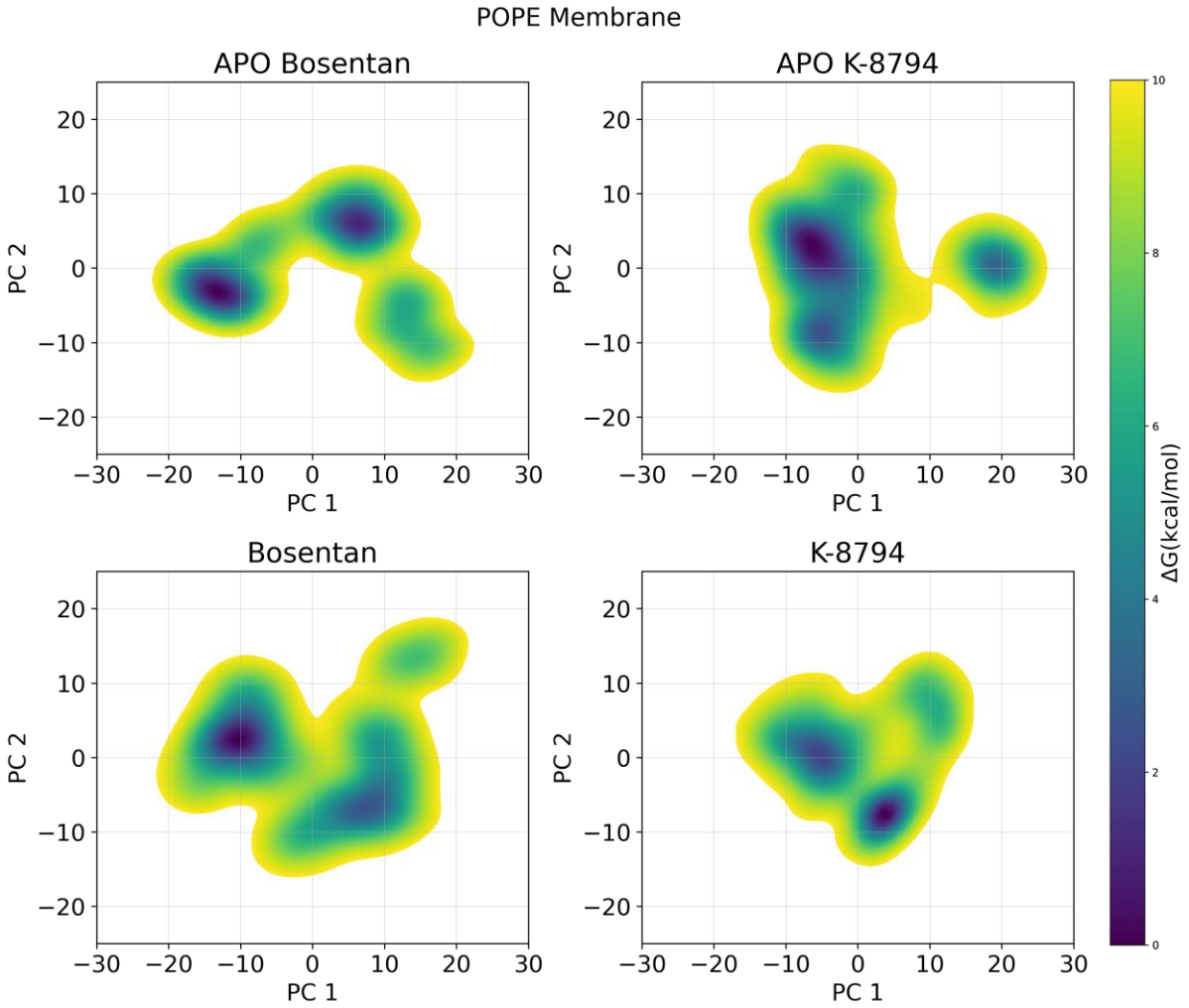


Figure 4.10. Free energy landscape (FEL) plots of bosentan and K-8794 bound and APO ET_B receptor embedded in POPE membrane, built by two principal components (PC1 and PC2) of MD trajectory. The landscapes are color-coded from yellow (energy maxima) to purple (energy minima).

As seen in Figure 4.11, bosentan-unbound (APO Bosentan) ET_B receptor structure formed one cluster, but spanned a wide conformational space, before transitioning to that cluster when embedded in the DMPC membrane. However, bosentan-bound (Bosentan) obtained two conformers, and displaying the structure has visited transitional space greatly. Unlike the bosentan-bound system, the K-8794-bound formed only one cluster but visited the conformational space extensively. Similar events were exhibited in the K-8794-free (APO K8794) system: one conformer visited, but spanned a relatively wide conformational space.

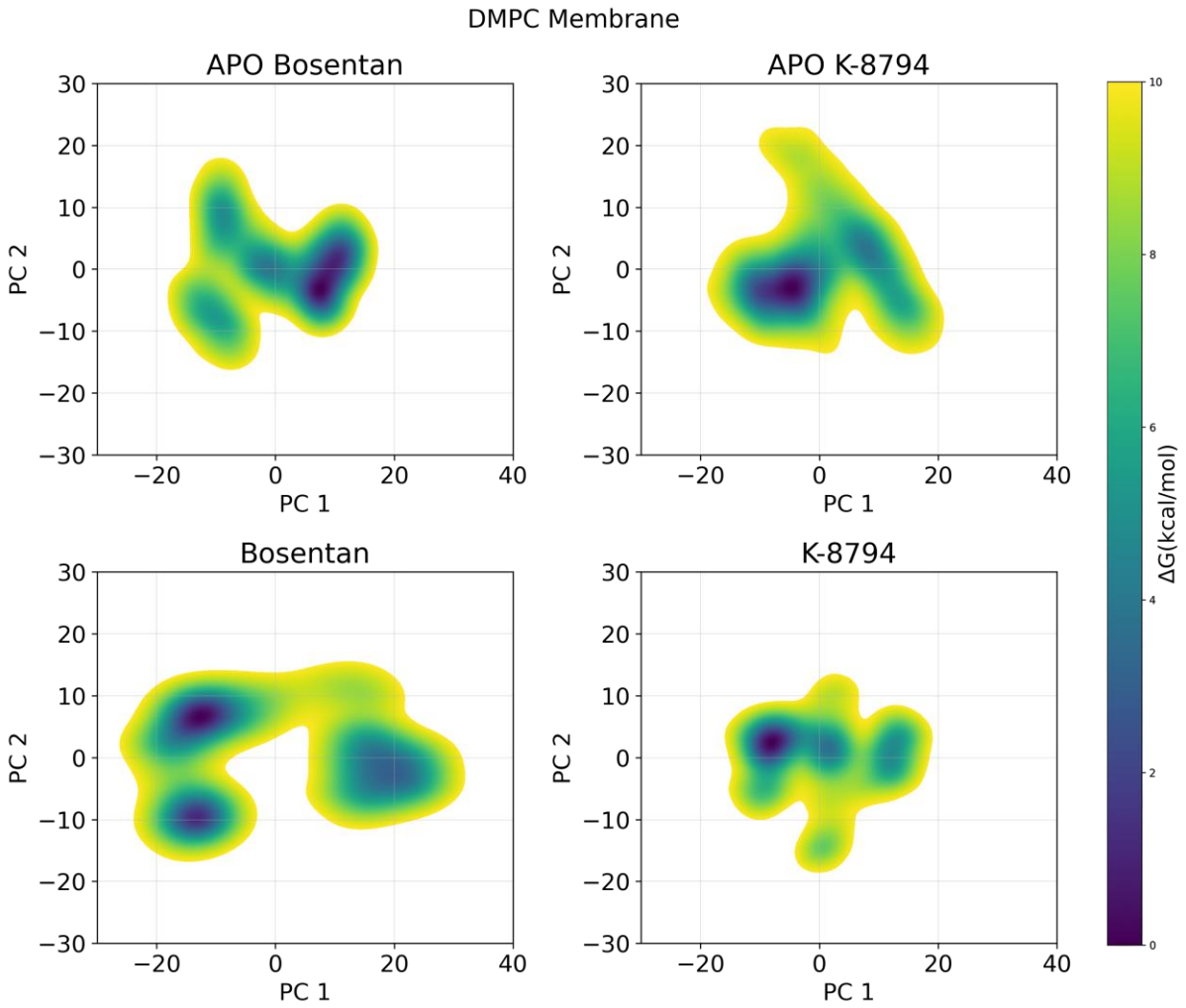


Figure 4.11. Free energy landscape (FEL) plots of bosentan and K-8794 bound and APO ET_B receptor embedded in DMPC membrane, built by two principal components (PC1 and PC2) of MD trajectory. The landscapes are color-coded from yellow (energy maxima) to purple (energy minima).

The overall analysis indicates that broadly visited conformational space over the simulation time suggests highly flexible systems. However, several clusters formed means the receptor obtained several different conformations, while one cluster indicates a formation of one conformation over the span of the MD simulation.

For the purpose of understanding protein structural changes and the direction of its motion, the PCAs are highlighted over the protein structure as arrows (porcupine plot) sampled from MD simulation trajectories (200 ns to 400 ns) for ET_B bound to bosentan and K-8794 embedded in the four membranes, were generated taking the first principal component mode (PC1) (Figure 4.12). The ET_B experienced major movement in the ECL and ICL area for all of the membrane systems, as expected. The only exception is visible in the Bosentan-ET_B-POPE system, as ICLs did not experience any major movement. A slight movement of the helix 8, which is the horizontal short helix in this case positioned on the

right, is exhibited by Bosentan-ET_B-POPC, K-8794-ET_B-POPC and K-8794-ET_B-DMPC systems. The flexibility of the intra and extracellular loops is highly expected because of their inherent flexibility, and this also agrees with the fluctuations observed in the RMSF analysis. Additionally, the two bound antagonists retained the ET_B receptor intact, by restraining its activity, as most of the intracellular loops seemed to be moving inwards. As expected, the secondary structure of the protein (transmembrane helices one to seven) did not experience any major movement and remained rigid, because the membrane phospholipid's one of the main functions is to retain the integrity of the cell walls and keep the correct placement of the biomolecules. Furthermore, the TM helices, especially the TM6, did not demonstrate any outer pendulum-like motion as the lower and upper helical bundle stayed closed inwards.

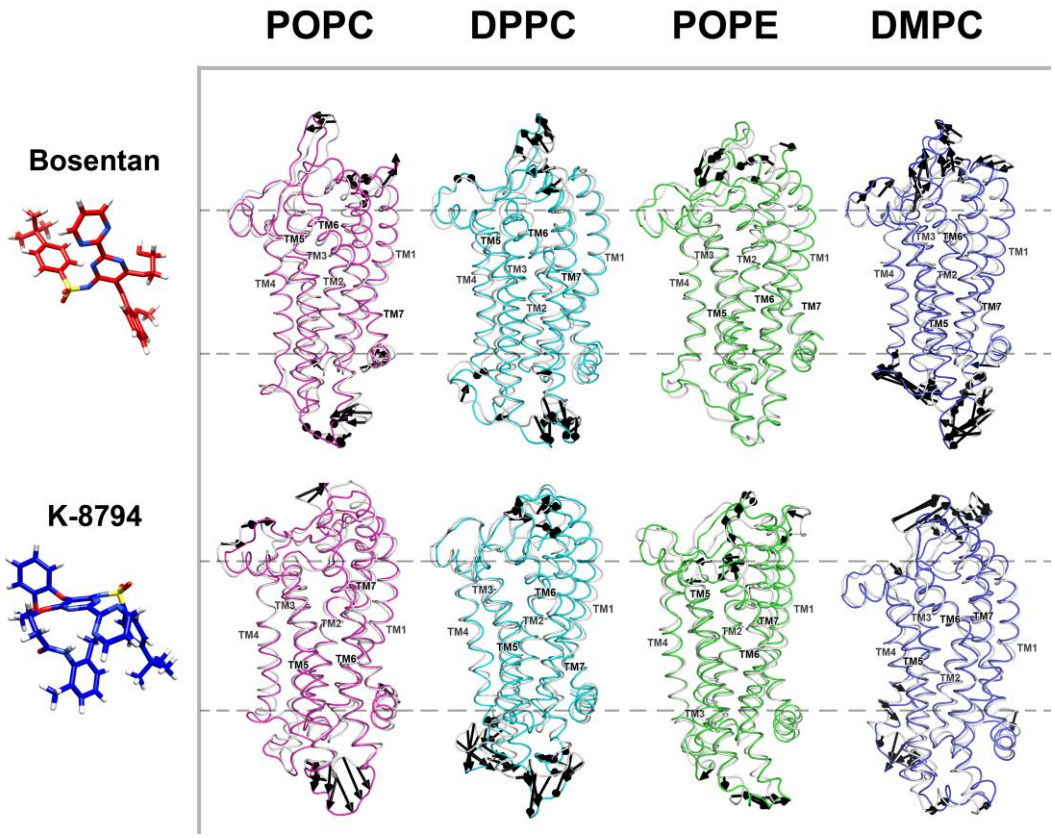


Figure 4.12. Porcupine plot highlighting the first PCA mode (PC1) of the bosentan (red) and K-8794 (blue) bound ET_B receptor complexes embedded in POPC, DPPC, POPE and DMPC membranes. The white cartoons indicate the initial secondary structure and loops before the occurrence of any structural changes. The purple (POPC), cyan (DPPC), green (POPE) and dark blue (DMPC) cartoons, representing the secondary structure and loops, indicate the finale of structural change. The black arrows represent the direction of motion of every Cα atom, while the length acts for the strength of the movement. The semi-transparent dashed black lines represent the outer and inner leaflets of the membrane.

4.4. Dihedral Angle analysis

For the purpose of understanding structural rearrangements and conformational diversity of antagonist-bound ET_B receptor, the dihedral angles of the hotspot residues Lys182 and Trp336 were extracted from the trajectories. The FEL representation was used to plot the residue angles of ligand-bound and compared with its corresponding APO ET_B receptor structures.

According to Figure 4.13, in the POPC membrane the angles for φ and ψ of Lys182 APO and the bosentan bound system did not demonstrate changes: (-70,-30). On the other hand, the same residue experienced changes with K-8794 bound compared to APO: backbone angles (φ, ψ) in APO were at (-70°,-30°), K-8794 were at (-70°,-40°). The angles φ and ψ for Trp336 in bosentan bound ET_B and APO did demonstrate slight changes when compared: in APO – (-90°,10°), in bosentan – (-90°,0°). For K-8794 bound and APO systems, Trp336 when compared between bound and unbound ligand, did have changes in angles: in APO – (-90°,-10°), in the K-8794-bound system the backbone angles (φ, ψ) were at (-95°,0°).

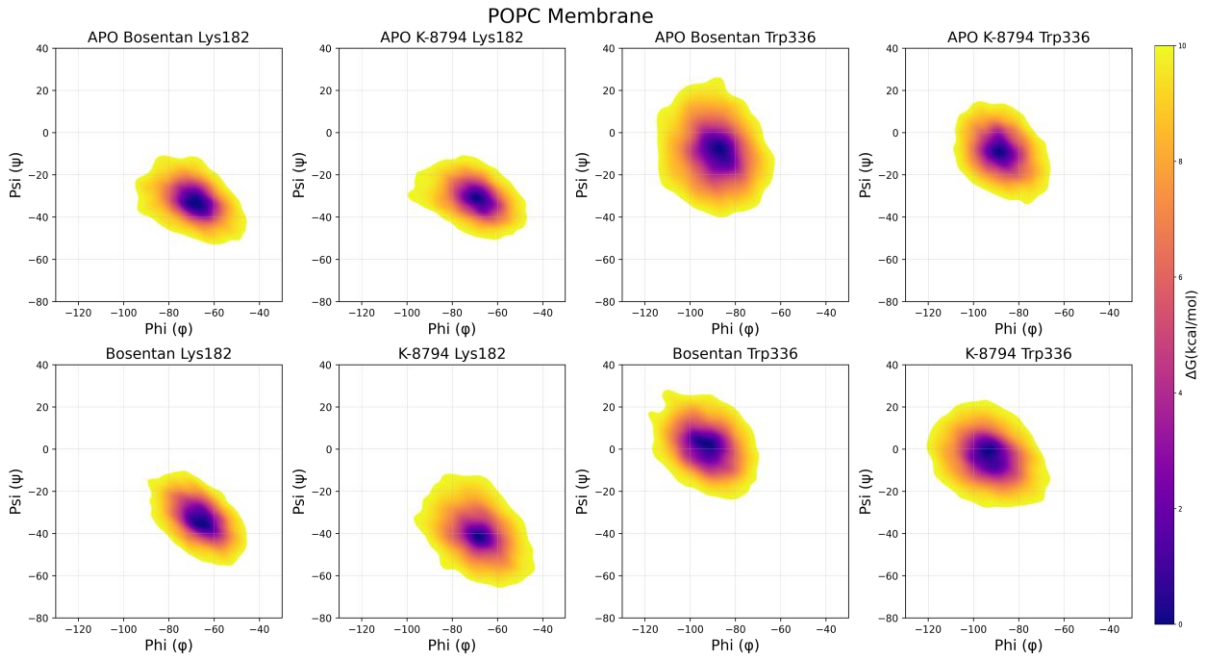


Figure 4.13. Free energy landscape (FEL) plots of hotspot residues Lys182 and Trp336 for bosentan and K-8794 bound, and APO ET_B receptor embedded in POPC membrane, built by two backbone angles φ (phi) and ψ (psi). The landscapes are color-coded from yellow (energy maxima) to blue (energy minima).

The two hotpot residue backbone angles (φ, ψ) of the ET_B receptor in the DPPC membrane are visualised in Figure 4.14. As seen in the figure, the Lys182 backbone angles, when compared with APO and bosentan bound receptor, did not experience any differences

as the residue angles in both systems were (-70°,-30°). However, when comparing with APO, the Lys182 residue of K-8794-bound ET_B did demonstrate high variation: in APO – (-70°,-30°), in K-8794-bound system – (-95°,-15°); the φ angle had a difference of 25°, and ψ angle had a difference of 15°. For Trp336 the φ and ψ angles in APO and bosentan-bound receptor system also exhibited different values: in APO the angles were (-90°,-10°), and in bosentan-bound receptor system the backbone angles were (-95°,0°). A similar phenomenon is exhibited in the K-879-bound systems, in APO Trp336 φ and ψ angles were (-90°,-15°), and in K-8794-bound ET_B receptor system the values were (-85°,-10°).

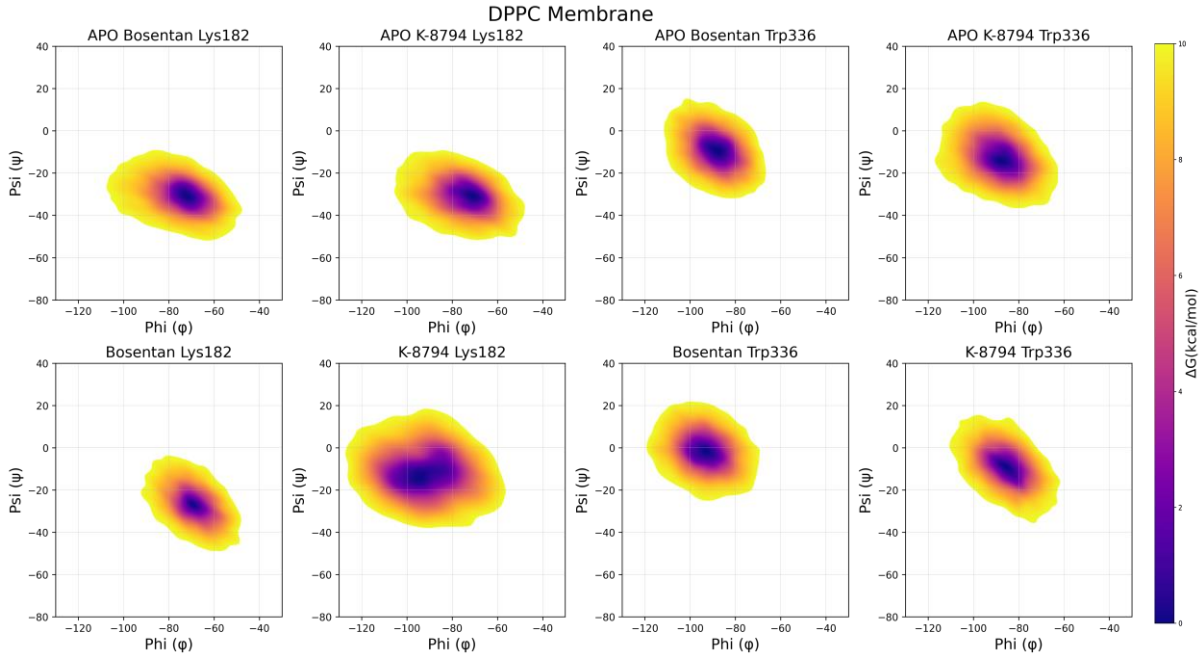


Figure 4.14. Free energy landscape (FEL) plots of hotspot residues Lys182 and Trp336 for bosentan and K-8794 bound, and APO ET_B receptor embedded in DPPC membrane, built by two backbone angles φ (phi) and ψ (psi). The landscapes are color-coded from yellow (energy maxima) to blue (energy minima).

The backbone angles (φ, ψ) of Lys182 and Trp336 of bosentan-bound and K-8794-bound ET_B receptor in the POPE membrane environment, and its APO structure is visualised in Figure 4.15. According to the FEL plot of Lys182 in the bosentan-bound system and APO, the φ and ψ angles had slight differences when comparing the system and APO: in APO structure Lys182 backbone angles were (-70°,-35°), and in receptor in complex with bosentan the residues angles were (-70°,-25°). For the same residue but in ET_B in complex with K-8794 and APO structure demonstrated larger differences. In the K-8794-free structure, Lys182 dihedral angles (φ, ψ) were (-75°,-40°), and in the K-8794-bound structure, the angles were (-65°,-35°). Trp336 residue backbone angles in the bosentan-free and bosentan-

bound structure were (-85°,-20°) and (-90°,-5°), respectively. Based on the previously mentioned results, the ψ angle of Trp336 had a significant difference of about 15°. Similarly in K-8794-bound and Apo structure, there was a high deviation between the angles of Trp336 in the two structures: in APO – (-80°,-20°), in K-8794-bound – (-95°,-5°). Both of the angles differed by 15°.

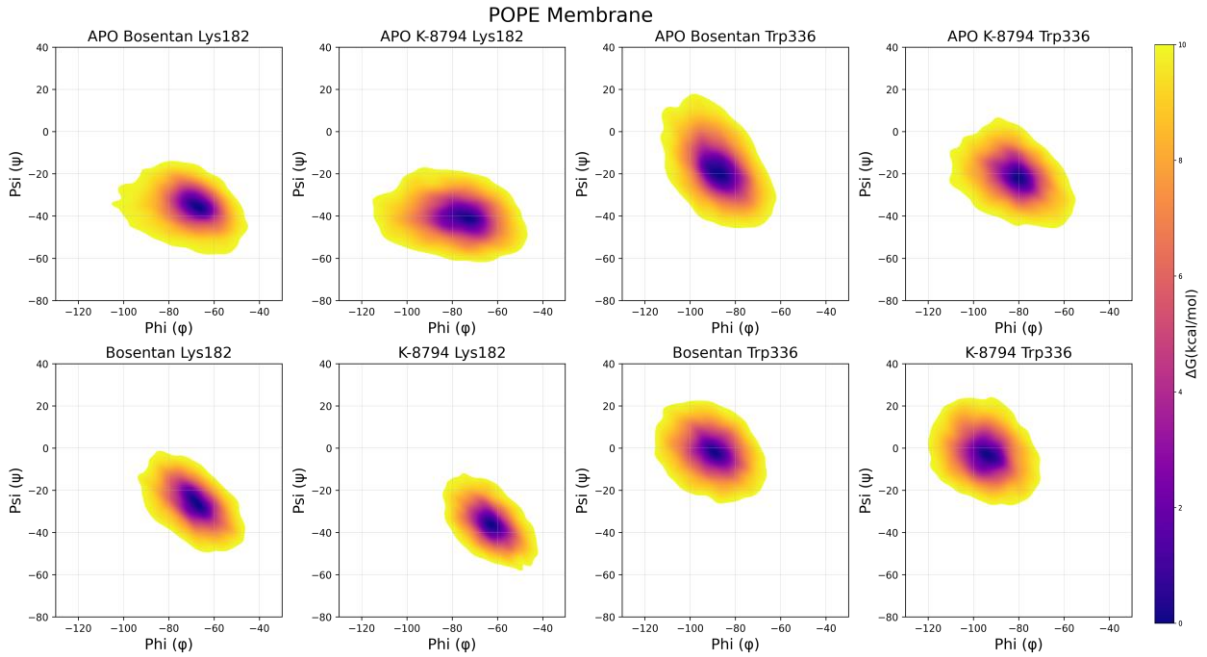


Figure 4.15. Free energy landscape (FEL) plots of hotspot residues Lys182 and Trp336 for bosentan and K-8794 bound, and APO ET_B receptor embedded in POPE membrane, built by two backbone angles φ (phi) and ψ (psi). The landscapes are color-coded from yellow (energy maxima) to blue (energy minima).

The dihedral angle results are presented in Figure 4.16 for Lys182 and Trp336 residues in bosentan-bound, K-8794-bound and APO ET_B structures in the DMPC membrane environment. When comparing Lys182 φ and ψ angles in bosentan-bound and ligand-free receptor structures, the angles did not show any significant differences, as they were at around (-70°,-30°) for both systems. The same phenomenon was seen in K-8794-bound and K-8794-free structures: the φ and ψ angles in both systems were (-70°,-30°). When comparing backbone angles between the two structures (APO and bosentan-bound) for Trp336 residue, it did not exhibit any significant contrast, because the angles were (-90°,-5°) for both systems. However, the φ and ψ angles Trps336 in K-8794-bound structure were (-85°,-10°), while in APO were (-98°,-10°); φ angle had a difference of around 13°.

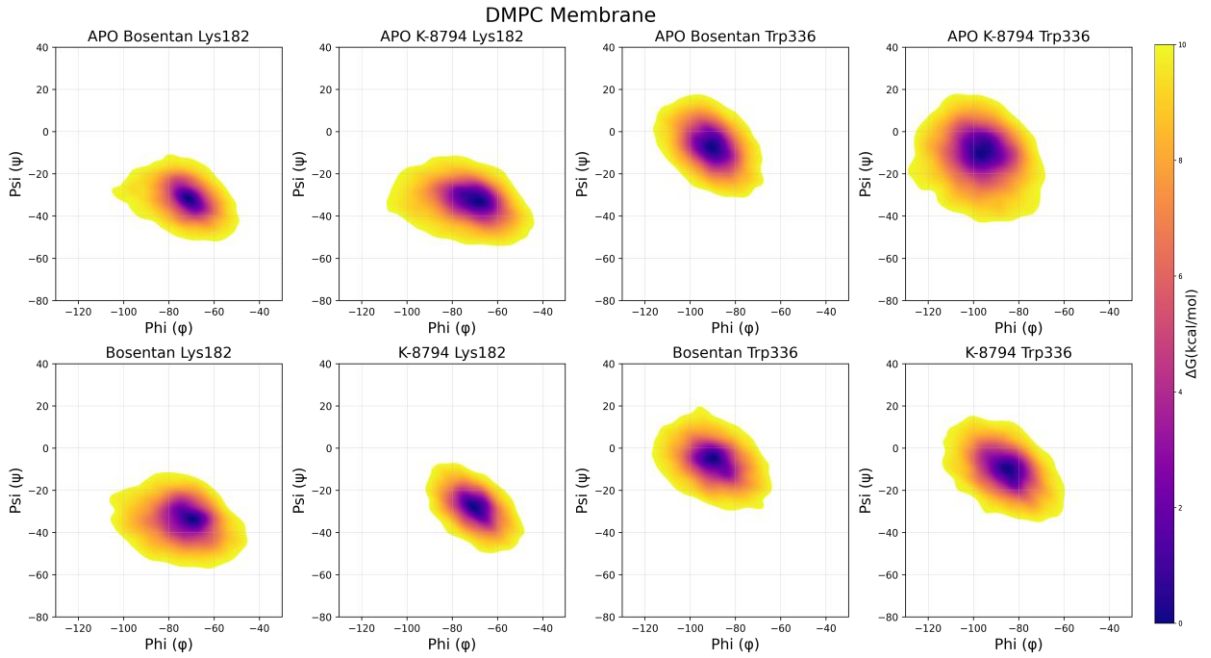


Figure 4.16. Free energy landscape (FEL) plots of hotspot residues Lys182 and Trp336 for bosentan and K-8794 bound, and APO ET_B receptor embedded in DMPC membrane, built by two backbone angles φ (phi) and ψ (psi). The landscapes are color-coded from yellow (energy maxima) to blue (energy minima).

On the whole, most of the φ (phi) and ψ (psi) angle differences were at around 5° to 10° in all membrane environments. However, higher than that of 5-10° differences in angles were seen in Lys182 of K-8794-bound and APO structures in the DPPC environment, where the φ angle differs by 25° between the two structures. Additionally, the 15° difference of ψ angle between the APO and bosentan-bound structures in the POPE membrane environment of Trp336 residue was exhibited. Also, both the φ and ψ angles differed by 15° for Trp336 residue when comparing the K-8794-bound and ligand-free structure in POPE membrane. A difference of around 13° in φ angle was exhibited by Trp336 residue of the previously mentioned structures, but when they were embedded in the DMPC environment. Overall, the differences in angles were not that significantly large, but they were important angle changes. These small differences in φ (phi) and ψ (psi) angles when comparing both ligand-bound structures with their equivalent APO, and bosentan with K-8794 meant that both inhibitors maintained similar conformational orientation of the side chains of Lys182 and Trp336 residues in all of the membrane environments, while the peptide bond remained rigid.

CONCLUSIONS

The Endothelin B (ET_B) receptor belongs to a Class A G protein-coupled receptor family. It consists of seven α-helical bundles and loop regions facing the external and internal side of the membrane. ET_B functions as a vasodilator and expands the blood vessels. When activated by ET-1 the receptor induces changes to the transmembrane helix 6 (TM6) and transmembrane helix 7 conformation, causing their extracellular region to move ca. 4 Å inwards, and the TM6 intracellular region to move outwards. The ET_B receptor is involved in several diseases such as pulmonary arterial hypertension (PAH), hepatopulmonary syndrome (HPS), chronic heart failure (CHF) and many more. Therefore, several ET_B receptor targeting drugs have been discovered. Two of them are the main targets of this study, and they are called bosentan and K-8794. Bosentan is the first of its kind inhibitory drug, targeting both the ET_A and ET_B receptors. K-8794 is bosetan's high-affinity analog specific to the ET_B receptor. When bound, both antagonistic ligands occupy the traditional orthosteric binding site of the ET_B.

This study has investigated the induced changes to ET_B receptor during bosentan and K-8794 binding when the receptor-ligand complexes were embedded into four membranes of POPC, POPE, DMPC and DPPC. The study was carried out by a series of 400 ns long molecular dynamics (MD) simulations. The RMSD analysis has revealed that the last 200 ns of the simulations (200 ns to 400 ns) were equilibrated and used for further analysis. RMSF analysis revealed (as expected) that the most fluctuating parts of the ET_B receptor structure are the extracellular and intracellular loops. The Rg analysis showed that the two ligand-receptor complexes were relatively compact in all the membranes, but the most compact was K-8794-ET_B in POPC and DPPC membranes. On the other hand, the least compact systems were the bosentan-ET_B complex in POPC and the K-8794-ET_B complex in the POPE membrane. During the H-bond analysis, the residues Lys182, Lys273, and Arg343 for both bosentan and K-8794 are in accordance with the fact that they form relevant H-bonds and contribute to the ligand binding. In addition, H-bonds established between the two ligands and the receptor in all of the membrane systems were 3 per average. The binding free energy (BFE) estimations suggested that the Lys182 and Trp336 contributed the most to the binding of both ligands (hotspot residues). Based on the BFE estimations, out of the four membranes, the most favourable membrane for bosentan binding to ET_B was the DPPC membrane, while for K-8794 was the POPE membrane, as it exhibited the highest BFE values (highest minima). The least favourable was the POPC membrane for both. Out of the two ligands, the

more favourable established interactions with both Lys182 and Trp336 residues were of the K-8794 ligand. Principal component analysis (PCA) suggested that the ligand-ET_B complexes are highly flexible systems that can visit a broad conformational space over the simulation time. Several clusters formed means the receptor obtained several different conformations, while one cluster indicates a formation of one conformation over the span of the MD simulation. Additionally, PCA revealed that the two inhibitors kept the ET_B receptor intact and restrained its activity as all of the intracellular and extracellular loops were moving inwards, trapping the antagonist in the protein's orthosteric binding pocket. Furthermore, the PCA analysis revealed that the TM6 did not experience any outer movement in its intracellular region, as expected when the two antagonists were bound. The dihedral angle analyses based on MD trajectories revealed that the dihedral angle changes of the hotspot residues are not large. The angles over all of the systems exhibited similarity, therefore bosentan and K-8794 maintained similar conformational orientation of the side chains of the two hotspot residues throughout the simulations while being embedded in the four membranes. Lastly, the study may provide theoretical knowledge and help with further drug development and discovery for ET_B receptor-related medical disorders.

RECOMMENDATIONS

The study revealed that the POPC membrane was the least favourable for a ligand binding to the ET_B receptor, out of the four membranes analysed. Therefore, it may indicate that the ET_B receptor favours more saturated membranes because they are less fluid and have higher constraining ability for a signalling protein, and the membrane can create more contacts with receptor helices. However, unsaturated membranes, such as POPC, can induce higher receptor flexibility. Hence, further studies may use saturated membranes firstly, for more accurate results, but it should not alienate unsaturated membrane types. But could use, for example, a mixture of two phospholipid types. However, it really depends on how the membranes of certain cell, which may be targeted, are constructed. Lastly, the ligand binding exploration study may provide theoretical knowledge and help with further drug development and discovery of ET_B receptor-related medical disorders.

REFERENCES

9.1C: Types of Receptors - Biology LibreTexts. (n.d.). Retrieved March 19, 2021, from https://bio.libretexts.org/Bookshelves/Introductory_and_General_Biology/Book%3A_General_Biology_(Boundless)/9%3A_Cell_Communication/9.1%3A_Signaling_Molecules_and_Cellular_Receptors/9.1C%3A_Types_of_Receptors

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). *Signaling through Enzyme-Linked Cell-Surface Receptors*. https://www.ncbi.nlm.nih.gov/books/NBK26822/

Alexander Yi, B., & Jan, L. Y. (2002). Ion Channels. In *Encyclopedia of the Human Brain* (pp. 599–615). Elsevier. https://doi.org/10.1016/B0-12-227210-2/00180-1

Amadei, A., Linssen, A. B. M., & Berendsen, H. J. C. (1993). Essential dynamics of proteins. *Proteins: Structure, Function, and Bioinformatics*, *17*(4), 412–425. https://doi.org/10.1002/prot.340170408

ambrisentan | C22H22N2O4 | ChemSpider. (n.d.). Retrieved May 6, 2021, from http://www.chemspider.com/Chemical-Structure.5293690.html

Amusengeri, A., Tata, R. B., & Tastan Bishop, Ö. (2020). Understanding the Pyrimethamine Drug Resistance Mechanism via Combined Molecular Dynamics and Dynamic Residue Network Analysis. *Molecules*, *25*(4), 904. https://doi.org/10.3390/molecules25040904

Angladon, M. A., Fossépré, M., Leherte, L., & Vercauteren, D. P. (2019). Interaction of POPC, DPPC, and POPE with the μ opioid receptor: A coarse-grained molecular dynamics study. *PLoS ONE*, *14*(3). https://doi.org/10.1371/journal.pone.0213646

Ashton, M. R., Postle, A. D., Hall, M. A., Smith, S. L., Kelly, F. J., Normand, C. S., & Normand, I. C. S. (1992). Phosphatidylcholine composition of endotracheal tube aspirates of neonates and subsequent respiratory disease A D Postle M A Hall. In *Archives ofDisease in Childhood* (Vol. 67).

Barton, M., & Yanagisawa, M. (2008). Endothelin: 20 Years from discovery to therapy. In *Canadian Journal of Physiology and Pharmacology* (Vol. 86, Issue 8, pp. 485–498). https://doi.org/10.1139/Y08-059

Bedan, M., Grimm, D., Wehland, M., Simonsen, U., Infanger, M., & Krüger, M. (2018). A Focus on Macitentan in the Treatment of Pulmonary Arterial Hypertension. *Basic & Clinical Pharmacology & Toxicology*, *123*(2), 103–113. https://doi.org/10.1111/bcpt.13033

Belge, C., & Delcroix, M. (2019). Treatment of pulmonary arterial hypertension with the dual

endothelin receptor antagonist macitentan: clinical evidence and experience. In *Therapeutic Advances in Respiratory Disease* (Vol. 13, p. 175346661882344). SAGE Publications Ltd. https://doi.org/10.1177/1753466618823440

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The Protein Data Bank. In *Nucleic Acids Research* (Vol. 28, Issue 1, pp. 235–242). Oxford University Press. https://doi.org/10.1093/nar/28.1.235

Bhattarai, A., Wang, J., & Miao, Y. (2020). G-Protein-Coupled Receptor–Membrane Interactions Depend on the Receptor Activation State. *Journal of Computational Chemistry*, *41*(5), 460–471. https://doi.org/10.1002/jcc.26082

Biagini, T., Petrizzelli, F., Truglio, M., Cespa, R., Barbieri, A., Capocefalo, D., Castellana, S., Tevy, M. F., Carella, M., & Mazza, T. (2019). Are Gaming-Enabled Graphic Processing Unit Cards Convenient for Molecular Dynamics Simulation? *Evolutionary Bioinformatics Online*, *15*, 1176934319850144. https://doi.org/10.1177/1176934319850144

Bosentan | C27H29N5O6S | ChemSpider. (n.d.). Retrieved May 6, 2021, from http://www.chemspider.com/Chemical-Structure.94651.html

Bowers, K. J., Chow, D. E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., Klepeis, J. L., Kolossvary, I., Moraes, M. A., Sacerdoti, F. D., Salmon, J. K., Shan, Y., & Shaw, D. E. (2007). *Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters*. 43–43. https://doi.org/10.1109/sc.2006.54

Bylund, D. B. (2014). Receptor Types and Subtypes☆. In *Reference Module in Biomedical Sciences*. Elsevier. https://doi.org/10.1016/b978-0-12-801238-3.05263-6

Chan, H. C. S., Li, Y., Dahoun, T., Vogel, H., & Yuan, S. (2019). New Binding Sites, New Opportunities for GPCR Drug Discovery. In *Trends in Biochemical Sciences* (Vol. 44, Issue 4, pp. 312–330). Elsevier Ltd. https://doi.org/10.1016/j.tibs.2018.11.011

Chen, J., Wang, J., Zhang, Q., Chen, K., & Zhu, W. (2015). Probing Origin of Binding Difference of inhibitors to MDM2 and MDMX by Polarizable Molecular Dynamics Simulation and QM/MM-GBSA Calculation. *Scientific Reports*, *5*(1), 17421. https://doi.org/10.1038/srep17421

Chen, J., Wang, J., & Zhu, W. (2016). Molecular Mechanism and Energy Basis of Conformational Diversity of Antibody SPE7 Revealed by Molecular Dynamics Simulation and Principal Component Analysis. *Scientific Reports*, *6*(1), 1–12. https://doi.org/10.1038/srep36900

Clustal Omega < Multiple Sequence Alignment < EMBL-EBI. (n.d.). Retrieved May 11, 2021, from https://www.ebi.ac.uk/Tools/msa/clustalo/

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

D'Orléans-Juste, P., Akide Ndunge, O. B., Desbiens, L., Tanowitz, H. B., & Desruisseaux, M. S. (2019). Endothelins in inflammatory neurological diseases. In *Pharmacology and Therapeutics* (Vol. 194, pp. 145–160). Elsevier Inc. https://doi.org/10.1016/j.pharmthera.2018.10.001

Darusentan | C22H22N2O6 | ChemSpider. (n.d.). Retrieved May 7, 2021, from http://www.chemspider.com/Chemical-Structure.154336.html

Davenport, A. P., Hyndman, K. A., Dhaun, N., Southan, C., Kohan, D. E., Pollock, J. S., Pollock, D. M., Webb, D. J., & Maguire, J. J. (2016). Endothelin. In *Pharmacological Reviews* (Vol. 68, Issue 2, pp. 357–418). American Society for Pharmacology and Experimental Therapy. https://doi.org/10.1124/pr.115.011833

Davies, M. N., Secker, A., Freitas, A. A., Mendao, M., Timmis, J., & Flower, D. R. (2007). On the hierarchical classification of G protein-coupled receptors. *Bioinformatics*, *23*(23), 3113–3118. https://doi.org/10.1093/bioinformatics/btm506

Echave, J. (2009). Textbook of Structural Biology by Anders Liljas, Lars Liljas, Jure Piskur, Göran Lindblom, Poul Nissen, and Morten Kjeldgaard. *Trends in Evolutionary Biology*, *1*(1), 2. https://doi.org/10.4081/eb.2009.e2

Endothelin-1 | C109H159N25O32S5 - PubChem. (n.d.). Retrieved May 22, 2021, from https://pubchem.ncbi.nlm.nih.gov/compound/16212950#section=Biologic-Description

Endothelin 2, human | C115H160N26O32S4 - PubChem. (n.d.). Retrieved May 22, 2021, from https://pubchem.ncbi.nlm.nih.gov/compound/Endothelin-2_-human#section=Biologic-Description

Enrasentan | C29H30O8 | ChemSpider. (n.d.). Retrieved May 6, 2021, from http://www.chemspider.com/Chemical-Structure.155044.html

Escribá, P. V., Wedegaertner, P. B., Goñi, F. M., & Vögler, O. (2007). Lipid-protein interactions in GPCR-associated signaling. In *Biochimica et Biophysica Acta - Biomembranes* (Vol. 1768, Issue 4, pp. 836–852). Elsevier. https://doi.org/10.1016/j.bbamem.2006.09.001

Fattal, E., Couvreur, P., & Dubernet, C. (2004). "Smart" delivery of antisense oligonucleotides by anionic pH-sensitive liposomes. *Advanced Drug Delivery Reviews*,

56(7), 931–946. https://doi.org/10.1016/j.addr.2003.10.037

Ferruz, N., Doerr, S., Vanase-Frawley, M. A., Zou, Y., Chen, X., Marr, E. S., Nelson, R. T., Kormos, B. L., Wager, T. T., Hou, X., Villalobos, A., Sciabola, S., & De Fabritiis, G. (2018). Dopamine D3 receptor antagonist reveals a cryptic pocket in aminergic GPCRs. *Scientific Reports*, *8*(1), 1–10. https://doi.org/10.1038/s41598-018-19345-7

Fiser, A., Do, R. K. G., & Šali, A. (2000). Modeling of loops in protein structures. *Protein Science*, *9*(9), 1753–1773. https://doi.org/10.1110/ps.9.9.1753

Gatfield, J., Mueller Grandjean, C., Sasse, T., Clozel, M., & Nayler, O. (2012). Slow Receptor Dissociation Kinetics Differentiate Macitentan from Other Endothelin Receptor Antagonists in Pulmonary Arterial Smooth Muscle Cells. *PLoS ONE*, *7*(10). https://doi.org/10.1371/journal.pone.0047662

Hanlon, C. D., & Andrew, D. J. (2015). Outside-in signaling - A brief review of GPCR signaling with a focus on the Drosophila GPCR family. In *Journal of Cell Science* (Vol. 128, Issue 19, pp. 3533–3542). Company of Biologists Ltd. https://doi.org/10.1242/jcs.175158

Hauser, A. S., Attwood, M. M., Rask-Andersen, M., Schiöth, H. B., & Gloriam, D. E. (2017). Trends in GPCR drug discovery: New agents, targets and indications. *Nature Reviews Drug Discovery*, *16*(12), 829–842. https://doi.org/10.1038/nrd.2017.178

Hauser, A. S., Chavali, S., Masuho, I., Jahn, L. J., Martemyanov, K. A., Gloriam, D. E., & Babu, M. M. (2018). Pharmacogenomics of GPCR Drug Targets. *Cell*, *172*(1–2), 41-54.e19. https://doi.org/10.1016/j.cell.2017.11.033

Hepatopulmonary Syndrome - NORD (National Organization for Rare Disorders). (n.d.). Retrieved March 12, 2021, from https://rarediseases.org/rare-diseases/hepatopulmonary-syndrome/

Hernández-Rodríguez, M., C. Rosales-Hernández, M., E. Mendieta-Wejebe, J., Martínez-Archundia, M., & Correa Basurto, J. (2016). Current Tools and Methods in Molecular Dynamics (MD) Simulations for Drug Design. *Current Medicinal Chemistry*, *23*(34), 3909–3924. https://doi.org/10.2174/0929867323666160530144742

Herrera-Zúñiga, L. D., Moreno-Vargas, L. M., Ballaud, L., Correa-Basurto, J., Prada-Gracia, D., Pastré, D., Curmi, P. A., Arrang, J. M., & Maroun, R. C. (2019). A complex view of GPCR signal transduction: Molecular dynamics of the histamine H3 membrane receptor. In *bioRxiv* (p. 604793). bioRxiv. https://doi.org/10.1101/604793

Hollingsworth, S. A., & Dror, R. O. (2018). Molecular Dynamics Simulation for All. In *Neuron* (Vol. 99, Issue 6, pp. 1129–1143). Cell Press.

https://doi.org/10.1016/j.neuron.2018.08.011

Hospital, A., Goñi, J. R., Orozco, M., & Gelpí, J. L. (2015). Molecular dynamics simulations: Advances and applications. In *Advances and Applications in Bioinformatics and Chemistry* (Vol. 8, Issue 1, pp. 37–47). Dove Medical Press Ltd. https://doi.org/10.2147/AABC.S70333

Houde, M., Desbiens, L., & D'Orléans-Juste, P. (2016). Endothelin-1: Biosynthesis, Signaling and Vasoreactivity. *Advances in Pharmacology (San Diego, Calif.)*, *77*, 143–175. https://doi.org/10.1016/bs.apha.2016.05.002

Human endothelin-3 | C121H168N26O33S4 - PubChem. (n.d.). Retrieved May 22, 2021, from https://pubchem.ncbi.nlm.nih.gov/compound/Human-endothelin-3#section=3D-Status

Ichikawa, O., Fujimoto, K., Yamada, A., Okazaki, S., & Yamazaki, K. (2016). G-Protein/β-Arrestin-Linked Fluctuating Network of G-Protein-Coupled Receptors for Predicting Drug Efficacy and Bias Using Short-Term Molecular Dynamics Simulation. *PLOS ONE*, *11*(5), e0155816. https://doi.org/10.1371/journal.pone.0155816

Iglarz, M., Landskroner, K., Bauer, Y., Vercauteren, M., Rey, M., Renault, B., Studer, R., Vezzali, E., Freti, D., Hadana, H., Schläpfer, M., Cattaneo, C., Bortolamiol, C., Weber, E., Whitby, B. R., Delahaye, S., Wanner, D., Steiner, P., Nayler, O., … Clozel, M. (2015). Comparison of macitentan and bosentan on right ventricular remodeling in a rat model of non-vasoreactive pulmonary hypertension. *Journal of Cardiovascular Pharmacology*, *66*(5), 457–467. https://doi.org/10.1097/FJC.0000000000000296

Ivankov, D. N., Bogatyreva, N. S., Lobanov, M. Y., & Galzitskaya, O. V. (2009). Coupling between Properties of the Protein Shape and the Rate of Protein Folding. *PLoS ONE*, *4*(8), e6476. https://doi.org/10.1371/journal.pone.0006476

Ivy, D. D., McMurtry, I. F., Colvin, K., Imamura, M., Oka, M., Lee, D. S., Gebb, S., & Jones, P. L. (2005). Development of occlusive neointimal lesions in distal pulmonary arteries of endothelin B receptor-deficient rats: A new model of severe pulmonary arterial hypertension. *Circulation*, *111*(22), 2988–2996. https://doi.org/10.1161/CIRCULATIONAHA.104.491456

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, *79*(2), 926–935. https://doi.org/10.1063/1.445869

K-8794 | Ligand page | IUPHAR/BPS Guide to PHARMACOLOGY. (n.d.). Retrieved May 7, 2021, from

https://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=9651

Kato, T., Mitani, Y., Masuya, M., Maruyama, J., Sawada, H., Ohashi, H., Ikeyama, Y., Otsuki, S., Yodoya, N., Shinohara, T., Miyata, E., Zhang, E., Katayama, N., Shimpo, H., Maruyama, K., Komada, Y., & Hirayama, M. (2020). A non-selective endothelin receptor antagonist bosentan modulates kinetics of bone marrow-derived cells in ameliorating pulmonary hypertension in mice. *Pulmonary Circulation*, *10*(2), 2045894020919355. https://doi.org/10.1177/2045894020919355

Katritch, V., Cherezov, V., & Stevens, R. C. (2013). Structure-function of the G protein-coupled receptor superfamily. *Annual Review of Pharmacology and Toxicology*, *53*(1), 531–556. https://doi.org/10.1146/annurev-pharmtox-032112-135923

Kawanabe, Y., & Nauli, S. M. (2011). Endothelin. In *Cellular and Molecular Life Sciences* (Vol. 68, Issue 2, pp. 195–203). https://doi.org/10.1007/s00018-010-0518-0

Kempner, E. S. (1993). Movable lobes and flexible loops in proteins Structural deformations that control biochemical activity. In *FEBS Letters* (Vol. 326, Issues 1–3, pp. 4–10). https://doi.org/10.1016/0014-5793(93)81749-P

Kenakin, T. (1987). Agonists, partial agonists, antagonists, inverse agonists and agonist/antagonists? *Trends in Pharmacological Sciences*, *8*(11), 423–426. https://doi.org/10.1016/0165-6147(87)90229-X

Khimji, A. karim, & Rockey, D. C. (2010). Endothelin-Biology and disease. In *Cellular Signalling* (Vol. 22, Issue 11, pp. 1615–1625). https://doi.org/10.1016/j.cellsig.2010.05.002

Klebe, G. (2013). Protein–Ligand Interactions as the Basis for Drug Action. In *Drug Design* (pp. 61–88). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-17907-5_4

Koch, D. G., Bogatkevich, G., Ramshesh, V., Lemasters, J. J., Uflacker, R., & Reuben, A. (2012). Elevated levels of endothelin-1 in hepatic venous blood are associated with intrapulmonary vasodilatation in humans. *Digestive Diseases and Sciences*, *57*(2), 516–523. https://doi.org/10.1007/s10620-011-1905-6

Koehl, B., Nivoit, P., El Nemer, W. E., Lenoir, O., Hermand, P., Pereira, C., Brousse, V., Guyonnet, L., Ghinatti, G., Benkerrou, M., Colin, Y., Le van Kim, C., & Tharaux, P. L. (2017). The endothelin B receptor plays a crucial role in the adhesion of neutrophils to the endothelium in sickle cell disease. *Haematologica*, *102*(7), 1161–1172. https://doi.org/10.3324/haematol.2016.156869

Krishnan, A., Almén, M. S., Fredriksson, R., & Schiöth, H. B. (2012). The origin of GPCRs: Identification of mammalian like rhodopsin, adhesion, glutamate and frizzled GPCRs in

fungi. *PLoS ONE*, *7*(1), 29817. https://doi.org/10.1371/journal.pone.0029817

Kusafuka, T., & Puri, P. (1997). Mutations of the endothelin-B receptor and endothelin-3 genes in Hirschsprung's disease. *Pediatric Surgery International*, *12*(1), 19–23. https://doi.org/10.1007/BF01194795

Latorraca, N. R., Venkatakrishnan, A. J., & Dror, R. O. (2017). GPCR Dynamics: Structures in Motion. *Chemical Reviews*, *117*(1), 139–155. https://doi.org/10.1021/acs.chemrev.6b00177

Lefkowitz, R. J. (2013). A brief history of G-protein coupled receptors (Nobel Lecture). In *Angewandte Chemie - International Edition* (Vol. 52, Issue 25, pp. 6366–6378). https://doi.org/10.1002/anie.201301924

Li, J., Abel, R., Zhu, K., Cao, Y., Zhao, S., & Friesner, R. A. (2011). The VSGB 2.0 model: A next generation energy model for high resolution protein structure modeling. *Proteins: Structure, Function and Bioinformatics*, *79*(10), 2794–2812. https://doi.org/10.1002/prot.23106

Ling, Y., Zhang, J., Luo, B., Song, D., Liu, L., Tang, L., Stockard, C. R., Grizzle, W. E., Ku, D. D., & Fallon, M. B. (2004). The role of endothelin-1 and the endothelin B receptor in the pathogenesis of hepatopulmonary syndrome in the rat. *Hepatology*, *39*(6), 1593–1602. https://doi.org/10.1002/hep.20244

Lipid Bilayer Membranes - Chemistry LibreTexts. (n.d.). Retrieved March 16, 2021, from https://chem.libretexts.org/Bookshelves/Biological_Chemistry/Supplemental_Modules_(Biological_Chemistry)/Lipids/Applications_of_Lipids/Lipid_Bilayer_Membranes

Litalien, C., & Beaulieu, P. (2011). Molecular Mechanisms of Drug Actions: From Receptors to Effectors. In *Pediatric Critical Care* (pp. 1553–1568). Elsevier Inc. https://doi.org/10.1016/B978-0-323-07307-3.10117-X

Lutz, P. E., & Kieffer, B. L. (2013). Opioid receptors: Distinct roles in mood disorders. In *Trends in Neurosciences* (Vol. 36, Issue 3, pp. 195–206). NIH Public Access. https://doi.org/10.1016/j.tins.2012.11.002

Lyne, P. D., Lamb, M. L., & Saeh, J. C. (2006). Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *Journal of Medicinal Chemistry*, *49*(16), 4805–4808. https://doi.org/10.1021/jm060522a

Macitentan | C19H20Br2N6O4S | ChemSpider. (n.d.). Retrieved May 6, 2021, from http://www.chemspider.com/Chemical-Structure.13134960.html

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey,

A. R. N., Potter, S. C., Finn, R. D., & Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, *47*(W1), W636–W641. https://doi.org/10.1093/nar/gkz268

Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R., & Sherman, W. (2013). Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design*, *27*(3), 221–234. https://doi.org/10.1007/s10822-013-9644-8

Maguire, J. J., & Davenport, A. P. (2015). Endothelin Receptors and Their Antagonists. In *Seminars in Nephrology* (Vol. 35, Issue 2, pp. 125–136). W.B. Saunders. https://doi.org/10.1016/j.semnephrol.2015.02.002

Manna, M., Nieminen, T., & Vattulainen, I. (2019). Understanding the Role of Lipids in Signaling Through Atomistic and Multiscale Simulations of Cell Membranes. *Annual Review of Biophysics*, *48*(1), 421–439. https://doi.org/10.1146/annurev-biophys-052118-115553

Marquardt, D., Williams, J. A., Kinnun, J. J., Kucerka, N., Atkinson, J., Wassall, S. R., Katsaras, J., & Harroun, T. A. (2014). *DMPC: A Remarkable Exception to the Tocopherol's Membrane Presence*. https://doi.org/10.1016/j.bpj.2013.11.304

Martí-Renom, M. A., Stuart, A. C., Fiser, A., Sánchez, R., Melo, F., & Šali, A. (2000). Comparative protein structure modeling of genes and genomes. In *Annual Review of Biophysics and Biomolecular Structure* (Vol. 29, pp. 291–325). Annu Rev Biophys Biomol Struct. https://doi.org/10.1146/annurev.biophys.29.1.291

Martyna, G. J., Klein, M. L., & Tuckerman, M. (1992). Nosé-Hoover chains: The canonical ensemble via continuous dynamics. *The Journal of Chemical Physics*, *97*(4), 2635–2643. https://doi.org/10.1063/1.463940

Martyna, G. J., Tobias, D. J., & Klein, M. L. (1994). Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics*, *101*(5), 4177–4189. https://doi.org/10.1063/1.467468

Masaki, T. (1998). The discovery of endothelins. In *Cardiovascular Research* (Vol. 39, Issue 3, pp. 530–533). Oxford Academic. https://doi.org/10.1016/S0008-6363(98)00153-9

Mazzuca, M. Q., & Khalil, R. A. (2012). Vascular endothelin receptor type B: Structure, function and dysregulation in vascular disease. In *Biochemical Pharmacology* (Vol. 84, Issue 2, pp. 147–162). Elsevier Inc. https://doi.org/10.1016/j.bcp.2012.03.020

Miller, W. H. (1995). Theoretical Overview of Chemical Dynamics. In *Frontiers of Chemical Dynamics* (pp. 1–19). Springer Netherlands. https://doi.org/10.1007/978-94-011-0345-

9_1

Mirzadegan, T., Benko, G., Filipek, S., & Palczewski, K. (2003). Sequence analyses of G-protein-coupled receptors: Similarities to rhodopsin. In *Biochemistry* (Vol. 42, Issue 10, pp. 2759–2767). NIH Public Access. https://doi.org/10.1021/bi027224+

Molecular Dynamics Panel. (n.d.). Retrieved April 18, 2021, from https://www.schrodinger.com/sites/default/files/s3/mkt/Documentation/current/docs/Documentation.htm#desmond_help/molecular_dynamics_panel.html?Highlight=NPγT ensemble

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. (2019). An online resource for GPCR structure determination and analysis. *Nature Methods*, *16*(2), 151–162. https://doi.org/10.1038/s41592-018-0302-x

Palczewski, K. (2006). G protein-coupled receptor rhodopsin. In *Annual Review of Biochemistry* (Vol. 75, pp. 743–767). NIH Public Access. https://doi.org/10.1146/annurev.biochem.75.103004.142743

Peddi, S. R., Sivan, S. K., & Manga, V. (2018). Molecular dynamics and MM/GBSA-integrated protocol probing the correlation between biological activities and binding free energies of HIV-1 TAR RNA inhibitors. *Journal of Biomolecular Structure and Dynamics*, *36*(2), 486–503. https://doi.org/10.1080/07391102.2017.1281762

Pharmacological Glossary | Tocris Bioscience. (n.d.). Retrieved March 5, 2021, from https://www.tocris.com/resources/pharmacological-glossary

Philippe, D., Chakass, D., Thuru, X., Zerbib, P., Tsicopoulos, A., Geboes, K., Bulois, P., Breisse, M., Vorng, H., Gay, J., Colombel, J. F., Desreumaux, P., & Chamaillard, M. (2006). Mu opioid receptor expression is increased in inflammatory bowel diseases: Implications for homeostatic intestinal inflammation. *Gut*, *55*(6), 815–823. https://doi.org/10.1136/gut.2005.080887

PyMOL | pymol.org. (n.d.). Retrieved April 18, 2021, from https://pymol.org/2/

Renault, P., Louet, M., Marie, J., Labesse, G., & Floquet, N. (2019). Molecular Dynamics Simulations of the Allosteric Modulation of the Adenosine A 2 a Receptor by a Mini-G Protein. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-41980-x

Rich, S., & McLaughlin, V. V. (2003). Endothelin Receptor Blockers in Cardiovascular Disease. In *Circulation* (Vol. 108, Issue 18, pp. 2184–2190). Lippincott Williams & Wilkins. https://doi.org/10.1161/01.CIR.0000094397.19932.78

Roos, K., Wu, C., Damm, W., Reboul, M., Stevenson, J. M., Lu, C., Dahlgren, M. K.,

Mondal, S., Chen, W., Wang, L., Abel, R., Friesner, R. A., & Harder, E. D. (2019). OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *Journal of Chemical Theory and Computation*, *15*(3), 1863–1874. https://doi.org/10.1021/acs.jctc.8b01026

Rosenbaum, D. M., Rasmussen, S. G. F., & Kobilka, B. K. (2009). The structure and function of G-protein-coupled receptors. In *Nature* (Vol. 459, Issue 7245, pp. 356–363). Nature Publishing Group. https://doi.org/10.1038/nature08144

Šali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, *234*(3), 779–815. https://doi.org/10.1006/jmbi.1993.1626

Satagopam, V. P., Theodoropoulou, M. C., Stampolakis, C. K., Pavlopoulos, G. A., Papandreou, N. C., Bagos, P. G., Schneider, R., & Hamodrakas, S. J. (2010). GPCRs, G-proteins, effectors and their interactions: human-gpDB, a database employing visualization tools and data integration techniques. *Database*, *2010*(0), baq019–baq019. https://doi.org/10.1093/database/baq019

Sawaki, M., Wada, A., Tsutamoto, T., Ohnishi, M., Fujii, M., Matsumoto, T., & Kinoshita, M. (2000). Chronic Effects of an Orally Active Selective Endothelin-B-Receptor Antagonist in Experimental Congestive Heart Failure. *Journal of Cardiovascular Pharmacology*, *36*(Supplement 1), S323–S326. https://doi.org/10.1097/00005344-200036051-00094

Schneider, M. P., Boesen, E. I., & Pollock, D. M. (2007a). Contrasting actions of endothelin ETA and ETB receptors in cardiovascular disease. In *Annual Review of Pharmacology and Toxicology* (Vol. 47, pp. 731–759). NIH Public Access. https://doi.org/10.1146/annurev.pharmtox.47.120505.105134

Schneider, M. P., Boesen, E. I., & Pollock, D. M. (2007b). Contrasting actions of endothelin ETA and ETB receptors in cardiovascular disease. In *Annual Review of Pharmacology and Toxicology* (Vol. 47, pp. 731–759). NIH Public Access. https://doi.org/10.1146/annurev.pharmtox.47.120505.105134

Seibt, B. F., Schiedel, A. C., Thimm, D., Hinz, S., Sherbiny, F. F., & Müller, C. E. (2013). The second extracellular loop of GPCRs determines subtype-selectivity and controls efficacy as evidenced by loop exchange study at A2 adenosine receptors. *Biochemical Pharmacology*, *85*(9), 1317–1329. https://doi.org/10.1016/j.bcp.2013.03.005

Shahane, G., Ding, W., Palaiokostas, M., & Orsi, M. (2019). Physical properties of model biological lipid bilayers: insights from all-atom molecular dynamics simulations.

Journal of Molecular Modeling, *25*(3), 76. https://doi.org/10.1007/s00894-019-3964-0

Shen, M., & Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein Science*, *15*(11), 2507–2524. https://doi.org/10.1110/ps.062416606

Shihoya, W., Nishizawa, T., Yamashita, K., Inoue, A., Hirata, K., Kadji, F. M. N., Okuta, A., Tani, K., Aoki, J., Fujiyoshi, Y., Doi, T., & Nureki, O. (2017). X-ray structures of endothelin ETB receptor bound to clinical antagonist bosentan and its analog. *Nature Structural and Molecular Biology*, *24*(9), 758–764. https://doi.org/10.1038/nsmb.3450

Simões, S., Nuno Moreira, J., Fonseca, C., Düzgüneş, N., & Pedroso De Lima, M. C. (2004). On the formulation of pH-sensitive liposomes with long circulation times. *Advanced Drug Delivery Reviews*, *56*(7), 947–965. https://doi.org/10.1016/j.addr.2003.10.038

Singh, K. D., Unal, H., Desnoyer, R., & Karnik, S. S. (2019). Mechanism of Hormone Peptide Activation of a GPCR: Angiotensin II Activated State of AT_1R Initiated by van der Waals Attraction. *Journal of Chemical Information and Modeling*, *59*(1), 373–385. https://doi.org/10.1021/acs.jcim.8b00583

sitaxentan | C18H15ClN2O6S2 | ChemSpider. (n.d.). Retrieved May 6, 2021, from http://www.chemspider.com/Chemical-Structure.187436.html

Takahashi, H., Soma, S., Muramatsu, M., Oka, M., Ienaga, H., & Fukuchi, Y. (2001). Discrepant distribution of big endothelin (ET)-1 and ET receptors in the pulmonary artery. *European Respiratory Journal*, *18*(1).

tezosentan | C27H27N9O6S | ChemSpider. (n.d.). Retrieved May 6, 2021, from http://www.chemspider.com/Chemical-Structure.133242.html

Tien, H. T., & Ottova, A. (2003). Chapter 1 The lipid bilayer concept: Experimental realization and current applications. In *Membrane Science and Technology* (Vol. 7, Issue C, pp. 1–73). Elsevier. https://doi.org/10.1016/S0927-5193(03)80025-6

Torrens-Fontanals, M., Stepniewski, T. M., Aranda-García, D., Morales-Pastor, A., Medel-Lacruz, B., & Selent, J. (2020). How do molecular dynamics data complement static structural data of GPCRs. In *International Journal of Molecular Sciences* (Vol. 21, Issue 16, pp. 1–26). MDPI AG. https://doi.org/10.3390/ijms21165933

Vila-Viçosa, D., Teixeira, V. H., Santos, H. A. F., Baptista, A. M., & Machuqueiro, M. (2014). Treatment of ionic strength in biomolecular simulations of charged lipid bilayers. *Journal of Chemical Theory and Computation*, *10*(12), 5483–5492. https://doi.org/10.1021/ct500680q

Waller, D. G., & Sampson, A. P. (2018). Principles of pharmacology and mechanisms of drug action. In *Medical Pharmacology and Therapeutics* (pp. 3–31). Elsevier.

https://doi.org/10.1016/b978-0-7020-7167-6.00001-4

Wanderlingh, U., Branca, C., Crupi, C., Nibali, V. C., La Rosa, G., Rifici, S., Ollivier, J., & D'Angelo, G. (2017). Molecular dynamics of POPC phospholipid bilayers through the gel to fluid phase transition: An incoherent quasi-elastic neutron scattering study. *Journal of Chemistry*, *2017*. https://doi.org/10.1155/2017/3654237

Watkins, H. A., Au, M., & Hay, D. L. (2012). The structure of secretin family GPCR peptide ligands: Implications for receptor pharmacology and drug development. In *Drug Discovery Today* (Vol. 17, Issues 17–18, pp. 1006–1014). https://doi.org/10.1016/j.drudis.2012.05.005

Webb, B., & Sali, A. (2016). Comparative protein structure modeling using MODELLER. *Current Protocols in Bioinformatics*, *2016*, 5.6.1-5.6.37. https://doi.org/10.1002/cpbi.3

What do all the Prime MM-GBSA energy properties mean? | Schrödinger. (n.d.). Retrieved May 31, 2021, from https://www.schrodinger.com/kb/1875?original_search=mmgbsa

WikiGenes - Adora1 - adenosine A1 receptor. (n.d.). Retrieved March 22, 2021, from https://www.wikigenes.org/e/gene/e/11539.html

Yeagle, P. L. (2016). Membrane Receptors. In *The Membranes of Cells* (pp. 401–425). Elsevier. https://doi.org/10.1016/b978-0-12-800047-2.00015-2

Ylilauri, M., & Pentikäinen, O. T. (2013). MMGBSA as a tool to understand the binding affinities of filamin-peptide interactions. *Journal of Chemical Information and Modeling*, *53*(10), 2626–2633. https://doi.org/10.1021/ci4002475

Yuan, X., & Xu, Y. (2018). Recent trends and applications of molecular modeling in GPCR–Ligand recognition and structure-based drug design. In *International Journal of Molecular Sciences* (Vol. 19, Issue 7). MDPI AG. https://doi.org/10.3390/ijms19072105

APPENDICES

```
# Comparative modeling by the automodel class
from modeller import *              # Load standard Modeller classes
from modeller.automodel import *    # Load the automodel class
import sys

log.verbose()    # request verbose output
env = environ()  # create a new MODELLER environment to build this model in

# directories for input atom files
env.io.atom_files_directory = ['.', '../atom_files']

# Read in HETATM records from template PDBs
env.io.hetatm = True

a = automodel(env,
              alnfile='align.ali',      # alignment filename
              knowns='ETB',                # codes of the templates
              sequence='model-ETB',              # code of the target
              assess_methods=assess.DOPE)  # request GA341 model assessment
#             assess_methods=(assess.GA341, assess.DOPE))  # GA341 and DOPE
#             assess_methods=soap_protein_od.Scorer())  # assess with SOAP

a.starting_model = 1                 # index of the first model
a.ending_model   = 20                # index of the last model
                                     # (determines how many models to calculate)
a.make()                             # do the actual comparative modeling

# Get a list of all successfully built models from a.outputs
ok_models = [x for x in a.outputs if x['failure'] is None]

# Rank the models by DOPE score
key = 'DOPE score'
if sys.version_info[:2] == (2,3):
    # Python 2.3's sort doesn't have a 'key' argument
    ok_models.sort(lambda a,b: cmp(a[key], b[key]))
else:
    ok_models.sort(key=lambda a: a[key])

# Get top model
m = ok_models[0]
print("Top model: %s (DOPE score %.3f)" % (m['name'], m[key]))
```

Figure A1. MODELLER script incorporating DOPE scoring per every model created. The .ali file is the alignment input, X-ray model sequence is the template, and the N and C-terminal truncated sequence is the target.

Table A1. All of the generated ET_B models using MODELLER and their DOPE scores. The highlighted rows in grey are the final selected models based on the DOPE score and Ramachandran Plot evaluation in parallel. DOPE score and the models are sorted in the ascending order.

Bosentan	DOPE score	K-8794	DOPE score
ETB-orig5xpr.B99990003.pdb	-40646.65234	model-ETB.B99990017.pdb	-42204.51172
ETB-orig5xpr.B99990004.pdb	-40636.69922	model-ETB.B99990019.pdb	-42179.47266
ETB-orig5xpr.B99990011.pdb	-40631.05859	model-ETB.B99990005.pdb	-42164.35547
ETB-orig5xpr.B99990019.pdb	-40584.01562	model-ETB.B99990015.pdb	-42137.91797
ETB-orig5xpr.B99990001.pdb	-40502.58203	model-ETB.B99990012.pdb	-41973.42188
ETB-orig5xpr.B99990013.pdb	-40434.46875	model-ETB.B99990003.pdb	-41940.69922
ETB-orig5xpr.B99990008.pdb	-40385.50781	model-ETB.B99990008.pdb	-41909.19141
ETB-orig5xpr.B99990014.pdb	-40356.78125	model-ETB.B99990014.pdb	-41899.94531
ETB-orig5xpr.B99990016.pdb	-40353.57031	model-ETB.B99990009.pdb	-41818.85156
ETB-orig5xpr.B99990015.pdb	-40327.36719	model-ETB.B99990004.pdb	-41733.78125
ETB-orig5xpr.B99990006.pdb	-40316.81641	model-ETB.B99990013.pdb	-41724.01953
ETB-orig5xpr.B99990009.pdb	-40315.63281	model-ETB.B99990020.pdb	-41723.10547
ETB-orig5xpr.B99990010.pdb	-40294.99609	model-ETB.B99990006.pdb	-41674.69922
ETB-orig5xpr.B99990012.pdb	-40248.45312	model-ETB.B99990011.pdb	-41658.35547
ETB-orig5xpr.B99990018.pdb	-40245.75781	model-ETB.B99990007.pdb	-41642.03125
ETB-orig5xpr.B99990005.pdb	-40205.01953	model-ETB.B99990016.pdb	-41629.85547
ETB-orig5xpr.B99990017.pdb	-40197.07422	model-ETB.B99990002.pdb	-41607.02344
ETB-orig5xpr.B99990002.pdb	-40090.97656	model-ETB.B99990018.pdb	-41541.39453
ETB-orig5xpr.B99990020.pdb	-40039.25000	model-ETB.B99990001.pdb	-41503.75000
ETB-orig5xpr.B99990007.pdb	-39963.46484	model-ETB.B99990010.pdb	-41493.09766