UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Structure-function studies of bombesin and bradykinin receptors and their ligands for new therapeutic opportunities

by
Bahareh Rasaeifar

UNIVERSITAT POLITECNICA DE CATALUNYA

Structure-Function Studies of Bombesin and Bradykinin Receptors and their ligands for new therapeutic opportunities

A Thesis submitted by:

Bahareh Rasaeifar

In fulfillment of the requirements of the Doctor's degree

Supervised by:

Prof. Juan Jesús Pérez González

Department of Chemical Engineering

Barcelona, 2020

Abstract

G-protein coupled receptors represents the largest family of membrane protein in eukaryotes cells with more than 800 members. The reason of their importance lies in their function, being responsible for the primary mechanism of signal transduction from the extracellular side to the cytosol of cells. Moreover, they are responsible for controlling vesicle transport, ion channel, and enzyme activity. These receptors can be activated from a range of stimuli. Upon activation the receptors undergo a conformational change that permits the stimulation of the second messengers and cellular response. Defect in the GPCRs signal transduction process cause disorder in body organs such as cancer, color blindness, obesity, pain, depression, hyperthyroid adenoma, diabetes, schizophrenia. Hence, due to the important role they play in the body, they are target for 30% of the drugs in the market. The development of novel drugs targeting GPCRs highly depends on a solid structure-function relationships knowledge.

Despite the variety of the GPCRs encoded in the human genome, they exhibit the same architecture, consisting of 7 transmembrane domain embedded in the width the lipid bilayer which is connected to the N-terminus and three extracellular loops located in the aqueous solution of the extracellular milieu and C-terminus and three intracellular loops located in the region for protein-protein interactions, such as G-Protein/arrestin or other subunits. Despite having the common characteristics of GPCRs, they represent a large variation in the structure. Also, despite efforts carried out to determine the crystal structures of the G-PCRs, still, there are numerous challenges of for crystallization of them such as limited availability of the receptor polar surface, difficult discovering of functional proteins, limited expression of recombinant receptors in the host and minimal existence of flexible receptors to have functional diversity.

Despite the publication of a number of crystallographic structures in the last years, there are many GPCrs which structure is not determined. Homology modeling is a method that can help to study both the structure and function of the GPCRs and it can provide more information when the process continues to the virtual screening. Homology modeling with the help of molecular dynamics simulation can discover the dynamic behavior of GPCRs by the means of discovering their conformational change. This method is an economic method for discovering novel drugs as well as explaining their structure-activity relationship and pharmacology in the step before experimental methods. In the present study, we constructed an 3Dstructure of six GPCRs, including M3 muscarinic receptor, bradykinin B1 receptor, Bradykinin B2 receptor, neuromedin receptor(BB1R), Gastrin-releasing peptide receptor (BB2R), and bombesin receptor

subtype 3(BB3R). Models were constructed using the closest the closest structure available in the phylogenetic tree as template.

In regard to M3 muscarinic, the aim of the study was the evaluation of structure refinement using molecular dynamics and the effect of the template selection, presence of a ligand in the refinement process on the accuracy of the models constructed. To evaluate the effect of the template selection, the model of the M3 muscarinic receptor was constructed from different templates such as the muscarinic M2 receptor, the histamine H1 receptor, and rhodopsin. The refinement process was carried out using molecular dynamics of a system consisting in the models embedded in a lipid bilayer for about 500 ns. Also, to evaluate the effect of the ligand in the accuracy of the models, the model constructed from the M2 muscarinic receptor was refined without any specific ligand and in complex to two different ligands. The ligands were used for this study were tiotropium and the N-methylscopolamine.

In regards to the bradykinin receptor, the models of bradykinin B1 receptor and bradykinin B2 receptors were constructed in a previous study of the group. Then, pharmacophore definition was done and subsequent virtual screening led to the discovery of some novel small molecules with antagonist activity. In the present study, the pharmacophore of the two receptors was compared to understand specific molecular features of the small molecules bradykinin ligands that make them selective for each one of these receptors.

Finally, for discovering the novel chemical scaffolds for the three bombesin receptors the models of each subtype of the bombesin receptor were constructed from neurotensin receptor (NTS1), and the refinement of 500 ns was performed for each one of the models. The refinement of BB1R and BB2R was done in complex with PD176252 and in regards to the BB3R, it was carried out in complex with AM-37.

The results of the present study show that in the refinement process, the closer the template is to the target receptor and the use of a ligand in the refinement process can lead to model accurate models. Also, the selectivity of the small molecules bradykinin receptors was identified properly by comparison of the pharmacophores of two receptors . Finally, the study of the features of bombesin receptors led to the discovery of novel small molecules with antagonist activity for BB1R and important residues of BB2R and BB3R for further mutagenesis study.

Acknowledgment

First and foremost, I would like to express the deepest appreciation to my supervisor professor. Juan Jesús Pérez González, who supports me with his genius in all steps of the research and teaches me how to have a new vision in all aspects of not only science but also life.

I have to thank my parents and my brother Mohammad for their love, help, and support throughout my life. Thank you for giving me the strength to reach the dream of my life. My sisters, Sahar and Maryam deserve my wholehearted thanks as well.

Also, I'm extremely grateful to Dr. Lourdes for the emotional support of me during all these years. always you make me motivated to keep going forward and don't give up my dreams.

my completion of this project could not have been accomplished without the support of my colleague, Dr. Patricia. Thanks for helping me with your knowledge and supporting me in all the difficult situations.

I would like to offer my special thanks to Ramona, who supports me in Barcelona such a family.

I sincerely place on record my thanks to the head, Azar, and Hajar who did not only help me with her knowledge also support me as friends emotionally in the smooth conduct of my research.

Finally, I would also like to thanks my friends Shabnam and Nogol. with whom I have shared moments of deep anxiety but also of big excitement. Their presence was very important in a process that is often felt as tremendously solitaire.

Table of Contents

List of Figures

List of Tables

Preface

The present work represents an interdisciplinary effort that walks between computer-aided drug design and pharmacology. Computer-aided drug design is a methodology born in the study of the structure of matter and has its roots in one of the most fundamental paradigms in chemistry: biological activity of molecules is connected to their structure. Actually, one of the first observed structure-activity relationships was reported in the middle of the XIX century during the course of a study carried out to determine the toxicity of different alcohols in mammals[1]. In that study, it could be established an inverse relationship between toxicity and alcohol solubility in water. Structure-Activity Relationships have been key in drug design to guide the process of lead generation and optimization as explained below. Computer-aided drug design represents a powerful method to help the design and discovery of novel active molecules and consequently, to pharmacology.

The Drug Discovery Process

The search for therapeutic agents has a long history and is associated with the origins of civilization. This knowledge is recorded in the traditional medicines treaties developed in India, China, or the Middle East. Before the 19th century, most of these drugs were prepared trough water/alcohol-based plant extracts as tinctures, decoctions, or infusions[2]. In the 19th century, the remarkable advance achieved in the chemistry of natural products permitted the isolation and characterization of the active compounds embedded in medicinal plants[3]. The first active natural compounds isolated were morphine, a powerful analgesic from *Papaver somniferum,* and salicylic acid, a precursor of aspirin from the willow bark[4]. Similarly, quinine became the basis of today's anti-malarial drugs isolated from *Cinchona officials*[5]. In addition to plants, other natural products were isolated from fungi and bacterial metabolites, such as cyclosporine and lovastatin[6].

The increasing number of bioactive compounds available produced the flourishment of drugstores that together with the organic synthesis capabilities of the well-established dye industry, gave rise, eventually, to pharmaceutical industries in the second half of the 19th century[7]. The race for drug design between big pharma started with the serendipitous discovery of penicillin and other therapeutic agents during World War II[8]. Immediately after, active compounds such as Streptomycin and Tetracycline[9],

Lysergide (LSD), and cisplatin as antitumor agents[10,11] were discovered. The 1988 Laureate of Nobel prize was awarded with the discovery of azathioprine, the first immunosuppressive agent. Further developments include acyclovir as the first antiviral compound, and trimethoprim for the treatment of bacterial infections[12].

A milestone in drug research was achieved at the beginning of the 20th century when the receptor hypothesis was established: "A drug interacts with the binding site on the surface of membrane protein, or a receptor, it can produce pharmacological action"[13,14]. Although it took time to characterize them since only neurotransmitter receptors had been recognized until the middle of 1970 as explained below. This concept was key in shifting away from serendipitous discoveries to rational drug design, and the concept of structure-activity relationships permitted the optimization of compounds before their test *in vitro* studies or in animals. Actually, the use of rational drug design caused the enhancement of drug discovery during the 1960s and 1970s. This decade is considered the golden period of drug discovery by producing histamine H2 receptor antagonists or antagonists and partial agonists of the β2 adrenergic receptor[6].

These were also the years when computer-aided drug discovery was born, getting its maturity in the 90s. It had its precursor in the quantitative structure-activity relationships, a technique that correlated molecular properties with their biological activity[2]. Advances in biophysical techniques such as X-ray crystallography and NMR provided hundreds of 3D structures of many targets at atomic resolution, permitting the development of more accurate molecular modeling methods. A clear success of a combined use of these techniques in drug discovery is exemplified by the design of the first HIV protease inhibitors in the late 90s. High Throughput Screening is another technique developed in the early 90s for improving drug discovery. Although the use of this technique permitted screening a large number of compounds[15,16], the process was associated with a high cost[17]. Hence, virtual screening as a low-cost procedure became an appealing approach[18].

G-Protein Coupled Receptors as Targets for Drug Design

This work has explored a few targets that belong to the superfamily of G-protein coupled receptors (GPCRs). Knowledge of GPCRs structure-activity accumulated during the last fifty years gives us a unique opportunity to advance on the pharmacology of the receptors and take this opportunity to improve the repertoire of therapeutic agents available. With more than 800 members identified, accounting for approximately 2% of the human genome, GPCRs are involved in multiple physiological functions, being targets for more than 30% of the drugs in the market[19]. First studies on GPCRs started in the 70s with the

development of radioligand binding methods. This permitted to study receptor regulation, the discovery receptor subtypes, and to develop theories concerning the mechanisms of receptor action. In parallel, advances in medicinal chemistry permitted the discovery of novel potent and selective compounds[20]. A breakthrough in the study of GPCRs came after the isolation and purification of the first GPCRs in the 90s. This permitted to carry out pharmacological studies directly with the receptors, overcoming the difficulties associated with animal tissues[21]. A subsequent breakthrough came after the crystallization of the first GPCR by the turn of the century and its subsequent 3D structure determination at atomic resolution[22]. From this first structure, there are currently more than 60 unique receptors solved at atomic resolution[23]. More recently, the discovery that the β-arrestin-GRK system is actually multifunctional opens an ample scenario of new opportunities[24].

Aim of the Present Study

The aim of this thesis is to contribute to a deeper understanding of the structure-activity relationships of G-protein coupled receptors (GPCRs) using computational tools. Specifically, we have first revised in detail some of the limitations of the homology modeling methodology used for the construction of atomistic models of GPCRs and then, make use of this methodology to construct diverse receptor-ligand complexes. Moreover, as a proof-of-concept, models constructed were subsequently used to discover novel hits by virtual screening.

The results of the present work are summarized in three chapters. In the first, the process of structure refinement using molecular dynamics is thoroughly revised. Actually, despite the number of GPCR 3D structures disclosed at atomic resolution, this represents only a fifth of the non-sensory GPCRs described. In the absence of a crystallographic structure, homology modeling represents a good alternative. However, the more accurate the model, the better results are expected. This is the reason to study in detail the process of structure refinement.

Specifically, we carried out the study of the construction of a model of the human M3 muscarinic receptor by homology modeling using diverse templates and using different refinement conditions. Analysis of the results of this study suggests that the closer the GPCR is used as a template from the target, the most accurate the 3D structure is produced. In addition, it can be seen that the 7-TM helix bundle can be modeled more accurately than the loops and finally, that the convergence of the refinement process is faster if the receptor has a ligand bound.

The second chapter begins with the results of a previous study devoted to determining the stereochemical features that characterize small molecule binders to the bradykinin B1 and B2 receptors. Starting with the 3D models previously developed in this laboratory, we analyzed the stereochemical features that make ligands being selective for the two receptors and explain the pharmacological profile of a set of compounds identified by virtual screening.

Finally, the third chapter regards the construction of atomistic models of the diverse subtypes of bombesin receptors. Models constructed by homology modeling and the initial models were refined through molecular dynamics simulation. Then based on the final selected receptor-ligand complexes from molecular dynamics simulation and best-constructed pharmacophore, screening of data has been done in the virtual libraries via virtual screening and the disclosure of a few BB1 selective ligands. In the case of the BB2 and BB3 receptors, we developed models for the bound conformations of a few ligands. However, due to the scarce information available we could not develop specific pharmacophores to be used in virtual screening.

References

1. Portoghese, P. S. Relationships between stereostructure and pharmacological activities. *Ann. Rev. Pharmacol.* **10**, 51–76 (1970).

2. Jones, W. P., Chin, Y.-W. & Kinghorn, A. D. The role of pharmacognosy in modern medicine and pharmacy. *Curr. Drug Targets.* **7**, 247–64 (2006).

3. W.O. Foye, T.L. Lemke and D.A. Williams. Foye's Principles of Medicinal Chemistry, 5th Ed. Lippincott Williams and Wilkins, Baltimore. 2002. pp. 819-864.

4. Weissmann, G. Aspirin. *Sci. Amer.* **264**, 84–90 (1991).

5. Cragg, G. M. & Newman, D. J. Plants as a source of anti-cancer agents. *J. Ethnopharmacol.* **100**, 72–9 (2005).

6. Houston, D. R., Eggleston, I., Synstad, B., Eijsink, V. G. H. & van Aalten, D. M. F. The cyclic dipeptide CI-4 [cyclo-(l-Arg-d-Pro)] inhibits family 18 chitinases by structural mimicry of a reaction intermediate. *Biochem. J.* **368**, 23–7 (2002).

7. Jones, A. W. Early drug discovery and the rise of pharmaceutical chemistry. *Drug Test. Anal.* **3**, 337–344 (2011).

8. Tsinopoulos, C. An Evolutionary Classification of the Strategies for Drug Discovery. Manufacturing Complexity Network Conference, Cambridge. 2002

9. Balaram, P. Drug discovery: Myth and reality. *Curr. Sci.* **87**, 847–848 (2004).

10. Celebrating Bicycle Day | OUPblog. Available at: https://blog.oup.com/2013/04/bicycle-day-lsd-albert-hoffman/. (Accessed: 21st October 2019)

11. Sneader, W. Drug prototypes and their exploitation. *Eur. J. Med. Chem.* **1**, 91 (1997).

12. Chabner, B. A. In celebration of a Nobel Prize. *J. Nat. Cancer Inst.* **80**, 1512–3 (1988).

13. Langley, J. N. On the reaction of cells and of nerve-endings to certain poisons, chiefly as regards the reaction of striated muscle to nicotine and to curari. *J. Physiol.* **33**, 374–413 (1905).

14. Paul Ehrlich - Biographical. Available at: https://www.nobelprize.org/prizes/medicine/1908/ehrlich/biographical/. (Accessed: 26th October 2019)

15. Bailey, D. & Brown, D. High-throughput chemistry and structure-based design: Survival of the smartest. *Drug Discov. Today* **6**, 57–59 (2001).

16. Ratti, E. & Trist, D. The continuing evolution of the drug discovery process in the pharmaceutical industry. *Farmaco* **56**, 13–19 (2001).

17. Sawyer, J. S. *et al.* Synthesis and activity of new aryl- and heteroaryl-substituted pyrazole inhibitors of the transforming growth factor-β type I receptor kinase domain. *J. Med. Chem.* **46**, 3953–3956 (2003).

18. Venkatesh, S. & Lipper, R. A. Role of the development scientist in compound lead selection and optimization. *J. Pharm. Sci.* **89**, 145–54 (2000).

19.   Hauser, A. S., Attwood, M. M., Rask-Andersen, M., Schiöth, H. B. & Gloriam, D. E. Trends in GPCR drug discovery: New agents, targets and indications. *Nat. Rev. Drug Discov.* **16,** 829-842. (2017).

20.   Miyano, K. *et al.* History of the G protein-coupled receptor (GPCR) assays from traditional to a state-of-the-art biosensor assay. *J. Pharmacol. Sci.* **126**, 302–309 (2014).

21.   Lefkowitz, R. J. A Brief History of G-Protein Coupled Receptors (Nobel Lecture). *Angew. Chem. Int. Ed.* **52**, 6366–6378 (2013).

22.   Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739–45 (2000).

23.   Kobilka, B. The Structural Basis of G-Protein-Coupled Receptor Signaling (Nobel Lecture). *Angew. Chem. Int. Ed.* **52**, 6380–6388 (2013).

24.   Wang, W., Qiao, Y. & Li, Z. New Insights into Modes of GPCR Activation. *Trends in Pharmacol. Sci.* **39**, 367–386 (2018).

CHAPTER 1: Introduction

1.1 G-protein Coupled Receptors

Cells consist of a gel-like substance containing proteins and nucleic acids together with diverse organelles and inclusions -collectively known as cytoplasm- enclosed within a lipid bilayer membrane. The cell surface is not smooth due to the presence of diverse proteins adrift embedded in the lipidic milieu. These proteins permit the cell to perform diverse biological functions such as communication with the surroundings, make distinctions between foreign and own cells, controlling cell adhesion, the development process, and controlling main metabolic processes like photosynthesis, transmission, and production of energy and salt balance[1–3].

Based on the way membrane proteins are connected to the membrane they are classified into two types, namely: integral and peripheral proteins. The former are stickily attached to the cell membrane and are difficult to be detached. They exhibit an integral part embedded in the hydrophobic middle layer of membrane cells. The latter have a weaker attachment to the membrane or occasionally, some of them are positioned into the membrane or attached to the lipid bilayer[4,5]. Receptors represent a specific kind of integral membrane proteins, engineered to transduce a signal that modifies a function in the cell. There are thousands of receptors in the human genome, so a profound knowledge of their function is critical to understand living cells and organisms[6]. Moreover, signal transduction processes are ligand-mediated, i.e. they are produced when specific ligands are recognized at the extracellular side of the membrane, producing a conformational change in the receptors that trigger the signal transduction process across the membrane.

Among the receptors, G-protein coupled receptors (GPCRs) represent a specific kind that makes use of a G protein to complete the signal transduction process they mediate. GPCRs represent one of the largest groups of membrane proteins in eukaryotes and one of the largest gene families in the mammalian genome, encompassing more than 800 members. Actually, they are responsible for the primary mechanism of signal transduction from the extracellular side to the cytosol. Moreover, they have the responsibility of controlling vesicle transport, ion channel, and enzyme activity. They can be activated by different stimuli, from small molecules to large proteins which can be spanning ions, small signal molecules, lipids, peptides and proteins such as neurotransmitters, neuropeptides, polypeptide hormones, inflammatory mediators, cytokines or other exogenous stimuli e.g. odors, tastes, pheromones or light[7,8]. Upon ligand binding, a conformational change is induced in the receptor, causing the stimulation of second messengers and eventually a cellular response[9–11].

The fundamental role played by GPCRs in controlling the process of life makes them important targets for therapeutic intervention. Abnormal GPCRs signals can lead to a diversity of disorders that affect tissue or body organs such as diverse cancer types, color blindness, obesity, pain, depression, hyperthyroid adenoma, diabetes, schizophrenia[12–15]. About one-half of members of the family have sensory functions, mediating olfaction, taste, light perception, and pheromone signaling whereas, the other half are targets of about 30% of drugs in the market, although only a minority of receptors (~10%) are exploited therapeutically[16]. This translates into ~30% of global sales drugs target a member of the GPCR family[17–24]. Interestingly, around 70% of the non-sensory GPCRs have endogenous ligands well characterized, whereas the rest, called orphan receptors do not have a native ligand yet identified. This makes that their function cannot be well understood[7], although they are very attractive targets for drug discovery.

The development of new drugs is intimately connected with our knowledge of the structure-function relationships of this superfamily of receptors. In this direction, careful studies comparing their sequences and function provided much of the knowledge presently accumulated. However, a major breakthrough came from the publication of the crystallographic structure of bovine rhodopsin[41] and the new structures subsequently disclosed. Recently, it has also been demonstrated the capability of a specific GPCR to transduce diverse signal pathways. A short description of the knowledge accumulated in the last years is outlined below.

1.2 The Architecture of GPCRs
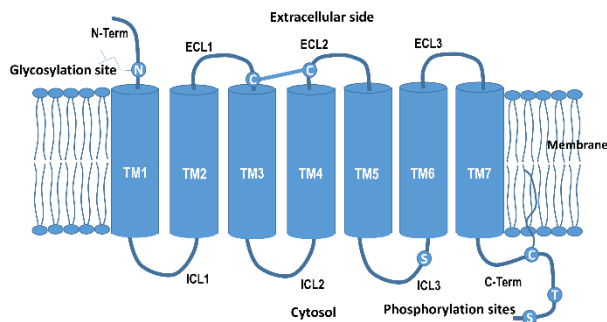
Figure 1.1 Schematic Structure of a GPCR

Despite of the variety of GPCRs encoded in the human genome, members of GPCRs exhibit a similar three-dimensional architecture, consisting of a 7-transmembrane helical domain (7TM or 7 membrane-spanning domain) forming a flattened two-layer structure known as the transmembrane bundle, oriented

roughly perpendicular to the plain of the membrane. GPCRs are single polypeptide chains, so those helices are interconnected. Accordingly, in addition to the 7-TM domain, GPCRs exhibit an extracellular amino-terminus (N-terminus), an intracellular carboxyl terminus (C-terminus), three loops located in the extracellular domain of the cell (ECL1-ECl3) and three intracellular loops (ICL1-ICL3) (Figure 1.1). The extracellular part of the receptor can be glycosylated, while the intracellular part can be phosphorylated.
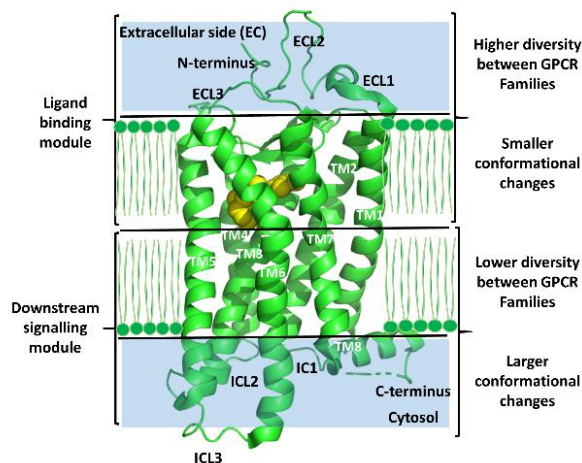
Figure 1.2 *Schematic Structure of a GPCR*

Figure 1.2 shows the way GPCRs are inserted in the membrane. Specifically, the 7-TM domain fits well within the width of the lipid membrane, and the extracellular loops are located in the aqueous solution of the extracellular milieu. Endogenous activators sit in a hydrophobic pocket (the orthosteric site) inside the helix bundle or bind to the extracellular loops or N-terminus, whereas the C-terminus is the region for protein-protein interactions, such as G-Protein/arrestin or other subunits[18]. On the other hand, the region of the helix bundle closer to the intracellular side is involved in the conformational changes leading to the binding of a G-protein at the intracellular side upon activation. Despite these common characteristics of GPCRs there is great variability of structures with major differences at the extracellular side, as explained below.

1.3 Classifications of GPCRs

Despite the similarity between members, GPCRs exhibit large differences in their sequence and function. So, it is difficult to have a coherent classification of GPCRs based on sequence similarity together with the knowledge accumulated for years to guess the function of a specific member.

Diverse methods have been used in the past to classify GPCRs including motif-based prediction[25,26], support-vector machines[27], and machine learning methods like Hidden Markov models. Currently, two overlapping classifications are used to predict the function of new sequences of GPCRs. A classification widely used in recent years is the A-F system, based on sequence and functional similarities. In this classification GPCRs are divided into six classes including Class A "rhodopsin-like family"; class B "secretin receptor family"; class C containing the metabotropic glutamate receptors; Class D containing the fungal mating pheromone receptors; class E or cyclic adenosine monophosphate (cAMP) receptors and class F "Frizzled and Smoothened receptors"[28]. Examples of some of the receptors from these groups can be seen in Figure1.3. Of these, classes D and E are not found in vertebrates. Another classification method is the GRAFS system based on phylogenetic analysis[29,30]. In this classification, GPCRs are divided into five classes named Glutamate(G), Rhodopsin(R), Adhesion(A), frizzled/taste2(F) and secretin (S). The largest and most diverse family in both classifications is the rhodopsin-like one which contains the common structural skeleton of the first structurally solved GPCRs, rhodopsin[31,32]. The difference between these two classifications lies in the adhesion family in the GRAFS system that is related to class B in the A-F system.
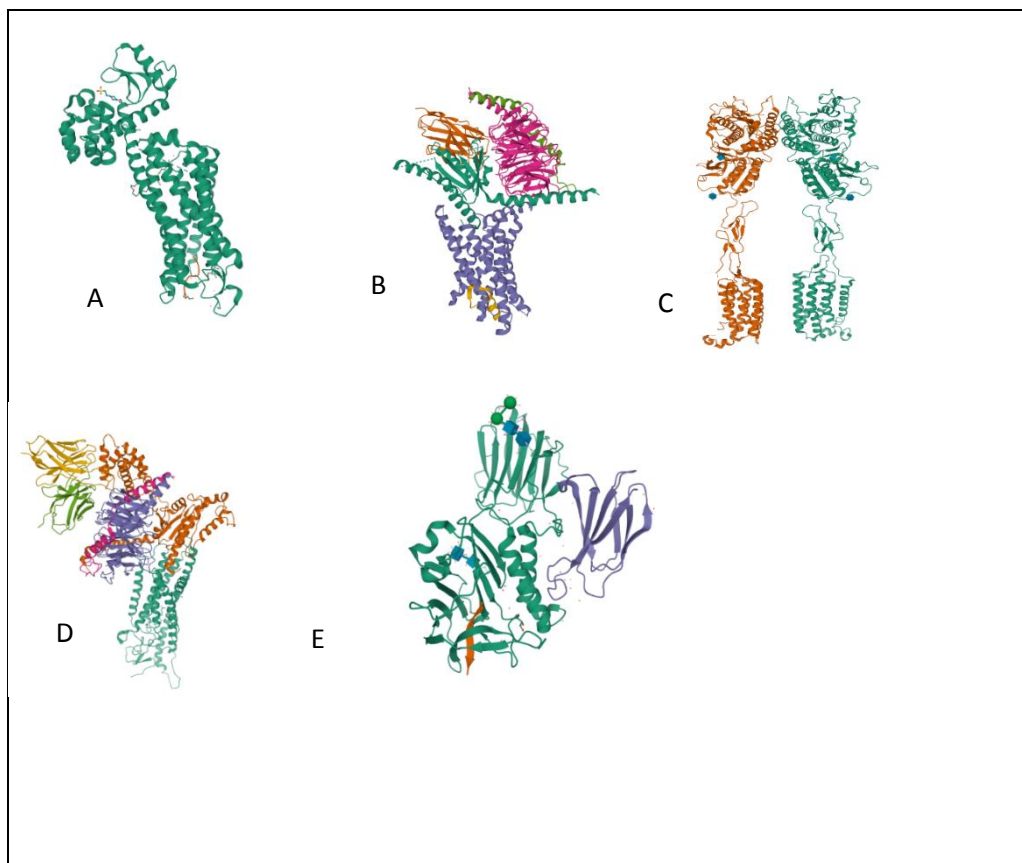
Figure 1.3 *Crystallographic structures of representative members of diverse classes of GPCRs. A) neurotensin receptor (pdb entry 4XEE); B) calcitonin receptor (pdb entry 6NIY); C) metabotropic glutamate receptor (pdb entry 6N52); D) Smoothened receptor (pdb (5KVM)*

1.3.1 Rhodopsin-like Class

This class contains 80% of the total number of receptors of the GPCRs superfamily, so it is the largest family of GPCRs and includes receptors binding hormones, neurotransmitters, neuropeptides, odors, light, taste and pheromones, and chemokines. They are well-characterized by highly conserved motifs as well as a disulfide bridge connecting first and second extracellular loops. In addition to their seven transmembrane helices, most of them have an eighth helix and a palmitoylated cysteine at the C-terminal domain, used as an anchor point of the receptor to the membrane. There is no much diversity at the N-terminus region of the receptor, being most of the diversity located in the transmembrane regions. Activation of the rhodopsin-like family mostly depends on the interaction of the ligand with the transmembrane helixes, the extracellular loops, and at least to the short N-terminal domain[33–35]. Figure 1.4A shows pictorially the neurotensin receptor as representative of this class of GPCRs.

Analysis of sequence alignments of GPCRs using multiple alignment algorithms shows that there are well-conserved residues or motifs in each one of the helices of these receptors as shown in Figure 1.4. Conserved residues in the rhodopsin family are an asparagine in TM1, an aspartate, an arginine, and a tyrosine in TM3(DRY motif), a tryptophan in TM4, a proline in TM5, the CxWP motif in TM6 and the NPxxY motif in TM7[36,37] and in most of the GPCRs there is a disulfide bridge between a cysteine in TM3 and a cysteine residue in ECl2[38,39].
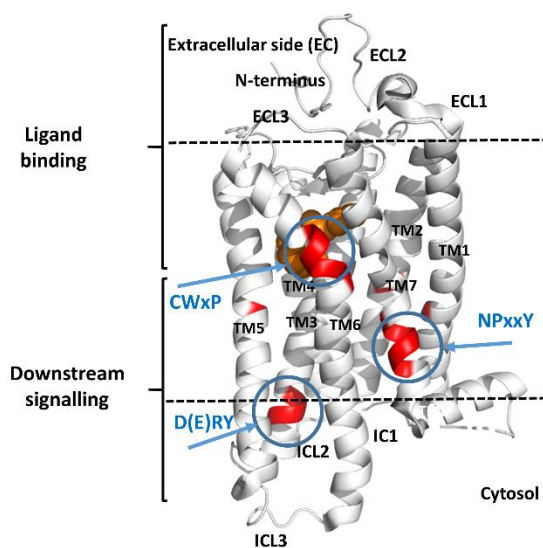
Figure1. 4 *Schematic view of the conserved residues in the members of the rhodopsin-like class of GPCRs. Structure of the adenosine A2AA receptor (PDB ID: 4EIY) (based on Lee et al., J. Med. Chem., 2018, 61, 1–46)*

The Class A rhodopsin-like family can be classified into different subfamilies based on the phylogenetic relationship between sequences[40] or alternatively, on the non-phylogenetic statistical method[41]. Phylogenetic relationships between sequences refer to the tree-based methods, which can be evolutionary model or binary. Examples of the evolutionary model are maximum parsimony or maximum likelihood that based on that Rhodopsin-like family can be divided into four groups: α, β, γ, and δ groups based on the maximum parsimony [40]. Binary or hierarchical classification of the sequences is based on a distance matrix such as neighbor-joining (NJ) method or the unweighted pair group method with arithmetic mean (UPGMA). Classification of the Rhodopsin-like family to the four mentioned group, is still under discussion in the NJ and UPGMA method as they produced different fan-like shape polygenetic tree [42,43]

The recent non-phylogenetic statistical method, that applies multidimensional scaling (MDS), can categorize the Rhodopsin-like family into four groups by the means of G0-G3[41]. G0 contains Peptide receptors, Opsins, and Melatonin receptors. The second group G1 is formed by somatostatin and opioid receptors, chemokine, and purinergic receptors, proteinase-activated receptors, and acid receptors. The G2 group include biogenic amine receptors and adenosine receptors and finally, the G3 group consists of melanocortin, phospholipids and cannabinoids, glycoprotein hormone receptors, and leucine-rich repeat (LRR) containing receptors, prostaglandin receptors, and Mas-related receptors. This classification of Rhodopsin GPCRs relies on the role of proline residues patterns in TM2 and TM5 that has been improved in the solved crystal structure of CXCR4[44]

1.3.2 Secretin-like Family

Class B receptors include 18 members that are important drug targets for the treatment of many human diseases, including type 2 diabetes, obesity, cardiovascular disease, and psychiatric disorders[30]. In addition to the 7-TM bundle, these receptors are characterized by a long N-terminal region with around 120 residues that fold in a globular domain and can be stabilized by a disulfide bond[28]. The native ligands of the class of receptors are polypeptide hormones with diverse lengths (27-141 residues). Figure 3B shows pictorially the calcitonin receptor as a representative member of this class of GPCRs.

Members of this class include glucagon and glucagon-like peptides (GLP1, GLP2), glucose-dependent insulinotropic polypeptide (GIP), secretin, vasoactive intestinal peptide (VIP), pituitary adenylate cyclase-activating polypeptide (PACAP) and the growth-hormone-releasing hormone (GHRH)[45]. According to the widely accepted 'two-domain' binding mode of class B GPCRs, the hormone peptide C

terminus initiates peptide recognition to the N-terminal domain of the receptor to bind the transmembrane domain ligand-binding pocket activating the receptor and triggering a downstream signaling cascade[46,47]. Recently, it has been discovered that the first extracellular loop (ECL1) also is important for recognizing peptide and activating of the receptor[48].

1.3.3 Glutamate Receptor Family

Members of this class are gamma-amino-butyric acid (GABA$_B$), metabotropic glutamate receptors, and extracellular calcium-sensing receptor[43,44,54–57]. Also, three taste type 1 receptors and the group of pheromone receptors, vomeronasal receptors (V2R) belong to this class[58]. Glutamate receptors participate in the modulation of synaptic transmission and neuronal excitability throughout the Central Nervous System. So, they play a major role in neural communication, memory formation, and learning and regulation. Accordingly, mutations of the genes of this receptor of this family or receptor autoantigen/antibody activity cause neurodegeneration diseases such as schizophrenia, Parkinson's disease, Alzheimer's disease, and multiple sclerosis[59–62].

These receptors form cysteine linked dimers. Their structures also include a 7-TM domain, a characteristic Venus flytrap extracellular agonist-binding domain, and a cysteine-rich domain that connects the two[49]. Figure 3C shows pictorially the stricture of the glutamate receptor as a representative of this class of GPCRs.

1.3.4 Adhesion GPCRs

The adhesion class of GPCRs comprises the second largest family with 33 and 31 members identified in the human and mouse genomes, respectively[50]. Moreover, they can be observed in invertebrates which among them sea urchins are remarkable because they contain around 100 different adhesion GPCRs[51]. It is one of the ancient evolutionary families of GPCRs and thought to be the evolutionary origin of the Secretin GPCR family.

Although there are no solved crystallographic structures yet. The specific feature of this class of GPCRs is the unusual long extracellular N-terminal containing several hundred to several thousand amino acids composed of multiple, diverse structural motifs/domains characteristically associated with cell adhesion functions. For example, autoproteolysis–inducing (GAIN) domain that trough autoproteolysis cleaves the receptor into two noncovalent-associate domains, N-terminal fragment (NTF) and cytoplasmic C-terminal fragments(CTF) (Figure 1.5). N-terminal domain is occupying most parts of the extracellular domain(ECD) while CTF contains one residue of the extracellular domain, the 7

transmembrane domain, and all regions of the intracellular domains [40,52,53]. There are 100 orphan receptors in GPCRs which most of them belong to the adhesion family[54].
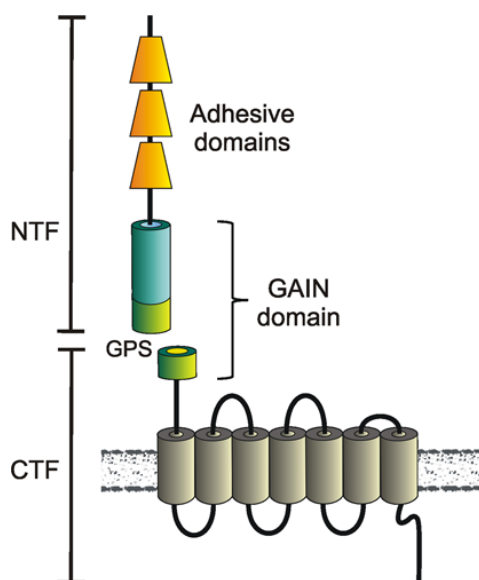
Figure 1.5 *Topology of Adhesion GPCRs (reproduced from I. Liebscher et al., Ann. N.Y. Acad. Sci. 1333 (2014) 43–64)*

1.3.5 Frizzled Class of GPCRs

The frizzled family includes 10 frizzled proteins (FZ (1-10)) and a smoothened (SMO) receptor that was isolated for the first time they from Drosophila[55]. Despite the SMO receptor has a structural similarity with the Frizzled receptors, its function is different from them[56]. Natural ligands of frizzled receptors are the group of WNT (Wingless e Int)-proteins causing the WNT/β/catenin signaling pathway that can manage embryonic development and tissue homeostases such as body shaping and regeneration of the small cells[57]. Mutation of the WNT/β/catenin pathway can lead to the appearance of different kinds of diseases like breast and colon cancer, chronic liver disease, hypertensive heart diseases such as fibrosis and neurodegeneration, etc.[58–60]. however, the SMO receptor mediates Hedgehog (HL) signaling pathway in tissue development[61] SMO receptor is recognized as an oncoprotein and antitumor agent[62].

Members of the frizzled class of GPCRs in addition to the 7TM domain they exhibit a large N terminus on the extracellular side that contains a cysteine-rich domain of ca. 120 residues used to bind their cognate ligands –the WNT family of lipoglycoproteins-. Also, there are some exceptions in the structures of the SMO receptor and recently resolved FZD4 transmembrane domain structure, which is more extension in the third extracellular (ECL2) and the sixth transmembrane (TM6) domain of SMO[63–66]. The existence of a shorter TM6 of the FD4 receptor plays a role in the connection between the cysteine-

rich domain and the transmembrane domain of the FD4 receptor[64,66]. Figure 3D shows pictorially the structure of the SMO receptor as a representative of this class of GPCRs.

1.4 Signal Transduction Pathways Mediated by GPCRs

GPCRs mediate one of the most important signal transduction pathways[67–69]. In the classical signaling, agonist binding to a GPCR causes a conformational change that permits binding to a G-protein on the cytosolic side. Subsequently, the α subunit of the heterothermic G-protein exchanges GDP for GTP, leading to its dissociation into Gα and Gβγ subunits that triggers signaling through second messengers such as adenylyl cyclases, phosphodiesterases, phospholipases, tyrosine kinases, and ion channel[70]. Also, in most of the biological systems, there is a negative feedback loop in second messenger signaling which means that when a ligand binds to the receptor and causes activation of the G-protein, the receptor can be phosphorylated in the cytoplasmic loops and the C-terminal domain by G-protein receptor kinases (GRKs), causing binding of β-arrestin and preventing interaction with the G-protein[71,72]. This competition leads to the desensitization of downstream signal pathway[73]. One of the roles of β-arrestin is the negative regulation of the G-protein signaling that prevents the harmful effect of the signaling pathway and others are desensitization, internalization, downregulation[73].

However, there are also at least five diverse activation modes different from the classical activation mechanism. These involve phenomena such as intracellular activation, dimerization activation, transactivation, biphasic activation, and biased activation[74], unveiling the complexity of signaling pathways modulated by such a family of receptors.

GPCRs can bind G-proteins, arrestins, and G-protein receptor kinases transducing diverse signals. The arrestin family includes four members: the visual arrestin involving arrestin-1 (visual arrestin) and arrestin-4 (cone arrestin) and non-visual arrestin containing β-arrestin1 and β-arrestin2 (arrestin-2 and arrestin-4) which the non-visual arrestin can express in most of the tissues and play an important role in GPCRs signal transduction pathway[75] moreover, seven GPKs (GPK1-7) exist in the human genome. GPK1 and GPK2 (expressed in the eye), GPK2, GPK3, GPK5 and GPK6 (are ubiquitous), and GPK4 (in reproductive track)[76] and about heterothermic subunits there are 21 Gα,6 Gβ and 12 Gδ subunits in human[77]. Classification of G-Proteins is based on the sequence similarity of the Gα subunit and they classified into four major classes (Gs, Gi/o, Gq/11, and G12/13)[77]. Combining different Gα, Gβ and Gδ subunits cause the production of multiple heterothermic complexes that effect the specificity of both GPCRs and their signal transduction pathway[78].

Analysis of the crystallographic structures of diverse GPCRs-agonist and GPCRs-G-protein complexes has shed light on the mechanism of ligand binding and conformational changes produced by ligand[79]. Specifically, conserved motifs of the receptors have a key role in transforming GPCRs back and forth inactive and active states by molecular micro switches. For example, the inactive form of the rhodopsin exhibits an ionic lock between R135 and E134 of the conserved E(D)RY350 motif of TM3 and E247 and T251 of TM6 that is broken in the activated form of the receptor. In addition, binding of the ligand causes a conformational change of W648 in the *CW648xP motif, which leads to the movement of the sixth transmembrane helix outwards in the active state[80,81]. Furthermore, in the activation process highly conserved proline residues in TM5, TM6, and TM7 induce deformation of TM5, translating and rotating of the TM6 and repositioning of TM7[82]. Similarly, prolines in TM4 and TM5 of class B and class F receptors and TM6 and TM7 of class C perform a similar role. Accordingly, it seems that proline residues play an important role in the conformational changes observed in the activated receptors.*

1.5 Pharmacological Profile of GPCRs Ligands

There are four groups of ligands that provide different signal transduction pathways when binding to the receptors including agonist, antagonist, partial agonist, and inverse antagonist. Agonists are kind of ligands that produce an appropriate response when binds to the receptor. An agonist increases the activity of its corresponding receptor above the basal level. The curve dose-response follows a sigmoid curve, as shown in Figure 6. There are ligands that despite producing a response, it is not as high as a full agonist, even binding to the same receptor. These molecules are known as partial agonists (Figure 6). Moreover, GPCRs normally exhibit a constitutive level of activity in the absence of a ligand, the activity of the receptor increases above the basal activity by binding to its corresponding agonist. In contrast, there are ligands that decrease the basal activity of a receptor. These ligands are called inverse agonists[83] (Figure 6). Thus, an inverse agonist is a ligand that induces a reversal of constitutive activity producing a pharmacological response opposite to the agonists[84]. When a molecule prevents the activity of an agonist or an inverse agonist it is called antagonist[85] (Figure 1.6).
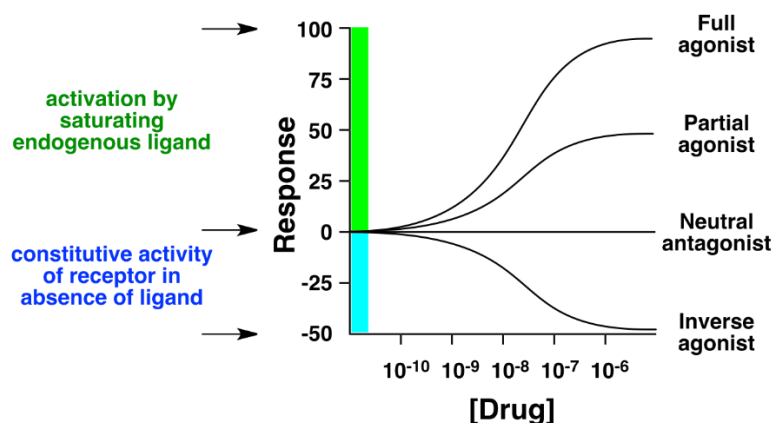
Figure 1.6 *Dose-response curves of diverse types of ligands.*

Most of the human diseases are correlated with the constitutive activity of specific GPCRs due to mutations[86]. Accordingly, the discovery of specific inverse agonists and antagonists can open doors for treating diverse human disorders[83,87].

During decades the pharmacological behavior of different ligands has been explained using a two-state model[88]. According to this model, receptors in the membrane are in an equilibrium between two states (conformations): the active and the non-active. Depending on the relative population between the two states, a specific receptor can display more or less basal activity. In this model, agonists increase the population of receptors in the active state, whereas inverse agonists lower its population. Antagonists, on the other hand, bind equally to the two states and then, keep the balance unaltered and preventing agonists or inverse agonists to produce any effect. Nowadays it is recognized that receptors occupy diverse discrete conformations of an ensemble and these conformations are associated with diverse signaling mechanisms. Thus, biased ligands bind to specific conformations that can promote pharmacological responses that differ from responses to traditional agonists.

1.6 GPCRs and the Drug Discovery Process

As mentioned above, GPCRs mediate a wide range of physiological responses affecting the immune, cardiovascular and endocrine systems, among others. Defectors signaling may cause pathophysiological disorders including neurodegenerative, immune, metabolic, cardiovascular, psychiatric, and oncologic diseases, among others. This situation makes GPCRs attractive targets for therapeutic intervention[89–92].

Whereas about 30-50% of drugs in the market target a GPCR, only 10% of the members of the superfamily are exploited therapeutically[16,93]. This underlines the opportunity of expanding to new druggable receptors in order to develop novel therapeutics. Most of the GPCRs targeted by approved

20

drugs belong to class A (~94%) and specifically, all aminergic receptors are established as drug targets. Members of the class B account only for a ~4% of the receptors targeted. On the other hand, several approved drugs are targeting the same receptor. Actually, it is estimated that there are about 10 distinct approved drugs, on average, targeting a specific GPCR. In contrast, there is a need to develop novel drugs for peptide receptors. Peptide receptors represent almost a third of the non-sensorial GPCRs and although their many compounds disclosed, they are basically peptide ligands[121]. However, peptides have many drawbacks as drugs, since their oral bioavailability is limited, they have high molecular mass, it is not easy for them to pass through the membrane which causes to having limited biodistribution and is easily degraded by proteases[122]. To overcome these problems non-peptidic or peptidomimetic ligands can be used in drug discovery[123–125].

This indicates the necessity of expanding to new druggable receptors in order to develop novel medications. Identification and exploitation of new targets are specially warranted for diseases with large unmet medical needs. The situation is changing by looking at the number of clinical trials on-going. If successful, it may bring up to 20% the number of GPCRs exploited therapeutically[16]

In addition, to target additional GPCRs, knowledge of the structure-activity relationships of these receptors gives opportunities to explore alternative ways of drug action. Presently, most of the approved drugs act as antagonists (~53%) and agonists (~42%), both binders to the orthosteric site. The percentage of recognized inverse antagonists is only ~1%, although it is thought that around 85% of all GPCR antagonists are actual inverse agonists[94].

Allosteric modulators can block the action of an agonist or antagonist (negative modulators) or modulate the signal transduction of a GPCR (positive modulators) by binding to a site that differs from the orthosteric. An advantage of designing allosteric modulators is the possibility to discover more selective drugs since the amino acid sequence in the extracellular loops is less conserved than the transmembrane domain among members of the same family. Around 3% of the approved drugs targeting GPCRs are allosteric modulators. Approved allosteric modulators include cinacalcet, a positive allosteric modulator of the calcium-sensing receptor for the treatment of hyperparathyroidism, and maraviroc, a negative allosteric modulator of the CCR5 chemokine receptor for prevention of cellular entry of HIV-1[95].

During the past decade, biased ligands have been discovered and developed for many GPCRs, such as the µ opioid receptor, the angiotensin II receptor type 1, the dopamine D2 receptor, and many others. These ligands exhibit selectivity to signal either through a specific G-protein or through a specific β-

arresting pathway for a given GPCR. For example, the endocannabinoid AEA is biased for Gi over Gs at the cannabinoid CB1R compared to the synthetic agonist WIN-55.

1.7 GPCRs Crystallographic Structure in Drug Discovery

The availability of high-resolution crystallographic structures of GPCRs has provided new insights into our understanding of GPCRs structure and function and has been key for the discovery of novel drugs. Due to its high endogenous concentration in bovine retina, rhodopsin was the first crystal structure reported at atomic resolution in 2000[96,97]. After this breakthrough, difficulties associated with expression, purification, and crystallization made the process difficult for other GPCRs. Technical advancements were necessary to overcome the numerous difficulties associated such as using lipid phases[98–100], lipid cubic phases (LCP) for crystal structure[101], and participating soluble fusion partners[101,102]. Other problems were also found regarding the stability of the proteins at the time of removing protein-membrane from the native lipid environment and using a mild detergent that prevents protein denaturation that was solved using different lipid phases[98–100]. Also, detergent optimization was done through vapor diffusion[97], and the phase of unknown targets of crystallographic structures was determined by molecular replacement[103]. besides, using site-directed mutagenesis studies helps for the stabilization of the receptor and recognizing the functional activity of GPCRs[101]. Also, the flexibility of the loops and termini of the GPCRs was handled by protein modification[101,103–106]. These advancements lead to the publication of the structure of the β-2A adrenergic receptor at atomic resolution [103,104].

Despite all the efforts carried out for determining the crystal structure of diverse GPCRs, still, there are numerous challenges for crystallization of the GPCRs such as the limited availability of the receptor polar surfaces to have crystal contacts because of embedding crystals in lipid bilayers, difficulty in discovering functional protein, limited expression of recombinant receptors in hosts and in final minimal existence of the flexible receptors which is necessary for having functional diversity[107] and coupling to various signaling mechanism. Hence, due to these obstacles in the crystallization and expression of the GPCRs, they are the largest "terra incognita" of biological structures. Despite the difficulties associated there are 64 unique structures of GPCRs reported, meaning that only 15% of the non-olfactory receptors are covered[21]. Most of the solved crystal structures correspond to class A (84%) with a few structures in other classes: B (10%), C(3%), and F (3%) classes. Actually, there is not yet a structure available of any member of the adhesion and taste family which have an interest in pharmacology since they are not activated by an endogenous agonist[11].

In the absence of the crystal structure, the structure of the proteins can be constructed by homology modeling and these structures can provide enough information for further study through virtual screening. Homology modeling[108] with the help of molecular dynamics simulation is a good method to discover conformational changes of the GPCRs or dynamic behavior of them[109–111] which will be explained in chapter 2. This represents an economic method to discover new drugs and explaining their structure-activity relationship and pharmacology of them before starting experimental method[19]. So this method is widely used as a substantial method of HTS.

1.8 Use of Computational Techniques in GPCRs Drug Discovery

1.8.1 Screening of Ligands Binding to the Orthosteric Site

The structure-based drug design method can identify ligands for the different orthosteric and allosteric binding sites for various receptors from the GPCRs family. Structure-based virtual screening was used to screen Lundbeck and Zinc databases using the inactive structure of the β_2A adrenergic receptor bound to an inverse agonist[112]. The study provided a list of hits that after tested showed 36% and 12% positive results from the Lundbeck and Zinc databases respectively. The same structure was used in a subsequent study that led to the discovery of 25 novel hits, 6 of which showed binding affinity less than 4 µM, one of them showed affinity of 9 nM, and five of them were registered as an inverse agonist.

After that the structure of the adenosine $A_{2A}R$ has been discovered, so most of the studies focused on finding hits that can bind to this receptor. In 2010, Katrichch et al[113] discovered 9 high-affinity hits (> 3.3 kcal/mole per heavy atoms) that were good starting points for further lead optimization, also another similar research about $A_{2A}R$ on zinc database could discover two novel hits[114].

One of the most successful studies using structure-based drug discovery corresponds to the identification of highly potent ligands for the adenine receptor[115]. In this study, the structure of $A_{2A}R$ was constructed from the β_{1A} adrenergic receptor by homology modeling that was subsequently used for virtual screening using a library of 545k available compounds. The study led to the identification of AZD4635 that is now in clinical assays as immunooncological drug[116]. Numerous works have reported the discovery of orthosteric ligands of diverse GPCRs using computational methods[108,116–119].

1.8.2 Discovery of Selective Ligands

Discovery of the selective compounds of the different subtypes of a receptor is one of the key points in drug discovery to consider, however, this is not an easy task due to the similarity of the binding

sites. Katritch et al.[120] reported the discovery of selective ligands for adenosine receptors based on homology models. However, in a similar study Kolb et al.[121] showed the limitations of docking studies for the discovery of selective ligands. They demonstrate that screening based on the single subtype has limitations to predict selectivity for different subtypes.

There are examples where structure-based discovery is used to identify new ligands with unexplored chemotypes and physical properties, leading to new biologic functions. This can be exemplified in the case of the identification of selective ligands using a structure-based docking for the muscarinic M_2R and M_3R^{122}. In that study, the authors were able to identify diverse ligands binders of each of the two receptors, being one of them a partial agonist at the M_3 receptor without measurable M_2 agonism. A similar protocol was used to screen molecules for the 5-HT_{1B} receptor, leading to the discovery of a compound with higher selectivity to 5-HT_{1B} (>300-fold) over the 5-HT_{2B}[123].

Recently, the discovering agonist-bound and biased-agonist bound GPCRs structures led to evaluating the effect of the molecular docking on discovering specific ligands. The first research that focuses on these structures, was done by Weiset et al.[124] in 2013. They used the active-state structure of $\beta_{2A}R$ to construct the active conformation of the dopamine D_2R by homology modeling and subsequently, they carried out a virtual screening using the ZINC database. The result of the study was the discovery of various low-affinity agonists and partial agonists for this receptor.

1.8.3 Screening of Compounds Binders to Allosteric Sites

Comparisons between the orthosteric and allosteric sites show that the former is more conserved than the latter, which is connected to the fact that endogenous ligands binding to the orthosteric site tend to activate all receptor subtypes[125]. There are three different profiles of allosteric ligands: they can produce a positive modulation (PAM), negative modulation (NAM), and a neutral effect (SAM). The effect produced by a ligand can be understood in terms of the effect it can produce on the receptor by freezing a specific conformation of the ensemble[126].

More than 75 small molecules with an allosteric mode of action have been reported for GPCRs. Most of them are ligands of members of the rhodopsin-like family, but a few for the B, C, and F[127]. Two NAMs allosteric ligands are presently in clinical trials: mozobile and selzentry have been discovered within the class A GPCRs[128]. The former is NAM ligand of the chemokine CXCR4 receptor and produces the release of stem cells into the bloodstream after autologous stem cell transplantation. The latter is a high-affinity NAM ligand of the chemokine CCR5 receptor and is used for HIV treatment together with an antiretroviral

agent. Reparixin[129] is a NAM modulator of the chemokine CXCR1/2 in preclinical studies that could be used for acute lung injury.

The newly discovered type of ligands are the bitopic which can bind to both allosteric and orthosteric binding sites simultaneously. These ligands have the two fragments binding to each site connected by a linker. However, the design of the linker requires to consider some molecular features in regard to the attachment point, length, flexibility, and composition[130]. The design of bitopic ligands has promising results. For example, the iperoxo-derived bitopic agonists: iper-6-path and iper-6-naph exploit the allosteric vestibule to control the extent of receptor movement to govern a hierarchical order of G-protein coupling. Thus, binding of the parent compound to the M_2R receptor results in Gi activation, whereas binding of that to both allosteric and orthosteric sites simultaneously permits to activate Gi- and Gs-mediated signaling events[131]. Compound 64 (SB269652) is another bitopic ligand discovered by Silvano et al.[132,133] from a non-selective antagonist of the dopamine D2/D3 receptors. When the ligand binds to the receptor it experiences a charge-charge interaction between the ligand amino nitrogen and carboxyl of D114 of the orthosteric site and forms and additional hydrogen bond with E95[265] of the allosteric site. Another compound discovered recently is the antipsychotic drug Brexpiparazole[134,135]. The discovery of bitopic ligands led to the recognition of a novel era of ligand design for the GPCRs family.

1.9 References

1.   Lodish H, Berk A, Zipursky SL, et al. Molecular Cell Biology. 4th edition. W. H. Freeman . New York, NY. 2000, pp. 579-594.

2.   Boyd, D., Schierle, C. & Beckwith, J. How many membrane proteins are there? *Protein Sci.* **7**, 201–205 (1998).

3.   Krivov, G. G., Shapovalov, M. V, Dunbrack, R. L. & Jr. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* **77**, 778–95 (2009).

4.   Alenghat, F. J. & Golan, D. E. Membrane protein dynamics and functional implications in mammalian cells. in *Curr. Top. Mem.* **72**, 89–120 (Academic Press Inc., 2013).

5.   Johnson, J. E. & Cornell, R. B. Amphitropic proteins: Regulation by reversible membrane interactions. *Mol. Mem. Biol.* **16**, 217–235 (1999).

6.   Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. Molecular biology of the cell. 4th edn. *Ann. Botany.* **91**, 401–401 (2003).

7.   Buck, L. Liberles, S. Methods and G proteins for screening and identifying ligands of G protein-coupled receptors. *U.S. Patent No. 9,658,208* (2017).

8.   Lismaa, T. P., Biden, T. J. & Shine, J. G Protein-Coupled Receptors. Springer Berlin Heidelberg, 1995.

9.   Weiss, F. U., Daub, H. & Ullrich, A. Novel mechanisms of RTK signal generation. *Curr. Opin. Genet. Dev.* **7**, 80–86 (1997).

10.   Wess, J. Molecular basis of receptor/G-protein-coupling selectivity. *Pharmacol. Ther.* **80**, 231–64 (1998).

11.   Sakmar, T. P. Introduction: G-Protein Coupled Receptors. *Chem. Rev.* **117**, 1–3 (2017).

12.   Drews, J. Drug discovery: a historical perspective. *Science* **287**, 1960–4 (2000).

13.   Karnik, S. S., Gogonea, C., Patil, S., Saad, Y. & Takezako, T. Activation of G-protein-coupled

receptors: a common molecular mechanism. *Trends Endocrinol. Metab.* **14**, 431–7 (2003).

14.   Insel, P. A., Tang, C.-M., Hahntow, I. & Michel, M. C. Impact of GPCRs in clinical medicine: Monogenic diseases, genetic variants and drug targets. *Biochim. Biophys. Acta - Biomembr.* **1768**, 994–1005 (2007).

15.   Spiegel, A. M. & Weinstein, L. S. Inherited Diseases Involving G Proteins and G Protein–Coupled Receptors. *Annu. Rev. Med.* **55**, 27–39 (2004).

16.   Hauser, A. S., Attwood, M. M., Rask-Andersen, M., Schiöth, H. B. & Gloriam, D. E. Trends in GPCR drug discovery: new agents, targets and indications. *Nat. Rev. Drug Discov.* **16**, 829–842 (2017).

17.   Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How many drug targets are there? *Nat. Rev. Drug Discov.* **5**, 993–996 (2006).

18.   Rask-Andersen, M., Masuram, S. & Schiöth, H. B. The Druggable Genome: Evaluation of Drug Targets in Clinical Trials Suggests Major Shifts in Molecular Class and Indication. *Annu. Rev. Pharmacol. Toxicol.* **54**, 9–26 (2014).

19.   Santos, R. *et al.* A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19–34 (2017).

20.   Hauser, A. S., Attwood, M. M., Rask-Andersen, M., Schiöth, H. B. & Gloriam, D. E. Trends in GPCR drug discovery: new agents, targets and indications. *Nat. Rev. Drug Discov.* **16**, 829–842 (2017).

21.   Pándy-Szekeres, G. *et al.* GPCRdb in 2018: adding GPCR structure models and ligands. *Nucleic Acids Res.* **46**, D440–D446 (2018).

22.   Eglen, R. M. Emerging concepts in GPCR function--the influence of cell phenotype on GPCR pharmacology. *Proc. West. Pharmacol. Soc.* **48**, 31–4 (2005).

23.   Dastmalchi, S., Church, W. B. & Morris, M. B. Modelling the structures of G protein-coupled receptors aided by three-dimensional validation. *BMC Bioinformatics* **9**, S14 (2008).

24.   Sriram, K. & Insel, P. A. G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? *Mol. Pharmacol.* **93**, 251–258 (2018).

25. Flower, D. R. & Attwood, T. K. Integrative bioinformatics for functional genome annotation: trawling for G protein-coupled receptors. *Semin. Cell Dev. Biol.* **15**, 693–701 (2004).

26. Holden, N. & Freitas, A. A. Hierarchical classification of G-protein-coupled receptors with a PSO/ACO algorithm. *Proc. IEEE Swarm Intell. Symp. (SIS'06) IEEE Press.* 77–84 (2006).

27. Karchin, R., Karplus, K. & Haussler, D. Classifying G-protein coupled receptors with support vector machines. *Bioinformatics* **18**, 147–159 (2002).

28. Hu, G.-M., Mai, T.-L. & Chen, C.-M. Visualizing the GPCR Network: Classification and Evolution. *Sci. Rep.* **7**, 15495 (2017).

29. Schiöth, H. B. & Fredriksson, R. The GRAFS classification system of G-protein coupled receptors in comparative perspective. *Gen. Comp. Endocrinol.* **142**, 94–101 (2005).

30. Afroze, S. *et al.* The physiological roles of secretin and its receptor. *Ann. Transl. Med.* **1**, 3 (2013).

31. Costanzi, S., Siegel, J., Tikhonova, I. G. & Jacobson, K. A. Rhodopsin and the others: a historical perspective on structural studies of G protein-coupled receptors. *Curr. Pharm. Des.* **15**, 3994–4002 (2009).

32. Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739–45 (2000).

33. Alexander, S. P. *et al.* The Concise Guide to Pharmacology 2017/18: Overview. *Br. J. Pharmacol.* **174**, S1–S16 (2017).

34. Attwood, T. K. & Findlay, J. B. Fingerprinting G-protein-coupled receptors. *Protein Eng.* **7**, 195–203 (1994).

35. Joost, P. & Methner, A. Phylogenetic analysis of 277 human G-protein-coupled receptors as a tool for the prediction of orphan receptor ligands. *Genome Biol.* **3**, RESEARCH0063-1 (2002).

36. Mirzadegan, T., Benkö, G., Filipek, S. & Palczewski, K. Sequence analyses of G-protein-coupled receptors: similarities to rhodopsin. *Biochemistry* **42**, 2759–67 (2003).

37. Ballesteros, J. A. & Weinstein, H. [19] Integrated methods for the construction of three-

dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* **25**, 366–428 (1995).

38.   Wheatley, M. *et al.* Lifting the lid on GPCRs: the role of extracellular loops. *Br. J. Pharmacol.* **165**, 1688–1703 (2012).

39.   Suku, E. & Giorgetti, A. Common evolutionary binding mode of rhodopsin-like GPCRs: Insights from structural bioinformatics. *AIMS Biophys.* **4**, 543–556 (2017).

40.   Fredriksson, R., Lagerström, M. C., Lundin, L.-G. & Schiöth, H. B. The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogon Groups, and Fingerprints. *Mol. Pharmacol.* **63**, 1256–1272 (2003).

41.   Pelé, J., Abdi, H., Moreau, M., Thybert, D. & Chabbert, M. Multidimensional Scaling Reveals the Main Evolutionary Pathways of Class A G-Protein-Coupled Receptors. *PLoS One* **6**, e19094 (2011).

42.   Devillé, J., Rey, J. & Chabbert, M. An Indel in Transmembrane Helix 2 Helps to Trace the Molecular Evolution of Class A G-Protein-Coupled Receptors. *J. Mol. Evol.* **68**, 475–489 (2009).

43.   Surgand, J.-S., Rodrigo, J., Kellenberger, E. & Rognan, D. A chemogenomic analysis of the transmembrane binding cavity of human G-protein-coupled receptors. *Proteins Struct. Funct. Bioinforma.* **62**, 509–538 (2005).

44.   Wu, B. *et al.* Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists. *Science,* **330**, 1066–107 (2010).

45.   Harmar, A. J. Family-B G-protein-coupled receptors. *Genome Biol.* **2**, REVIEWS3013-1 (2001).

46.   Miller, L. J., Dong, M. & Harikumar, K. G. Ligand binding and activation of the secretin receptor, a prototypic family B G protein-coupled receptor. *Br. J. Pharmacol.* **166**, 18–26 (2012).

47.   Singh, K., Leprince, J., Lefranc, B., Vaudry, D. & Chow, B. Discovering small compound modulators for Secretin receptor. Cambridge Healthtech Institute's 13th Annual Drug Discovery Chemistry, San Diego, CA, 2-6 (2018).

48.   Cao, C., Zhang, H., Yang, Z. & Wu, B. Peptide recognition, signaling and modulation of class B G

protein-coupled receptors. *Curr. Opin. Struct. Biol.* **51**, 53–60 (2018).

49.   Congreve, M., Oswald, C. & Marshall, F. H. Applying Structure-Based Drug Design Approaches to Allosteric Modulators of GPCRs. *Trends Pharmacol. Sci.* **38**, 837–847 (2017).

50.   Paavola, K. J. & Hall, R. A. Adhesion G protein-coupled receptors: signaling, pharmacology, and mechanisms of activation. *Mol. Pharmacol.* **82**, 777–83 (2012).

51.   Whittaker, C. A. *et al.* The echinoderm adhesome. *Dev. Biol.* **300**, 252–266 (2006).

52.   Monk, K. R. *et al.* Adhesion G Protein-Coupled Receptors: From In Vitro Pharmacology to In Vivo Mechanisms. *Mol. Pharmacol.* **88**, 617–623 (2015).

53.   Bjarnadóttir, T. K., Fredriksson, R. & Schiöth, H. B. The Adhesion GPCRs: A unique family of G protein-coupled receptors with important roles in both central and peripheral tissues. *Cell. Mol. Life Sci.* **64**, 2104–2119 (2007).

54.   Hamann, J. *et al.* International Union of Basic and Clinical Pharmacology. XCIV. Adhesion G Protein-Coupled Receptors. *Pharmacol. Rev.* **67**, 338–367 (2015).

55.   Chan, S. D. *et al.* Two homologs of the Drosophila polarity gene frizzled (fz) are widely expressed in mammalian tissues. *J. Biol. Chem.* **267**, 25202–7 (1992).

56.   Schulte, G. International Union of Basic and Clinical Pharmacology. LXXX. The Class Frizzled Receptors. *Pharmacol. Rev.* **62**, 632–667 (2010).

57.   Logan, C. Y. & Nusse, R. THE WNT SIGNALING PATHWAY IN DEVELOPMENT AND DISEASE. *Annu. Rev. Cell Dev. Biol.* **20**, 781–810 (2004).

58.   Zhao, Y. *et al.* An essential role for Wnt/β-catenin signaling in mediating hypertensive heart disease. *Sci. Rep.* **8**, 8996 (2018).

59.   Majidinia, M., Aghazadeh, J., Jahanban-Esfahlani, R. & Yousefi, B. The roles of Wnt/β-catenin pathway in tissue development and regenerative medicine. *J. Cell. Physiol.* **233**, 5598–5612 (2018).

60.   Markham, N. O., Fry, W. H., Cao, Z., Huh, W. J. & Coffey, R. J. Su2036 - TNKS2 Promotes WNT/β-

Catenin Signaling through Parylation of NKD2 in Colon Cancer Cells. *Gastroenterol.* **154**, S-1365 (2018).

61.   Gorojankina, T. Hedgehog signaling pathway: a novel model and molecular mechanisms of signal transduction. *Cell. Mol. Life Sci.* **73**, 1317–1332 (2016).

62.   Sharpe, H. J., Wang, W., Hannoush, R. N. & de Sauvage, F. J. Regulation of the oncoprotein Smoothened by small molecules. *Nat. Chem. Biol.* **11**, 246–255 (2015).

63.   Byrne, E. F. X. *et al.* Structural basis of Smoothened regulation by its extracellular domains. *Nature* **535**, 517–522 (2016).

64.   Zhang, X. *et al.* Crystal structure of a multi-domain human smoothened receptor in complex with a super stabilizing ligand. *Nat. Commun.* **8**, 1-10 (2017).

65.   Huang, P. *et al.* Structural Basis of Smoothened Activation in Hedgehog Signaling. *Cell* **174**, 312-324 (2018).

66.   Yang, S. *et al.* Crystal structure of the Frizzled 4 receptor in a ligand-free state. *Nature* **560**, 666–670 (2018).

67.   Luttrell, L. M. Transmembrane Signaling by G Protein-Coupled Receptors. in *Trans. Sig. Prot.* **332**, 3-49 (Humana Press, 2006).

68.   Marinissen, M. J. & Gutkind, J. S. G-protein-coupled receptors and signaling networks: emerging paradigms. *Trends Pharmacol. Sci.* **22**, 368–76 (2001).

69.   Pierce, K. L., Premont, R. T. & Lefkowitz, R. J. Seven-transmembrane receptors. *Nat. Rev. Mol. Cell Biol.* **3**, 639–650 (2002).

70.   Cabrera-Vera, T. M. *et al.* Insights into G Protein Structure, Function, and Regulation. *Endocr. Rev.* **24**, 765–781 (2003).

71.   DeGraff, J. L., Gurevich, V. V. & Benovic, J. L. The Third Intracellular Loop of α_2-Adrenergic Receptors Determines Subtype Specificity of Arrestin Interaction. *J. Biol. Chem.* **277**, 43247–43252 (2002).

72.	Marion, S., Oakley, R. H., Kim, K.-M., Caron, M. G. & Barak, L. S. A β-Arrestin Binding Determinant Common to the Second Intracellular Loops of Rhodopsin Family G Protein-coupled Receptors. *J. Biol. Chem.* **281**, 2932–2938 (2006).

73.	Smith, J. S. & Rajagopal, S. The β-Arrestins: Multifunctional Regulators of G Protein-coupled Receptors. *J. Biol. Chem.* **291**, 8969–8977 (2016).

74.	Wang, W., Qiao, Y. & Li, Z. New Insights into Modes of GPCR Activation. *Trends Pharmacol.l Sci.* **39**, 367–386 (2018).

75.	Patwari, P. & Lee, R. T. An expanded family of arrestins regulate metabolism. *Trends Endocrinol. Metab.* **23**, 216–222 (2012).

76.	Premont, R. T. & Gainetdinov, R. R. Physiological Roles of G Protein–Coupled Receptor Kinases and Arrestins. *Annu. Rev. Physiol.* **69**, 511–534 (2007).

77.	Baltoumas, F. A., Theodoropoulou, M. C. & Hamodrakas, S. J. Interactions of the α-subunits of heterotrimeric G-proteins with GPCRs, effectors and RGS proteins: A critical review and analysis of interacting surfaces, conformational shifts, structural diversity and electrostatic potentials. *J. Struct. Biol.* **182**, 209–218 (2013).

78.	Oldham, W. M. & Hamm, H. E. Heterotrimeric G protein activation by G-protein-coupled receptors. *Nat. Rev. Mol. Cell Biol.* **9**, 60–71 (2008).

79.	Zhang, D., Zhao, Q. & Wu, B. Structural Studies of G Protein-Coupled Receptors. *Mol. Cells* **38**, 836–42 (2015).

80.	Park, J. H., Scheerer, P., Hofmann, K. P., Choe, H.-W. & Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **454**, 183–187 (2008).

81.	Scheerer, P. *et al.* Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **455**, 497–502 (2008).

82.	Venkatakrishnan, A. *et al.* Structured and disordered facets of the GPCR fold. *Curr. Opin. Struct. Biol.* **27**, 129–137 (2014).

83. Milligan, G. Constitutive Activity and Inverse Agonists of G Protein-Coupled Receptors: a Current Perspective. *Mol. Pharmacol.* **64**, 1271–1276 (2003).

84. Nutt, D. *et al.* Inverse agonists – What do they mean for psychiatry? *Eur. Neuropsychopharmacol.* **27**, 87–90 (2017).

85. Kenakin, T. Principles: receptor theory in pharmacology. *Trends Pharmacol. Sci.* **25**, 186–92 (2004).

86. Spiegel, A. Mutations in G proteins and G protein-coupled receptors in human endocrine diseases. in *Insights into Receptor Function and New Drug Development Targets* 139–150 (Springer Berlin Heidelberg, 2006).

87. de Ligt, R. A., Kourounakis, A. P. & IJzerman, A. P. Inverse agonism at G protein-coupled receptors: (patho)physiological relevance and implications for drug discovery. *Br. J. Pharmacol.* **130**, 1–12 (2000).

88. Park, P. S. H., Lodowski, D. T. & Palczewski, K. Activation of G protein-coupled receptors: Beyond two-state models and tertiary conformational changes. *Ann. Rev. Pharmacol. Toxicol.* **48**, 107–141 (2008).

89. GPCRs, Desirable Therapeutic Targets in Oncology. Available at: https://www.laboratoryequipment.com/article/2017/01/gpcrs-desirable-therapeutic-targets-oncology. (Accessed: 13th September 2018)

90. Lynch, J. R. & Wang, J. Y. G Protein-Coupled Receptor Signaling in Stem Cells and Cancer. *Int. J. Mol. Sci.* **17**,707 (2016).

91. O'Hayre, M., Degese, M. S. & Gutkind, J. S. Novel insights into G protein and G protein-coupled receptor signaling in cancer. *Curr. Opin. Cell Biol.* **27**, 126–135 (2014).

92. Sever, R. & Brugge, J. S. Signal transduction in cancer. *Cold Spring Harb. Perspect. Med.* **5**, a006098 (2015).

93. Lagerström, M. C. & Schiöth, H. B. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat. Rev. Drug Discov.* **7**, 339–357 (2008).

94. Greasley, P. J. & Clapham, J. C. Inverse agonism or neutral antagonism at G-protein coupled receptors: A medicinal chemistry challenge worth pursuing? *Eur. J. Pharmacol.* **553**, 1–9 (2006).

95. Wold, E. A. & Zhou, J. GPCR Allosteric Modulators: Mechanistic Advantages and Therapeutic Applications. *Curr. Top. Med. Chem.* **18**, 2002–2006 (2019).

96. Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739–45 (2000).

97. Okada, T. *et al.* X-Ray Diffraction Analysis of Three-Dimensional Crystals of Bovine Rhodopsin Obtained from Mixed Micelles. *J. Struct. Biol.* **130**, 73–80 (2000).

98. Stroud, R. M. New tools in membrane protein determination. *F1000 Biol. Rep.* **3**, 8 (2011).

99. Ujwal, R. & Bowie, J. U. Crystallizing membrane proteins using lipidic bicelles. *Methods* **55**, 337–341 (2011).

100. Denisov, I. G. & Sligar, S. G. Nanodiscs for structural and functional studies of membrane proteins. *Nat. Struct. Mol. Biol.* **23**, 481–486 (2016).

101. Xiang, J. *et al.* Successful Strategies to Determine High-Resolution Structures of GPCRs. *Trends Pharmacol. Sci.* **37**, 1055–1069 (2016).

102. Rawlings, A. E. Membrane proteins: always an insoluble problem? *Biochem. Soc. Trans.* **44**, 790–5 (2016).

103. Cherezov, V. *et al.* High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318**, 1258–65 (2007).

104. Rasmussen, S. G. F. *et al.* Crystal structure of the human β2 adrenergic G-protein-coupled receptor. *Nature* **450**, 383–387 (2007).

105. Isberg, V. *et al.* GPCRdb: an information system for G protein-coupled receptors. *Nucleic Acids Res.* **44**, D356–D364 (2016).

106. Thompson, A. A. *et al.* Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic. *Nature* **485**, 395–399 (2012).

107.   Ghosh, E., Kumari, P., Jaiman, D. & Shukla, A. K. Methodological advances: The unsung heroes of the GPCR structural revolution. *Nat. Rev. Mol. Cell Biol.* **16**, 69–81 (2015).

108.   Cavasotto, C. N. & Palomba, D. Expanding the horizons of G protein-coupled receptor structure-based ligand discovery and optimization using homology models. *Chem. Commun.* **51**, 13576–13594 (2015).

109.   Hillisch, A., Pineda, L. F. & Hilgenfeld, R. Utility of homology models in the drug discovery process. *Drug Discov. Today* **9**, 659–669 (2004).

110.   Bissantz, C., Bernard, P., Hibert, M. & Rognan, D. Protein-based virtual screening of chemical databases. II. Are homology models of g-protein coupled receptors suitable targets? *Proteins Struct. Funct. Bioinforma.* **50**, 5–25 (2002).

111.   Becker, O. M., Shacham, S., Marantz, Y. & Noiman, S. Modeling the 3D structure of GPCRs: advances and application to drug discovery. *Curr. Opin. Drug Discov. Devel.* **6**, 353–61 (2003).

112.   Sabio, M., Jones, K. & Topiol, S. Use of the X-ray structure of the β2-adrenergic receptor for drug discovery. Part 2: Identification of active compounds. *Bioorg. Med. Chem. Lett.* **18**, 5391–5395 (2008).

113.   Katritch, V. *et al.* Structure-based discovery of novel chemotypes for adenosine A2A receptor antagonists. *J. Med. Chem.* **53**, 1799–1809 (2010).

114.   Carlsson, J. *et al.* Structure-based discovery of A2A adenosine receptor ligands. *J. Med. Chem.* **53**, 3748–3755 (2010).

115.   Langmead, C. J. *et al.* Identification of Novel Adenosine A $_{2A}$ Receptor Antagonists by Virtual Screening. *J. Med. Chem.* **55**, 1904–1909 (2012).

116.   Jazayeri, A., Andrews, S. P. & Marshall, F. H. Structurally enabled discovery of adenosine a2a receptor antagonists. *Chem. Rev.* **117**, 21–37 (2017).

117.   Andrews, S. P., Brown, G. A. & Christopher, J. A. Structure-based and fragment-based GPCR drug discovery. *ChemMedChem* **9**, 256–275 (2014).

118. Ngo, T. *et al.* Identifying ligands at orphan GPCRs: current status using structure-based approaches. *Br. J. Pharmacol.* **173**, 2934–2951 (2016).

119. Negri, A. *et al.* Discovery of a novel selective kappa-opioid receptor agonist using crystal structure-based virtual screening. *J. Chem. Inf. Model.* **53**, 521–526 (2013).

120. Katritch, V., Kufareva, I. & Abagyan, R. Structure based prediction of subtype-selectivity for adenosine receptor antagonists. *Neuropharmacol.* **60**, 108–115 (2011).

121. Kolb, P. *et al.* Limits of Ligand Selectivity from Docking to Models: In Silico Screening for A1 Adenosine Receptor Antagonists. *PLoS One.* **7**, (2012).

122. Kruse, A. C. *et al.* Muscarinic receptors as model targets and antitargets for structure-based ligand discovery. *Mol. Pharmacol.* **84**, 528–540 (2013).

123. Rodríguez, D., Brea, J., Loza, M. I. & Carlsson, J. Structure-based discovery of selective serotonin 5-HT1B receptor Ligands. *Structure* **22**, 1140–1151 (2014).

124. Weiss, D. R. *et al.* Conformation guides molecular efficacy in docking screens of activated β-2 adrenergic G protein coupled receptor. *ACS Chem. Biol.* **8**, 1018–1026 (2013).

125. Mohr, K., Schmitz, J., Schrage, R., Tränkle, C. & Holzgrabe, U. Molecular alliance - From orthosteric and allosteric ligands to dualsteric/bitopic agonists at G protein coupled receptors. *Angew. Chem. Int. Ed.* **52**, 508–516 (2013).

126. Brogi, S., Tafi, A., Désaubry, L. & Nebigil, C. G. Discovery of GPCR ligands for probing signal transduction pathways. *Front. Pharmacol.* **5**, (2014).

127. Gentry, P. R., Sexton, P. M. & Christopoulos, A. Novel Allosteric Modulators of G Protein-coupled Receptors. *J. Biol. Chem.* **290**, 19478–88 (2015).

128. Allegretti, M., Cesta, M. C. & Locati, M. Allosteric modulation of chemoattractant receptors. *Front. Immunol.* **7**, (2016).

129. Zarbock, A., Allegretti, M. & Ley, K. Therapeutic inhibition of CXCR2 by Reparixin attenuates acute lung injury in mice. *Br. J. Pharmacol.* **155**, 357–364 (2008).

130. Fronik, P., Gaiser, B. I. & Sejer Pedersen, D. Bitopic Ligands and Metastable Binding Sites: Opportunities for G Protein-Coupled Receptor (GPCR) Medicinal Chemistry. *J. Med. Chem.* **60**, 4126–4134 (2017).

131. Bock, A. *et al.* The allosteric vestibule of a seven transmembrane helical receptor controls G-protein coupling. *Nat. Commun.* **3**,1-11 (2012).

132. Silvano, E. *et al.* The tetrahydroisoquinoline derivative SB269,652 is an allosteric antagonist at dopamine D3 and D2 receptors. *Mol. Pharmacol.* **78**, 925–934 (2010).

133. Lane, J. R. *et al.* A new mechanism of allostery in a G protein-coupled receptor dimer. *Nat. Chem. Biol.* **10**, 745–752 (2014).

134. Garnock-Jones, K. P. Cariprazine: A Review in Schizophrenia. *CNS Drugs* **31**, 513–525 (2017).

135. Frankel, J. S. & Schwartz, T. L. Brexpiprazole and cariprazine: distinguishing two new atypical antipsychotics from the original dopamine stabilizer aripiprazole. *Ther. Adv. Psychopharmacol.* **7**, 29–41 (2017).

CHAPTER 2: Methodology

2.0 Computer aided drug design

The process of drug discovery is long, expensive and full of risks. It is estimated that the introduction of a new drug into the market requires 10-15 years and around 500-800 M US $ investment. In addition, taking into account the huge competition among pharmaceutical companies, there is a great interest in finding novel technologies that can reduce the time required as well as the cost. In this direction, computer-aided drug design got momentum in the 80s and it is now widely used at the present in the drug discovery process at the pharmaceutical industry.

Computer-aided drug design (CADD) or molecular modelling use computers to design or discover new molecules based on the concept that biochemical processes are receptor mediated and subsequently, on the ligand-receptor interaction concept and the necessary stereochemical complementarity between them. The process follows a virtuous circle that requires three steps to finalize improving the initial hypothesis: analyse the available experimental information, produce a hypothesis of the ligand-receptor interaction and identify the next generation of molecules that will be synthesised based on the produced hypothesis and compare with the previous results to eventually improve the starting hypothesis. Depending on the experimental information available, computer-aided drug design can have categorized into two main classes: structure-based drug design (SBDD) and ligand-based drug design (LBDD). The former can be carried out when there is enough structural information on the target and the latter, is used when we only have knowledge of a collection of active compounds.

In order to produce a hypothesis of the ligand-receptor interaction, this technique uses diverse molecular parameters such as the surface of the molecules, electrostatic force, hydrophobic and hydrogen bond interactions. These parameters are important because they play an important role in the evaluation and prediction of the interactions between the receptor and the ligand. CADD studies can be used to provide information about drug-target complexes, analyse the structure of the target for possible binding pocket, generation of new molecules, test their affinity, docking the molecules to the target, ranking of them based on the best affinity and optimizing them to enhance their binding affinity.

2.1 Ligand based drug design (LBDD)

Ligand-based drug design (LBDD) needs to be applied when there is no information about the 3D structure of the target. In this technique, a hypothesis of the ligand-receptor interaction is produced by comparison of the features of a collection of active and non-active compounds. More specifically, this method is based on analysing the 2D or 3D structures of the set of reference structures with known pharmacological activity to the target. So, while the structure of the target protein is not available,

structure of the ligand of the target is taking into account as the starting point of the drug discovery which is based on this fact that molecules with the same structure will produce the same biological response[1,2].

One way to go is to superimpose the structures of the ligands available and identify commonalities. This can be used to find structures with similar features in data bases. Thus, similarity searching can be done among the compounds that have similar characteristics as the same as reference structures such as physicochemical properties and 2D or 3D structure. This approach depends highly to the input molecule as well as their structure[3].

Another procedure is ligand-based pharmacophore modelling. In this procedure, diverse molecular features of the ligands are evaluated. Features to compute are hydrogen-bond acceptors or donors, negative or positive charged functional groups, ring centers and volume spheres[4] . At the final stage, information from active and inactive compounds apply to build a model trough structure-activity relationships (QSAR) making use of the physio-chemical properties of the compounds[5]. Comparison of the structures of the active compounds permits to define a hypothesis of the features that make the ligands complementary to the receptor and consequently, gives directions on how the compounds could have higher affinity. The process is shown schematically in Figure 2.1.
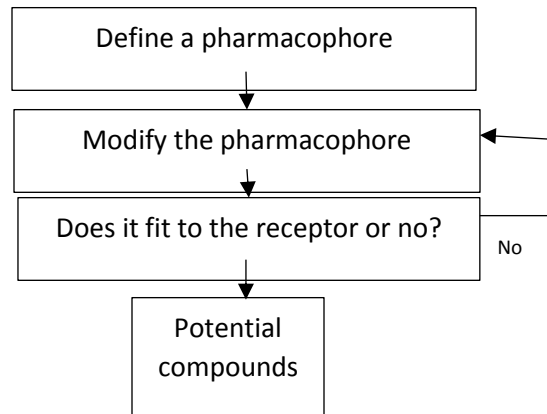
Figure 2. 1 *Flux diagram of Ligand-based drug design (Pharmacophore modelling)*

2.2 Structure-based drug discovery

Structure-based drug design is based on the study of the stereochemical complementarity between a ligand and its receptor. The accuracy of this procedure depends on the available information on the 3D structure of the target protein. Ideally, the procedure requires knowledge of the 3D structure, determined by experimental methods, although in its absence it can be constructed a 3D model by homology modelling. Fortunately, the number of available 3D structures of proteins, determined either by X-ray crystallography or nuclear magnetic resonance (NMR) is steadily increasing, so the application of

structure-based drug design is nowadays common. Accordingly, the procedure involves prepare the 3D structure of the target or construct it by homology modelling if the structure of the target protein is not available by experimental methods, and perform molecular docking studies of a set of ligands onto the receptor. The diverse steps to follow include, building structure of the target protein and identify the binding pocket of the receptor which should have the interest interaction with the ligand, perform docking process and select the best docking based on the scoring function and interesting interaction, developing pharmacophore mapping of receptor-ligand complex, virtual screening of compounds or hits, filtering and selecting novel hits and experimental evaluation of them to compare with the computer-simulated results (Figure 1.2).

```
┌─────────────────────────────────┐
│   Define the Target structure   │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│     Docking of the ligand       │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│  Select the best protein-ligand │
│            complex              │
└─────────────────────────────────┘
                 │
                 ▼
         ┌────────────────┐
         │       MD       │
         └────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│      Pharmacophore modeling     │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│        Virtual Screening        │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│    Select Candidate compounds   │
└─────────────────────────────────┘
```
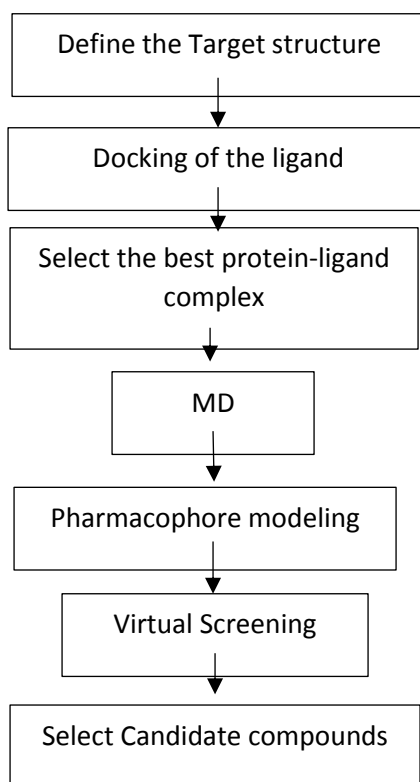
Figure 2. 2 *Flux diagram of Structure-based drug design*

The goal of the molecular docking step is to find compounds that better fit the streochemical characteristics of the binding pocket. Docking studies are performed at different stages in the drug discovery process such as hit identification, lead optimization, and library for target-ligand complex. The docking process is a fast and cheap computational technique, although its success depends on the quality of the structure of the target protein[6–8]. There are some challenges associated with structure-based drug

design, such as lack of 3D crystallographic structure of target protein and the inclusion or not of receptor-ligand flexibility[9].

As mentioned above, sometimes the availability of high-resolution crystal structure of protein target is limited. Special case is the family of G-protein coupled receptors (GPCRs) with a few experimental structures available, whereas they are highly extended in the human genome[10,11]. When there is no experimental structure available of the target protein of interest, a 3D structure of the target protein at atomic resolution, it can be constructed by homology modelling. While constructed protein models can provide insight into the biochemistry and the function of the biomolecules, they have limited accuracy and it is difficult to use them in structure-based drug discovery process. So that having an accurate model structure of the target protein plays an important role in drug discovery[10,12,13].

2.2.1 Homology modelling

Homology modelling is a method to construct a 3D model of a target protein at atomic-resolution by threading its sequence into the known structure of a protein with high sequence identity that is used as template. The method is based on the observation that proteins with high sequence identity share common 3D structures.

The procedure consists of four main steps. The first step is to select a set of proteins that share a high sequence identity with the target. For this purpose, the sequence of the target protein will be compared with the sequences a diverse proteins whose 3D structure is known. Second, identify the protein that will be used as template, based on the multiple sequence alignment. This is based on the alignment of conserved residues (motifs) of the diverse sequences. This step plays an important role in homology modelling. Sequences that differ in a large number of residues or there are frequent insertions and deletions should not be considered. In the case of GPCRs, where sequence identity is not very high this selection can be guided by the phylogenetic tree of the family. In the third step the sequence of the target protein is threated into the structure of the template. Finally, the constructed model is optimised and assessed. Normally, several models with small differences can be constructed. Their subsequent assessment permits to select the one with the best characteristics. Differences between models normally occur in the regions of the protein that have less sequence identity with the template, even though the rest of the protein share high sequence identity with the template. The percentage of the sequence identity between target and template plays an important role in the quality of the constructed model. If sequence identity is between 60-100% the result can be compared to structure solve by NMR with medium resolution. Lower sequence identity requires to rely in conservation motifs and are less reliable[14].

As mentioned above, sequence alignment is an important step in the process. In this work we have used the multiple sequence alignment algorithm CLUSTALW[15,16]. This is an open-source code that uses the UPGMA/Neighbor-joining method to generate a distance matrix between sequences. For comparison purposes we have also used sequence alignment module of MOE (the basic program used in this thesis) that uses the BLOSUM40 scoring function to construct a distance matrix. When it is difficult to align regions of protein and target with the less sequence identity it is important to consider the alignment with less gaps in the alignment of both sequences.

Once the alignment is finalized, the coordinates of the template can be transferred to the residues of the target sequence to produce a rough or crude model of the target protein. Construction of the backbone of the target is the first step of the process. Sometimes in the alignment there is a gap by deletion or insertion of a few residues. Normally, this occurs in loop regions. In this case, is necessary to carry out an extra modelling effort. Loop modelling is based on database searching. In database searching loops can construct from protein structure and replacing them to the target sequence, while in the data base searching method, many random loops construct, and we should search for the reasonable one with the lower energy and Φ and Ψ angles in the allowable parts of the Ramachandran plot. Another important point during the process is side-chain modelling which plays an important role in the quality of the constructed model. It is important for evaluating interaction of the protein-ligand on the binding site and it can be built trough searching of different conformation of the side chain torsion angles and selecting the ones with the lowest interaction energy within the next atoms. All the interesting side-chain torsion angles can be deducted from the protein-crystal structure.

The homology modelling process within MOE has two steps: the first one is partial geometry whereas all coordinates copy to the target if conservation of residue identity can be observed. The second one is Boltzmann-weighted randomized sampling which is data- collection and model building stages. During the data collection, backbone fragments can collect from the high-resolution crystal structure and high chain conformation of non -identical residues can assemble from the rotamer library. Models can be constructed based on the backbone and loop placement and they can score by a contact energy function[17].

The final step of constructing a model in homology modelling is model optimisation. Which is the energy minimization process of constructed models by positioning the atoms in a way that produces the lowest energy potential for the overall conformation of the model. This step can be summarized in the scoring function table and the best model can be selected based on that.

After construction, model validation is done by checking Φ and Ψ angles, bond, length, chirality and interatomic contacts. Constructing a good model depends on the template selection, applied algorithm and model validation. Interatomic contacts can be eliminated by energy minimization. Model validation can be evaluated using the Ramachandran plot which is a way to visualize backbone conformations of every amino acids in the structure through a representation of their Φ and Ψ dihedral angles. And it shows which values or conformations of these angles are possible for each residue of the protein. If the angles are in the allowable region of the Ramachandran plot, it can assume that a good model has been constructed.

Inaccuracies during the modelling effect the quality of the final model. These errors can be produced in sequence alignment step and it can be distinguished by comparison of the final constructed model with the corresponding crystal structure trough root mean square deviation (*rmsd*). The produced models should exhibit *rmsd*< 2Å in comparison to the crystal structure. These kinds of errors can influence the binding site by producing wrong side-chain orientation and consequently week receptor-ligand interaction. These kind of models cannot be used for structure-based drug design. So, it is really important to apply methods for flexibility of the receptor during modelling for having precise side-chain orientation and enhancing the quality of the constructed model[18,19].

Several successful examples of drug discovery using homology modelling have been reported in the literature. For example, a 3D structure of the human histone deacetylases(HDAC10) produced by homology modelling was used to design a novel inhibitor for the treatment of cancer[20]. Similarly, the mycobacterium tuberculosis Mary(RV2156C) integral membrane protein was constructed by homology modelling for the design antibiotics for the treatment of tuberculosis[21]. Inhibitors of the 5-Alpha-Reductase 2, target for the treatment of several diseases including prostate cancer, benign prostatic hyperplasia, male pattern baldness, acne, and hirsutism were discovered using a protein structure constructed by homology modelling[22]. Other example is the discovery of α-glucosidase inhibitors for the treatment of diabetes[23]. Homology modelling has been also used to study the role of certain residues in the experimental mutagenesis studies[24–26].

A special attention should be paid to GPCRs. Despite of the low sequence identity between the members of class A GPCR superfamily (~20%), these proteins share several structural features such as having 7 transmembrane helixes and conserved motifs in the transmembrane region[27]. However, still, there are some ambiguous questions about homology modelling such as selecting the best template for constructing the model[28], the presence of loops and insertion of that to the model without using template data[29], using single or multiple templates for having an accurate model in the absence of the low sequence

identity between template and target[30] and identification of the best model among various models based on the different analysis such as template modelling score[29], RMSD[31] and energy score[32]. Chapter 3 of this thesis deals with the topic of template selection for the homology modelling of these proteins. Several effects like the effect of the environment are evaluated trough molecular dynamics simulation in the refinement process of the structures[33].

2.2.2 Molecular docking

Molecular docking techniques were developed during the early 80s and are key methods in structure-based drug design[54,58]. Although molecular docking was initially developed for protein-ligand studies, it was later extended to study protein-protein, and protein-nucleotide complexes. In this thesis, we used protein-ligand molecular docking for different proteins.

The goal of protein-ligand docking studies is to simulate the interaction process between a ligand and a protein at the atomic level, to yield the best complementary pose and evaluate its relevance in comparison with other ligand-receptor complexes binding to the same target[36]. More specifically, molecular docking aims to optimize conformation of both target and ligand, also optimization of the orientation between them for achieving the minimized free energy for the whole system[34,35]. The process involves the ligand to explore different manners to adjust to the binding site of the protein. For this purpose, the orientation and position of ligand to the receptor should define correctly. During docking searching in high-dimensional space is searched thoroughly using a scoring function that permits the best fitting of the ligand into the binding pocket, i.e. its binding pose and also, a *binding affinity* of the ligand is roughly provided through the values of the scoring function.

The docking process can be carried out rigid or flexible. The former process is based on the lock and key theory of molecular recognition and provides the right orientation of a ligand as a key that fits onto the receptor that acts as a lock[35,37]. The latter, is based on the "induced fit" theory that suggests that the active site of the protein is reshaping constantly by interaction with the ligand as the ligand interactions with the receptor. In this theory both protein and ligand are flexible.

The actual process of docking can be summarized in two steps: first, compute diverse conformations of the ligand and second, docking the diverse conformations and rank them through scoring function. The basic tools for this aim are search algorithm and scoring function to generate and evaluate conformations of the ligand. The search algorithm can identify different poses of the ligand in the binding mode and scoring function rank the highest-affinity of the produced conformation by predicting the strength of non-contact interaction between two molecules after ending up the process of

docking[35,38] . Then the possible conformations of ligand can rank and prepare for the selection based on the best scores and interesting binding poses[34].

The success of protein-ligand docking depends on the selected target, quality of protein-ligand interaction and software suit dependent[39]. The software used in this thesis for molecular docking includes the genetic-optimization for ligand docking (GOLD)[40] and the molecular operating environment (MOE)[41]. They are robust docking programs and are reliable among other software programs[42].

2.2.2.1 Genetic-optimization for ligand docking (GOLD)

Gold is an automatic-docking program that uses a genetic algorithm strategy for ligand placement. It allows full flexibility of ligand and partial flexibility of protein, including side-chain and backbone part of the protein, during molecular dynamic[43]. The docking accuracy of the gold was validated on different complexes[44,45]. Gold can predict the poses with 90% accuracy[46]. Like all the other programs, gold can be tested in three parts: initialization of the protein and ligand, genetic algorithm mechanism and fitness function.

The first step of docking is the preparation of both ligand and receptor and define the docking sphere. For the preparation of the receptor, at first it should have protonated by protonation part of the program and also, its tautomer should be defined. And if there are any ligand or metal ions in the protein they should remove. Also some cases should evaluate such as ligand hydrogen atoms, ionization, and tautomer state and if it is necessary hydrogen will be added and incorrect ionization and tautomeric state will correct.

Docking spheres refers to the placing in the binding site of the protein with a determined size and centre. The centre can be defined as either point or atoms. The centre coordinates are x, y and z-direction and the radius of the binding site have defined by gold docking calculation.

Next step of the docking is genetic algorithm implementation to sample for both conformation spaces and ligand binding modes[45]. The genetic algorithm can modify and optimize parameters such as dihedral angels of routable bonds ligand and geometric (flipping ring corners), dihedral angels of the protein-OH group and NH3 groups and mapping of fitting points of the ligand in the binding site.

After generation of protein-ligand binding cleft, the mechanism of the placing the ligand will be done using at least square fitting procedure and is based on fitting points, which means the program adds a fitting point to the hydrogen binding group at both protein and ligand and map the acceptor point in the protein and vice versa. Also, the generation of hydrophobic fitting points will be done in the protein cavity where ligand CH group is mapped.

And at final Gold fitness score will evaluate which predict ligand binding poses by attention to the different factors such as hydrogen binding energy, van der Waals energy, metal interaction, and ligand torsion strain. The main formula for this function summarize as:

$$\text{Gold fitness: } S_{hb\text{-}ext} + S_{dvw\text{-}ext} + S_{hb\text{-}int} + S_{vdw\text{-}int}$$

Where:

$S_{hb\text{-}ext:}$ is the protein-ligand hydrogen bond score

$S_{dvw\text{-}ext}$: is the protein-ligand van der Waals score

$S_{hb\text{-}int}$: is the contribution to the fitness due to intramolecular hydrogen bonds in ligand[40,47].

$S_{vdw\text{-}int:\ Is}$ the contribution due to the intramolecular strain in ligand.

In this study, the gold function was used for the main docking calculation by using methods such as Gold score, scoring function. Different parameters for these methods were chosen including flipping ring corners, planer, and pyramid nitrogen, allowing have routable H2O molecule, diverse pose during pacing in the pocket and rotating interacting hydrogen bonds[48].

2.2.2.2 Molecular Operating Environment (MOE)

MOE is the another software to support structure-based drug design steps such as molecular modelling, energy minimization, and molecular docking by using scientific vector language (SVL) as a scripting and application development language of MOE. In this thesis, MOE was used for the preparation of the molecular structure, energy minimization, molecular docking and post docking refinement, and visualization. Multiple methods are applied for molecular docking with MOE which easily is integrated for this software including conformational analysis and placement, first scoring, refinement, second scoring and pharmacophore constraint.

In the conformational analysis step, conformations can apply from a single ID conform or conformation database to design preferred torsion angle to the routable bonds. During placement, generation of the set of poses will be done from ligand conformations by using one of the different placement methods such as alpha triangle, alpha PMI, proxy triangular and triangular method. Rescoring of generated poses is based on the scoring function that insists on hydrophobic, ionic and hydrogen bonds contact. For having the correct docking framework, the assignment of all new scores will be done based on having a good score for the best pose. During refinement, positioning of conformations will refine by using explicit molecular mechanism for grid-based energy method and after that rescoring will be done for the results from the refinement process and it will appear trough the scoring scheme and in final if it is necessary for the user, preparation of pharmacophore point will be done from the final selected pose. Molecular system optimization can be used for preparing the structure to be docked for further

simulation. In this thesis, it was used for the identification of the missing atoms from different residues, energy minimization, and assignment of protonation states. Unfortunately, little or no hydrogen coordinates data can be seen in most of the macromolecular structures due to the low resolution or disorder of the crystal structure, while explicit hydrogen atoms are applied for further atoms molecular mechanism, docking, and electrostatic calculation. Moreover, the initial state of the hydrogen bond and ionization state of the treatable group produce dramatic effects on the results of docking.

Since some hydrogen atoms can be seen in PDB files, a protonation state should be done for amino acids, as well as centre ions and also solvents. MOE uses a protonate 3D method to do this step by the purpose of assign ionization state and position of the hydrogens in the macromolecular structure of input 3D coordinates (typically from a crystal structure). The main task in this thesis are defining:

- ✓ The rotamers of –SH-OH-CH$_3$ and NH$_3$ groups in Cys, Ser, Tyr, Thr, Met, and Lys.
- ✓ The tautomers of imidazole (His) and carboxyl acids (Arg, Glu).
- ✓ The protonation state of metal ligand atoms Cys, His, Asp, Glu, etc…
- ✓ The ionization state of metals
- ✓ The element identity in His, terminal amides (Asn, Glu) and sulphonamide.
- ✓ The orientation of each water molecule[41].

After the adding of the missing atoms and residues to the receptor and preparation of the ligand structure, they submit to the energy minimization calculation. Energy minimization is a necessary step to provide a local energy minimum because lower energy states are more stable and so they can consider as a native state of the system. The MOE force field MMFF94X was used for the minimization of the systems. Using a good force field is necessary to describe the quality of the ligand and the target individually as well as the bond. For this aim, having a good force field is required to predict conformational energy and molecular geometries to prevent modeling the wrong conformation of the receptor or ligand upon binding. Among all force fields, MMFF94X is one of the best to compute behaviour of the proteins, amino acids, and small molecule structures. Moreover, MMFF94X has special performance in using studies of binding molecules[49].

One of the objectives of this thesis is to study the behaviour of M3 muscarinic receptor as well as the Bombesin receptor trough molecular dynamics simulations, so different models from each study has been used to develop the molecular dynamics simulation. The main purpose of extending models with molecular dynamics simulation (MD) is modelling the molecular motions as well as the optimization of all particles of the system. MOE software shows precise results in comparison to the other programs such as

results in docking accuracy for different receptors and finding the best binding modes for further test studies such as virtual screening. Thus, the final refined model from molecular simulations was used MOE for pharmacophore-based virtual screening of novel compounds.

2.2.2.3 Glide-based ligand docking with energetics (GLIDE)

This program designed to search for possible positional, orientation, and conformational space that is close to the ligand by a screening of a large library with sufficient computational speed. For this purpose, it uses a series of hierarchical filters consisting of the following steps: (a) Grids or pre-processing step for demonstration of the receptor shape and its properties (b) Pre-screening for initial screening of set of ligand conformations (c) Minimization of the ligand in the receptor surface using OPLS-AA force field (d) selection of the best candidates or lowest energy poses (e) Final scoring based on the predicting binding affinities which order ligand in one table(Figure 2.3)[50]. Glide generates and docks many cores and rotamer conformations through the docking process. For ligand searching, those atoms use in diameter tests which is close to the ligand diameters.

Gide: "Funnel"

Ligand conformationStep1: Site point search

Step2a: Diameter test

Step2b: Subset test

Step2c: Greedy score

Step2d: Refinement

Step 3: Grid minimization

Step4: Final scoring (Glide score)

Top hits
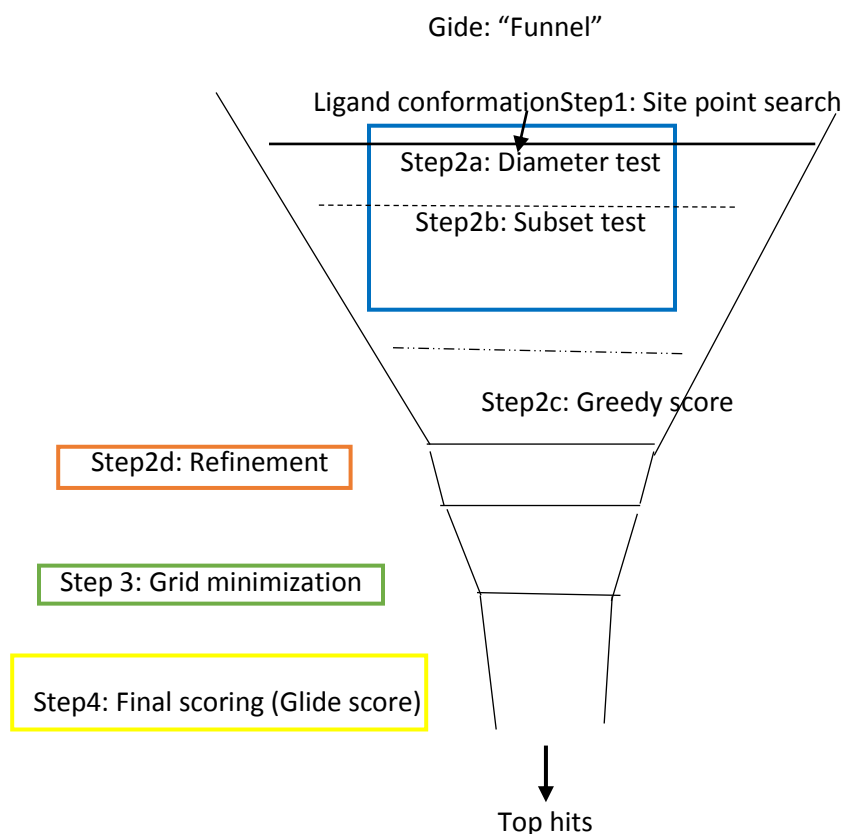
Figure 2. 3 The Glide docking Hierarchy

In the first step or fast site point search, grid of site points generate in the binding site, then pre-compute is done for a histogram of the distance between site points and receptor surface, subsequently,

a comparison is done between site-point receptor surface and histogram of ligand center-ligand surface and in final rejection is done among initial mismatched site point. The second scoring is rough scoring which consists of a diameter test, subset test, and greedy scoring and refinement process. In this step, steric clashes of atoms near the ligand diameter check trough diameter test for different orientation of the ligand diameter, then rotation of ligand diameters perform in subset test and scoring is done for atoms with the capability of making H-bond and ligand-metal interaction. Scoring of greedy scoring is for all atoms in the position of ±1Å in x, y, and z directions and rescoring of that is doing to reduce the number of poses for energy minimization. The third step is energy minimization which uses the OPLS-AA force field to calculate van der Waals and electrostatic grids. Energy minimization permits free movements of contacts. Moreover, another process is done in this step such as optimization of torsional angles through flexible docking and performing Monte Carlo moves to explore nearby torsional minima for a small number of low energy poses. In the final step, the best poses choose for each ligand based on a model energy score (E-model) which is a combination of Coulomb van der Waals energy, glide score, and strain energy. Glide score can calculate based on this formula:

$$Gscore: 0.065 * vdW + 0.130 * Coul + Lipo + Hbond + Metal + BuryP + RotB + Site$$

Whereas:

Vdw: van der Waals energy term

Coul: coulomb energy term

Lipo: lipophobic term taken from hydrophobic grid potential or favourable hydrophobic interaction

H-bond: Hydrogen bonding term

Metal: Metal binding term

Bury P: penalty for freezing rotatable bonds

Rot B: Penalty for freezing rotatable bonds.

Site: polar interaction in the active site

There are two types of scoring based on Glide score including Glide Score SP and HTVS ("softer" function,) which use in database screening application and Glide Score XP ("harder" function") which use in lead optimization and docking process. The latter one that used in this thesis consist of this formula.

$$\text{XP Glide score: } E_{coul} + E_{vdW} + E_{bind} + E_{penalty}$$

The hierarchical search of glide makes it an accurate program for predicting the binding mode of the ligand. Also, it is a cost-effective program to complete a systematic search. Because its algorithm permits optimization of the ligand at the same time that search of best conformation and location of the ligand is done. So, a large library can screen with this cost-effective computational approach. We used GLIDE software to dock the different conformations of antagonists to the orthosteric site of the crude models after the refinement process. In the chapter 5 of this thesis, Glide has been used to dock different conformations of the antagonist ligand to the orthosteric binding site of the crud model using XP scoring function.

2.2.3 Molecular Dynamics simulations

Molecular dynamics is a method to calculate the time-dependent behaviours of molecules by evaluating the dynamics of the system and providing information about fluctuations and conformation changes of proteins[51-76]. Molecular dynamics simulation can apply in different studies such as study stability and folding of proteins, rational design of biologically active molecules e.g., drugs, enzyme reactions, conformational changes in small or large scale, determine constructed 3D structures of homology modelling, X-ray diffraction, and NMR. With molecular dynamics simulations, it is possible to study both thermodynamic and/or time-dependent properties[54]. In this thesis, molecular dynamics simulations (MD) were used in chapters 3 and 5 for the refinement of the crude models resulting from homology modelling.

In this work molecules are represented by a classical potential, so that chemical bonds are fixed although flexible and the atoms of a molecule are subjected to weak interactions that determine their conformations. The set of parameters that determine the diverse interactions between atoms are collected in a force field, so that every atom feels a potential energy as a result of the diverse atoms surrounding it. Accordingly, it is possible to determine the force on every atom and using Newton's equations of motion, compute their accelerations, speeds and positions as a function of time. The integration of the equation of the motion provides a trajectory for describing position and velocities and acceleration of each particle along the sampling time. This is a deterministic method that describes

position and velocities of particles and also with this method, it is possible to predict the state of the systems at any time in the past or future [52,53,55,56].

The starting point of molecular dynamics simulation is to consider the atom coordinates from the starting structure and compute the interactions between atoms. Then, it is possible to calculate and evaluate velocity for each atom by integration of Newton's equations of motion applying a numerical method like for example the Verlet algorithm[57].

According to Newton's equation of motion, the acceleration of an atom can be computed as:

$$\vec{F}_i = m_i . \vec{a}_i$$

Where F_i is the force extracted on particle i, m_i is mass of the particle i and a_i is the acceleration of the particle i. The force on an atom can also be expressed as the gradient of the potential energy:

$$\vec{F}_i = -\vec{\nabla}V_i$$

Where V is the potential energy and it is function of the atomic positions of all atoms in the system. Since the velocity is the derivative of position and acceleration is derivative of velocity, so the equation of the motion can be written as:

$$\frac{d\vec{v}_i}{dt} = \frac{\vec{F}_i}{m_i}$$

Solving the equations of motion, it should be considered that it consists of a system of ordinary of equations. If N is the number of atoms, there are 3N position of coordinates and also 3N of velocity coordinates. The system of equations cannot be solved analytically, so we use a numerical method that is straight forward[81]. This requires to use an integration step Δt:

$$\vec{v}_i(t + \Delta t) = \vec{v}_i(t) + \frac{\vec{F}_i}{m_i}\Delta t$$

$$\vec{r}_i(t + \Delta t) = 2\vec{r}_i(t) - \vec{r}_i(t - \Delta t) + \frac{\vec{F}_i}{m_i}\Delta t^2$$

Time step needs to be chosen in such a way that is shorter than the time required for an atom to conduct half a cycle for the highest frequency vibration of the system. This corresponds to the vibration of bonds that involve hydrogen atoms and it is normally taken as 1 fs. In order to use larger time steps, these bonds are constrained so that time step can be doubled. Examples of algorithms to constrain specific vibrations include settle[58], shake[59], M-Shake [60] and Lincs[61] algorithms. Taking into account the computers available and the time step that can be used, typical trajectories performed nowadays are of the order of a microsecond[53,55,56,62,63].

There are different factors that influence in the outcomes of the MD simulation such as accuracy of the starting configuration of the whole system, statistical ensemble of the system (e.g. temperature, pressure), type of solvation, chosen degrees of freedom, bonded and non-bonded interactions, time-step integration and parameters of selected force fields or describing the atoms in the system[53,55,56].

2.2.3.1 Force field

Force fields are used to describe the intramolecular interaction potential energy of the system[56]. A force field is a mathematical function describing the potential energy of a system which depends on its corresponding particle coordinates. Thus, it includes an analytical expression of the intramolecular interaction together with a set of parameters for each of the atom types of the system. These parameters can be obtained from *ab initio* calculations or from experimental results such as x-ray electron diffraction, NMR, Raman and neutron spectroscopy, etc[64]. The most commonly used force fields for molecular dynamics include Amber[65], CHARMM[66–68], Gromos[69], and OPLSAA[70]. For the calculations reported in this work we used OPLSAA force field implemented in the GROMACS package[98].

The basic function of the force field includes bonded and non-bonded interactions, including electrostatic and van der Waals. Accordingly, the general form of the total energy can be written as:

$E_{total} = E_{bonded} + E_{nonbonded}$

$E_{nonbonded} = E_{electrostatic} + E_{van der Waals}$

Normally, bonds and angles are modelled with a harmonic potential or quadratic energy function to prevent bond breaking, while dihedral angles exhibit several minima and cannot compute as harmonic oscillators. Dihedral functions are normally modelled with a Fourier series. Non-covalent interaction are computed between every pair of atoms. van der Waals interactions are modelled with a Lennard-Jones potential, whereas the functional form of the electrostatic term corresponds to the Coulomb potential.

The potential energy of the system is computed as a summation of the bonded terms such as bond elongation, angle, and dihedral together with the non-bonded interactions. The functional form of the OPLSAA force field is the following[71]:

$$E(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N) = E_{bond} + E_{angles} + E_{dih} + E_{non-bonded}$$

For any i-j pairs that are connected by bond, the bond stretching energies can calculate as:

$$E_{bond} = \sum_{bond} k_r (r - r_{eq})^2$$

Where k_r is the stretching force constant and distance between atoms and its equilibrium bond length show as r and r_{eq}.

The angle term between three bonded atoms can calculate as:

$$E_{angle} = \sum_{angle} k_\theta (\theta - \theta_{eq})^2$$

Whereas binding force constant represents KƟ and angle between atoms and its equilibrated value shows as Ɵ – Ɵeq respectively.

For any group of four bonded atoms, the dihedral angles can calculate as [72]:

$$E_{tor} = \sum_i \left[\frac{1}{2} V_{1i}(1 + cos\varphi_i) + \frac{1}{2} V_{2i}(1 + cos2\varphi_i) + \cdots \right]$$

Where V is the dihedral constant that effect on the barrier light and ɸ shows the dihedral angle and the equilibrium value according to the biochemical convention((trans ɸ=180º, cis ɸ=0º and gauche ɸ=60º/300º)

The first non-bonded term represented by a 6-12 Lennard-jones potential which is a simple mathematical function[73] which accounts for two distinct attractive and repulsive forces which are applied for neutral atoms and molecules as:

$$E_{vdW} = \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}} \right)^{12} - \left(\frac{C_{ij}}{r_{ij}} \right)^{6}$$

where Cij stands for attractive forces at long range (van der Waals or dispersion); Aij represent the repulsive at short range producing for overlapping of electronic orbitals anf rij is the distance between atoms I and j. Finally, the Coulomb potential which describes the electrostatic interaction express as:

$$E_{elec} = \sum_{i,j} \frac{q_i q_j}{4\pi\varepsilon_o r_{ij}}$$

where qi and qj are the charges of atoms i and j and rij is the distance between these atoms and ε_o is the vacuum permittivity.

2.2.3.2 Statistical mechanics and molecular dynamics simulations

Molecular Dynamics (MD) is a very powerful sampling tools that permits the study of both kinetic or time-dependent phenomena and also thermodynamic properties of a system. The kinetic study can describe the mechanism of the action of chemical processes, while thermodynamic explains the driving force of chemical reactions. However, MD simulations provide microscopic level information such as atomic positions and velocities that need to be converted into macroscopic observables such as pressure, energy, and heat. For this purpose is it necessary to use the concepts derived in Statistical Mechanics. In order to convert microscopic information into macroscopic one, time-independent statistical averages are required such as the thermodynamic state. The thermodynamic state can define a set of parameters such as temperature (T), pressure (P), and the number of particles (N), whereas the microscopic state of the

system is described in a multidimensional spaces called the phase space, where the individual atomic positions (q) and momenta (p) are specified. This 6xN dimensional space for systems, being N the number of particles. An ensemble is a collection points describing conditions of a specific thermodynamic state in the phase space. There are three main ensembles in the molecular dynamics simulations that contain their specific characteristics which are the microcanonical ensemble(NVE), the canonical ensemble (NVT) and the Isobaric-Isothermal ensemble (NPT)[74].

The microcanonical ensemble or NVE is a thermodynamic state that has a fixed number of atoms (N), volume (v), and energy and correlated to the system that is completely isolated from its surrounding environment and there isn't any energy transfer between this system and its environment. So, the volume of the system remains fixed. Because its total energy is constant, it is not possible to define the temperature of this ensemble because the temperature can only describe in the system that has interaction with its surroundings. So the Microcanonical ensemble is described in terms of its number of atoms, volume, and energy. This ensemble is using to show possible states of a system with specific total energy. During the system equilibrium state of this system, enthalpy is maximized while the total energy is conserved.

Canonical ensemble or NVT is a collection of a system in which their thermodynamic states can define by fixed number of atoms (N), volume (V), and temperature (T). However, in this ensemble, the energy can transfer across the system and its surrounding but matters can't transfer. Because this system is in thermal contact, it is possible to transfer heat (q) between this system and its surroundings until it reaches its thermal equilibration. So, the temperature of this system, unlike the Microcanonical ensemble, can be defined as constant. This system describes states of thermal equilibrium of a mechanical system with a heat bathed fix temperature. Also, in this ensemble, the energy isn't constant. It can change based on temperature[75]. The importance of this system is describing Helmholtz free energy of a system which in this system it is possible to work maximum at a constant volume (v) and temperature (T).

Isothermal-Isobaric ensemble (NPT) has the characteristic of fix number of atoms (N), fix pressure (P), and fix temperature (T). In this system transformation of the energy is possible from the boundary, but it is impossible for matters. Also, the volume of the system can be changed because the external pressure of the system should match the external surrounding environment. The NPT ensemble has heat bath temperature (T) as same as the canonical ensemble, in which the heat bath of the surrounding is more than the internal environment. So, an increase in the system's heat energy does not have any effect on the surrounding environment. This ensemble has important role in chemistry because most of the

chemical reactions can occur under constant pressure. Also, the importance of that is because of describing Gibbs free energy of a system, which is the maximum effort of a system to do work at constant pressure (p) and constant temperature (T). For lipid bilayers, control of pressure is possible upon constant membrane area (NPAT) or constant surface torsion "gamma" (NPγT)(Figure 2.4).

system walls impermeable to heat and matter	$T_{in}=T_{ext}$ system heat bath	p_{ext} p_{in} system heat bath $p_{in}=p_{ext}$ $T_{in}=T_{ext}$
Constant number of particles (N), volume (V), and total energy (E).	Constant number of particles (N), volume (V), and temperature (T).	Constant number of particles (N), pressure (p), and temperature (T).
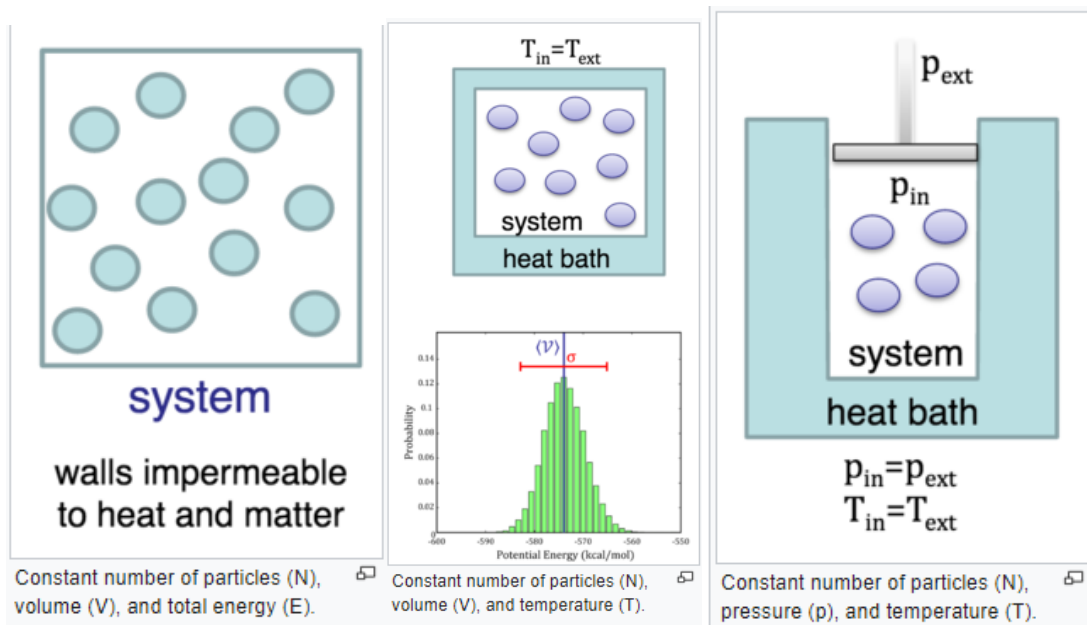
Figure 2. 4 Different ensembles in the systems

The components of the system in this studies are protein, lipid, water molecules, and ions and for all of the systems the NPT ensemble was used. The type of water used in this study in TP3P water is one of the 3 site water models and is popular because of its simplicity and water efficiency[76] . This water model has been re-parameterized to improve energy and density for the liquid water. Also, it has good quality and computational cost which is why we used that in this thesis. And in final the lipid that was applied in the systems is 1-palmitoyl-2-oleoyl-sn-glycerol-3-phosphocholine (POPC). This lipid is widely applied in molecular dynamic simulation of lipid bilayers and its force field parameters were calibrated carefully and they were found in the liquid crystalline phase at 300k temperature condition.

2.3 References

1.   Bacilieri, M. & Moro, S. Ligand-Based Drug Design Methodologies in Drug Discovery Process: An Overview. *Curr. Drug Discov. Technol.* **3**, 155–165 (2007).

2.   Guido, R. V. C., Oliva, G. & Andricopulo, A. D. Virtual screening and its integration with modern drug design technologies. *Curr. Med. Chem.* **15**, 37–46 (2008).

3.   Drwal, M. N. & Griffith, R. Combination of ligand- and structure-based methods in virtual screening. *Drug Discov. Today. Technol.* **10**, e395-401 (2013).

4.   Wang, H., Duffy, R. A., Boykow, G. C., Chackalamannil, S. & Madison, V. S. Identification of novel cannabinoid CB1 receptor antagonists by using virtual screening with a pharmacophore model. *J. Med. Chem.* **51**, 2439–2446 (2008).

5.   Cumming, J. G., Davis, A. M., Muresan, S., Haeberlein, M. & Chen, H. Chemical predictive modelling to improve compound quality. *Nature Reviews Drug Discovery* **12**, 948–962 (2013).

6.   Klebe, G. Virtual ligand screening: strategies, perspectives and limitations. *Drug Discov. Today* **11**, 580–94 (2006).

7.   Leach, A. R., Shoichet, B. K. & Peishoff, C. E. Prediction of protein-ligand interactions. Docking and scoring: successes and gaps. *J. Med. Chem.* **49**, 5851–5 (2006).

8.   Lundstrom, K. Structural genomics and drug discovery: Molecular Pharmacology. *Journal of Cellular and Molecular Medicine* **11**, 224–238 (2007).

9.   B-Rao, C., Subramanian, J. & Sharma, S. D. Managing protein flexibility in docking and its applications. *Drug Discovery Today* **14**, 394–400 (2009).

10.   Anderson, A. C. The process of structure-based drug design. *Chemistry and Biology* **10**, 787–797 (2003).

11.   Bajorath, J. Rational drug discovery revisited: interfacing experimental programs with bio- and chemo-informatics. *Drug Discov. Today* **6**, 989–995 (2001).

12.   Becker, O. M., Shacham, S., Marantz, Y. & Noiman, S. Modeling the 3D structure of GPCRs: advances and application to drug discovery. *Curr. Opin. Drug Discov. Devel.* **6**, 353–61 (2003).

13.   Cavasotto, C. N. & Phatak, S. S. Homology modeling in drug discovery: current trends and

applications. *Drug Discov. Today* **14**, 676–83 (2009).

14. Hillisch, A., Pineda, L. F. & Hilgenfeld, R. Utility of homology models in the drug discovery process. *Drug Discov. Today* **9**, 659–669 (2004).

15. Higgins, D. G., Thompson, J. D. & Gibson, T. J. [22] Using CLUSTAL for multiple sequence alignments. *Methods Enzymol.* **266**, 383–400 (1996).

16. CLUSTALW. https://www.genome.jp/tools-bin/clustalw

17. Nayeem, A., Sitkoff, D. & Krystek, S. A comparative study of available software for high-accuracy homology modeling: from sequence alignments to structural models. *Protein Sci.* **15**, 808–24 (2006).

18. Burley, S. K., Joachimiak, A., Montelione, G. T. & Wilson, I. A. Contributions to the NIH-NIGMS Protein Structure Initiative from the PSI Production Centers. *Structure* **16**, 5–11 (2008).

19. Ginalski, K. Comparative modeling for protein structure prediction. *Curr. Opin. Struct. Biol.* **16**, 172–7 (2006).

20. Ibrahim Uba, A. & Yelekçi, K. Homology modeling of human histone deacetylase 10 and design of potential selective inhibitors. *J. Biomol. Struct. Dyn.* (2019). doi:10.1080/07391102.2018.1521747

21. Mallavarapu, B. D., Abdullah, M., Saxena, S. & Guruprasad, L. Inhibitor binding studies of Mycobacterium tuberculosis MraY (Rv21 56c): Insights from molecular modeling, docking, and simulation studies. *J. Biomol. Struct. Dyn.* **37**, 3751–3763 (2019).

22. Shamsara, J. Homology Modeling of 5-alpha-Reductase 2 Using Available Experimental Data. *Interdiscip. Sci.* **11**, 475–484 (2019).

23. Park, H. *et al.* Discovery of novel α-glucosidase inhibitors based on the virtual screening with the homology-modeled protein structure. *Bioorganic Med. Chem.* **16**, 284–292 (2008).

24. Muhammed, M. T. & Aki-Yalcin, E. Homology modeling in drug discovery: Overview, current applications, and future perspectives. *Chemical Biology and Drug Design* **93**, 12–20 (2019).

25. Luther, K. B., Schindelin, H. & Haltiwanger, R. S. Structural and mechanistic insights into lunatic fringe from a kinetic analysis of enzyme mutants. *J. Biol. Chem.* **284**, 3294–305 (2009).

26. Gagnidze, K. *et al.* Homology modeling and site-directed mutagenesis to identify selective

inhibitors of endothelin-converting enzyme-2. *J. Med. Chem.* **51**, 3378–87 (2008).

27.   Ballesteros, J. A. & Weinstein, H. [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* **25**, 366–428 (1995).

28.   C`Özlem, A., Bishop, T., De Beer, T. A. P. & Joubert, F. *Protein homology modelling and its use in South Africa*.

29.   Kufareva, I. & Abagyan, R. Methods of protein structure comparison. *Methods Mol. Biol.* **857**, 231–57 (2012).

30.   Shen, M. & Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **15**, 2507–2524 (2006).

31.   Nayeem, A., Sitkoff, D. & Krystek, S. A comparative study of available software for high-accuracy homology modeling: From sequence alignments to structural models. *Protein Sci.* **15**, 808–824 (2006).

32.   Zhang, Y. & Skolnick, J. Scoring function for automated assessment of protein structure template quality. *Proteins Struct. Funct. Genet.* **57**, 702–710 (2004).

33.   Topiol, S. & Sabio, M. Use of the X-ray structure of the Beta2-adrenergic receptor for drug discovery. *Bioorg. Med. Chem. Lett.* **18**, 1598–602 (2008).

34.   Sousa, S. F., Fernandes, P. A. & Ramos, M. J. Protein-ligand docking: current status and future challenges. *Proteins* **65**, 15–26 (2006).

35.   Meng, X.-Y., Zhang, H.-X., Mezei, M. & Cui, M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Curr. Comput. Aided-Drug Des.* **7**, 146–157 (2012).

36.   McConkey, B. J., Sobolev, V. & Edelman, M. The performance of current methods in ligand–protein docking. *Current Science* **83**, 845–856

37.   Fischer, E. Einfluss der configuration auf die wirkung der enzyme [The influence of configuration on the effect of enzymes]. *Ber. Dtsch. Chem. Ges.* **27**, 2985–2993 (1894).

38.   Ferreira, L. G., Dos Santos, R. N., Oliva, G. & Andricopulo, A. D. Molecular docking and structure-based drug design strategies. *Molecules* **20**, 13384–13421 (2015).

39. Śledź, P. Protein structure-based drug design: from docking to molecular dynamics. *Curr. Opin. Struct. Biol.* **48**, 93–102 (2018).

40. Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W. & Taylor, R. D. Improved protein-ligand docking using GOLD. *Proteins Struct. Funct. Bioinforma.* **52**, 609–623 (2003).

41. Molecular Operating Environment (MOE) | MOEsaic | PSILO. Available at: https://www.chemcomp.com/Products.htm. (Accessed: 27th October 2019)

42. Pagadala, N. S., Syed, K. & Tuszynski, J. Software for molecular docking: a review. *Biophys. Rev.* **9**, 91–102 (2017).

43. Jones, G., Willett, P., Glen, R. C., Leach, A. R. & Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **267**, 727–748 (1997).

44. Friesner, R. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).

45. Jones, G., Willett, P., Glen, R. C., Leach, A. R. & Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **267**, 727–748 (1997).

46. Wang, Z. *et al.* Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: the prediction accuracy of sampling power and scoring power. *Phys. Chem. Chem. Phys.* **18**, 12964–75 (2016).

47. Vasseur, C. *et al.* A chorionic gonadotropin-sensitive mutation in the follicle-stimulating hormone receptor as a cause of familial gestational spontaneous ovarian hyperstimulation syndrome. *N. Engl. J. Med.* **349**, 753–9 (2003).

48. Verdonk, M. L. *et al.* Modeling water molecules in protein-ligand docking using GOLD. *J. Med. Chem.* **48**, 6504–15 (2005).

49. Kini , J., et al. Automation of process of devising newer Chromone derivatives from existing Natural and Synthetic Chromones by their Study of Structure Activity Relationship – a novel approach. *Inernational J. Sci. Res.* **01**, (2013).

50. Richard A. Friesner, *,† *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. (2004). doi:10.1021/JM0306430

51. Hospital, A., Goñi, J. R., Orozco, M. & Gelpí, J. L. Molecular dynamics simulations: Advances and applications. *Advances and Applications in Bioinformatics and Chemistry* **8**, 37–47 (2015).

52. Xie, Q. & Tinker, R. Molecular dynamics simulations of chemical reactions for use in education. *J. Chem. Educ.* **83**, 77–83 (2006).

53. Molecular Dynamics for Everyone: A Technical Introduction to the Molecular Workbench Software. (2010). doi:10.13140/RG.2.2.15942.32325

54. Wereszczynski, J. & McCammon, J. A. Statistical mechanics and molecular dynamics in evaluating thermodynamic properties of biomolecular recognition. *Quarterly Reviews of Biophysics* **45**, 1–25 (2012).

55. Zhang, Z. H. & Linn, M. C. Can Generating Representations Enhance Learning With Dynamic Visualizations? *J Res Sci Teach* **48**, 1177–1198 (2011).

56. Ercolessi, F. A molecular dynamics primer. *Spring Coll. Comput. physics, ICTP, Trieste* 19 (1997).

57. Katchalski-Katzir, E. *et al.* Molecular surface recognition: Determination of geometric fit between proteins and their ligands by correlation techniques. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 2195–2199 (1992).

58. Miyamoto, S. & Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **13**, 952–962 (1992).

59. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).

60. Kräutler, V., Van Gunsteren, W. F. & Hünenberger, P. H. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J. Comput. Chem.* **22**, 501–508 (2001).

61. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).

62. Hansson T, Oostenbrink C, van G. W. Molecular dynamics simulations - ScienceDirect. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0959440X02003081?via%3Dihub. (Accessed: 27th October 2019)

63.   Karplus, M. & McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nature Structural Biology* **9**, 646–652 (2002).

64.   González, M. A. Force fields and molecular dynamics simulations. *École thématique la Société Française la Neutron.* **12**, 169–200 (2011).

65.   Pearlman, D. A. *et al.* AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.* **91**, 1–41 (1995).

66.   Brooks, B. R. *et al.* CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217 (1983).

67.   MacKerell, A. D. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).

68.   Mackerell, A. D., Feig, M. & Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulation. *J. Comput. Chem.* **25**, 1400–1415 (2004).

69.   Van Gunsteren, W. F. & Berendsen, H. J. C. Algorithms for macromolecular dynamics and constraintdynamics. *Mol. Phys.* **34**, 1311–1327 (1977).

70.   Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118**, 11225–11236 (1996).

71.   Kaminski, G. A., Friesner, R. A., Tirado-Rives, J. & Jorgensen, W. L. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **105**, 6474–6487 (2001).

72.   Sambasivarao, S. V. & Acevedo, O. Development of OPLS-AA force field parameters for 68 unique ionic liquids. *J. Chem. Theory Comput.* **5**, 1038–1050 (2009).

73.   Lennard-Jones, J. E. Cohesion. Proceedings of the Physical Society. **43**, 461 (1931).

74.   *Molecular Dynamics Molecular Dynamics Chapter 2. Molecular Dynamics in Various Ensembles*.

75.   Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).

76.   Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

CHAPTER 3: Using Molecular Dynamics for the refinement of atomistic models of GPCRs by homology modeling

3.1 Introduction

GPCRs are key mediators of the communication of a cell with its environment. These receptors are able to recognize a diverse array of molecules of the extracellular milieu that may trigger a variety of intracellular signaling cascades in response. Their biological function is so relevant that almost a 50% of the drugs in the market target GPCRs[1–5]. Progress in understanding their function, as well as the details of the mechanisms involved in signal transduction are of major importance for designing new drugs, more selective and with minimal side effects. Indeed, to fully understand GPCRs function, a detailed knowledge of their structure as well as their plasticity is of surmount importance [7]. Actually, the relationship between the physicochemical properties of molecules and biological activity to their structure is a well-established paradigm in chemistry[6]. The idea that a more complete knowledge on GPCRs structure can provide us a better understanding of their function was recognized in the 2012 edition of the Nobel Prize in Chemistry, awarded to Brian Kobilka and Robert Lefkowitz for their contribution to present knowledge of GPCRs structure from crystallographic studies. Indeed, in their announcement, the Swedish Academy underlined that structural studies were "crucial for understanding how G protein-coupled receptors function"[8].

In the recent years, we have witnessed an important increase of the number of crystallographic structures of GPCRs available. Since the determination of the structure of bovine rhodopsin at atomic resolution in the year 2000[17], today we have available experimental structures of different 63 unique receptors[9]. Although an important figure, it only represents less than a 10% of the number of GPCRs already described (more than 800). Actually, determination of the crystallographic structure of membrane proteins still presents big challenges that hamper the availability of novel structures, including their low-expression yields, low receptor stability after detergent extraction from native membranes, and high conformational heterogeneity[10].

The considerable number of available GPCR structures permits to identify commonalities and differences. Thus, analysis of the structures reveals that although GPCRs share a bundle of seven transmembrane (TM) α-helices connected by three intracellular loops and three extracellular loops, they show specific structural differences that may be relevant for drug design studies. Ligands from the external milieu bind primarily to the orthosteric site, a pocket located on the extracellular side of the helix bundle that is highly conserved among the members of a subfamily. Furthermore, allosteric ligands have also been described. These ligands bind on the extracellular surface of the receptor and modulate the action of orthosteric ligands[11,12]. Moreover, the structures of GPCRs are not rigid. Actually, GPCRs are very

plastic and different ligands can affect the intrinsic dynamics of a receptor and activate distinct G protein-mediated signaling pathways, a feature defined as "functional selectivity"[13]. This unprecedented knowledge on the structure of GPCRs has been key for carrying out structure-based drug design studies and for the discovery of more selective ligands[1–3]. To extend the use of these methods to other proteins which structure is not available, it is necessary to have a robust method for modeling the structures of GPCRs at atomic resolution that is reliable. These models are useful tools to get insight into the structure-activity relationships of the receptor-ligand complex for the discovery of novel molecules with therapeutic activity[14–21].

Homology modeling is a technique widely used for modeling 3D structures of GPCRs which crystallographic structure is not known[22]. The procedure is based on the paradigm that two proteins with high sequence identity have similar 3D structures. Actually, protein structure is more conserved than its sequence so that structures can be constructed from the template with the highest sequence identity[6,9,10,14,15]. Specifically, it consists of using the backbone structure of a template protein to thread the sequence of the target protein. Obviously, the procedure requires a careful selection of the template and optimization of side chains conformation of the target protein. Despite its apparent simplicity, homology modeling of GPCRs presents many difficulties including the low sequence identity among receptors that spite the common structure, makes challenging to get an accurate structure, reliable enough to carry out rational drug design[23,24]. Fortunately, results on mutagenesis studies and biophysical methods are important to validate the models[20,21,25].

Since bovine rhodopsin was the first crystal structure of a GPCR solved at atomic resolution and had no competitor for several years[19,26], it was routinely used as template to construct 3D models of other GPCRs of the rhodopsin-like family by homology modeling[19,20,24,27]. Despite that sequence identity with other GPCR is less than 20 %, modeling was guided by some conserved motifs in specific positions exhibited by all receptors of the rhodopsin-like family[19,28]. However, major criticism to these models came from the low sequence identity with other GPCRs[25,29,30] and the diversity of binding pockets they exhibit[29]. These models, although useful, do not have the accuracy necessary for reliable structure-based drug design. So that when the crystallographic structures of other GPCRs were available, they were used as templates in such a way that more accurate models could be constructed, by using the closest template to the target.

Presently, modeling approaches achieve close-to-experimental accuracy for small rigid orthosteric ligands when templates with high sequence identity are used[7]. However, there are still many issues that

need to be addressed to improve the quality of the models constructed by homology modeling, such as selection of the best template[31,32] , importing the previous constraints from knowledge-based study to the process of the modeling[32,33] , constructing the extracellular loop[34,35] and refinement of the starting model[18,36–38].

In the previous study, we modeled the human muscarinic M3 receptor using rhodopsin as template and refining the structure produced in a lipid membrane environment using molecular dynamics simulations[18]. After its publication, a few months later the crystallographic structure of the rat muscarinic M3 receptor bound to the bronchodilator drug tiotropium was available[39]. This permitted us to assess the accuracy of the model constructed. The analysis suggested that the homology modeling process needed improvements. However, it was not clear the source of the inaccuracies during the modeling process. They could have been due to the election of the template or due to the refinement process.

In order to understand the effect of template selection and the refinement process on the construction of an accurate model of the muscarinic M3 receptor, we planned a set of calculations described in the present chapter and published elsewhere[40,41]. The goal of the present study is to construct diverse models of the muscarinic M3 receptor using different templates and refinement conditions and compare with its crystallographic structure. Specifically, we constructed models of the muscarinic M3 receptor using three different templates: the muscarinic M2 receptor (very close in the phylogenetic tree), the histamine H1 receptor (at intermediate distance in the phylogenetic tree) and rhodopsin (distant in the phylogenetic tree). Refinement of the structures was carried out using molecular dynamics of the models constructed embedded in a lipid bilayer for about 500 ns. In order to understand the effect on the refinement of having a ligand bound to the receptor, we performed five different refinements. In the case of using the muscarinic M2 receptor as template, a model was refined without ligand, another one with tiotropium bound (the ligand bound to the muscarinic M3 crystallographic structure), and another one with N-methylscopolamine (a non-selective muscarinic antagonist) bound. Refinements using the histidine H1 and rhodopsin as templates were carried out with tiotropium bound.

3.2 Computational Procedure

3.2.1 Constructing M3 Muscarinic Receptor Model by Homology Modeling

The steps necessary for the construction of a model by homology modeling are: multiple sequence alignment, selection of the template, alignment of the sequences of template and target, first model construction and refinement of the constructed model [42,43]. These are the steps detailed below for the present case.

Homology models of the rat muscarinic M3 receptor were constructed using diverse templates by means of the Molecular Operating Environment (MOE) package[44]. The protein sequence of the rat M3 muscarinic receptor (corresponding to the available crystallographic structure) was retrieved from the Uniprot database (ID: P08484). For the rest of the templates, sequences were taken from the corresponding crystallographic structures, taking care of removing those residues corresponding to the T4 Lysozyme inserted between TM5 and TM6 from the respective sequences, if applicable: PDB ID: 3UON for the human muscarinic M2[45]. PDB ID: 3RZE for the human histamine H1[46] and PDB ID: 1GZM for bovine rhodopsin[17].

Multiple sequence alignment of the sequences was carried out using the MOE-Align tool.[35] For this purpose, diverse sequences taken from diverse phylogenetic branches of the Rhodopsin like family were selected to carry out a multiple sequence alignment. Once the alignment was completed and checked for unwanted gaps within the TM regions, we proceeded to construct diverse atomistic models for each template. In order to simplify the construction of the atomistic models of the rat muscarinic M3 receptor without compromising their accuracy, the long intracellular loop (ICL3) was substituted by a stretch of a few residues. Specifically, in the case of both the M2 muscarinic and the histamine H1 receptors, since both crystallographic structures correspond to fusion proteins of the receptor with the T4-lysozyme inserted in the ICL3, we simply removed the sequence of lysozyme and joined the residues left. Thus, in the case of the M2 muscarinic receptor four residues from the C-terminus of TM5 and seven residues from the N-terminus of TM6 and five residues from the C-terminus of TM5 and three residues from the N-terminus of TM6 in the case of the histamine H1 receptor. In the case of rhodopsin, we simply considered all residues of the corresponding ICL3. Subsequently, diverse crude homology models for each of the templates were constructed by threading the sequence of the target receptor to each of the crystallographic structures with the subsequent incorporation of alternative sidechain conformations using an extensive rotamer library generated from a high-resolution structural database embedded in MOE[35]. Once hydrogens were added using the protonate3D method[36], crude models were energy minimized using a contact energy function to relieve any serious steric strains. The diverse models generated for each template were scored according to their rmsd to the average structure using the Cα atoms for the calculation. The model with the highest score was considered as a crude model and used for further refinement. Finally, the stereochemical quality of the models constructed was assessed by the distribution of the backbone dihedral angles in the Ramachandran map.

3.2.2 Molecular Docking

Before performing the refinement process, the antagonist tiotropium (1) (Figure 1) was docked onto the orthosteric binding site of the constructed model of the M3 muscarinic receptor using the Glide algorithm[47]. The compound was docked in multiple orientations and multiple conformations. The resulted poses were rank ordered by their binding/docking score using the glidescore function. The best poses were analyzed visually and validated in accordance to the information available on the involvement of specific residues in binding from diverse site directed mutagenesis studies[38]. Crude homology models with tiotropium bound were used as starting models for refinement. In addition, in order to understand the effect of using a ligand bound in the refinement process, two additional starting models for the M2 template were constructed: one bound to the antagonist NMS (2) (Figure 3.1) and another without any ligand bound. Henceforth, four protein-ligand complexes of the M3 muscarinic receptor and one without ligand were refined using molecular dynamic simulations.

(1) **(2)**
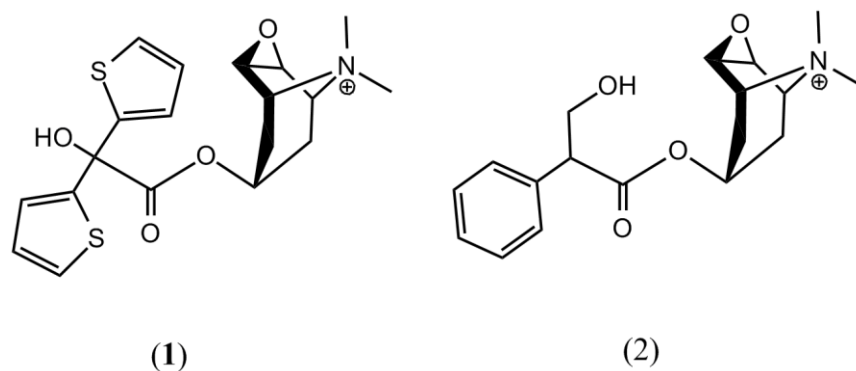
Figure 3. 1 *Chemical structures of tiotropium (1) and N-methylscopolamine (NMS) (2).*

3.2.3 Molecular Refinement

Models were refined using MD simulations[39,40]. For each model, the starting structure was embedded into a box of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipids and water molecules previously equilibrated according to the procedure described elsewhere[41]. The box had an initial size of 10.3 x 8.0 x 10.2 nm3 (XYZ), organized in such a way that the bilayer plane was oriented on the XY plane. Before protein insertion, the box contained 256 lipids (corresponding to an area per lipid of 0.64 nm2) and circa 17,000 water molecules. The protein was placed in the center of the box, and the overlapping molecules were removed. Specifically, all water molecules with oxygen atoms closer than 0.40 nm to a non-hydrogen atom of the protein, as well as all lipid molecules with at least one atom closer

74

than 0.25 nm to a non-hydrogen atom of the protein were removed. This resulted in a final system containing 188 lipids and circa 14,655 water molecules. Removal of these atoms introduced small voids between the protein and water or lipid molecules that disappeared during the first part of the MD simulation, in which a progressive adjustment of the lipid bilayer and water molecules to the protein takes place. Next, 105 randomly selected water molecules were replaced by 45 sodium and 60 chloride ions, providing a neutral system with a concentration approximately 0.2 M on sodium chloride. This concentration is fairly similar to that found in biological organisms, although they exhibit different intra- and extra-cellular ion concentrations.

For each model, a 500 ns MD simulation was carried out at constant pressure using the GROMACS package 4.6[42]. The OPLSAA force field,[43] currently implemented in GROMACS, was used to describe all molecules of the system, except for water that was modeled using the TIP3P model[44]. The systems were subjected to periodic boundary conditions in the three coordinate directions. The temperature was kept constant at 300 K using separate thermostats for the protein, water, ions, and lipid molecules. The time constant for the thermostats was set to 0.1 ps except for water, for which a smaller value of 0.01 ps was used. The pressure in the three coordinate directions was kept at 0.1 MPa by independent Berendsen barostats using a time constant of 1.0 ps. The equations of motion were integrated using the leapfrog algorithm with a time step of 2 fs. All bonds involving hydrogen atoms within the protein and lipid molecules were kept frozen using the LINCS algorithm[45]. The bonds and the angle of water molecules were fixed using the analytical SETTLE method. Lennard–Jones interactions were computed using a cutoff of 1.0 nm. The electrostatic interactions were treated either using the PME technique[46].

3.3 Results and Discussion

3.3.1 Construction of Crude Models

As mentioned above, the goal of the present study was to investigate the performance of molecular dynamics simulations used in the refinement process. For this purpose, we carried out a "semi-blind" homology modeling study of the rat M3 muscarinic receptor whose crystallographic structure is available, using the human M2 muscarinic receptor with and without an antagonist bound to it, the human H1 histamine receptor and bovine rhodopsin as templates.

Sequence alignment of the rat muscarinic M3 receptor with the human muscarinic M2, the human histamine H1, and bovine rhodopsin shows sequence identities of 60%, 24% and 17%, respectively (see Figure 3.2). These values are larger if only the transmembrane (TM) regions are considered, with values of 79%, 32% and 23%, respectively. Interestingly, the best sequence identity score is found in the TM3

segment, whereas the poorer is found in the TM1 segment in all the cases[47].As in the rest of the muscarinic receptor subtypes, the intracellular loop ICL3 of M3 is very long and was omitted for modeling purposes. The deletion of this large segment bears no consequences to the overall structure of the GPCR or the orthosteric binding pocket as in the muscarinic M3 receptor-phage T4 lysozyme fused protein used for crystallographic studies where, receptor ability to bind agonist or antagonist ligands is not modified.

As mentioned above, crude models were constructed by threading the sequence of the target receptor on the crystallographic structure of the templates with the subsequent incorporation of the corresponding side chains using a library of conformers. For each template alternative models were generated using template backbone coordinates with alternative side chain conformations using an extended rotamer library implemented in the MOE package[35]. Following this process a number of independent models, based on loop and side chain placements were scored using a contact energy function. Among these, the model with the highest score was selected for refinement process. These crude models presented similar distances to the target M3 crystallographic structure than the corresponding crystallographic structures of the diverse templates. Thus, the muscarinic M2 receptor exhibits a root-mean square deviation (rmsd) 1.6 Å; 1.8 Å in the case of the histamine H1 receptor; and 2.3 Å in the case of rhodopsin (using the backbone Cα for the measure).

```
Alignment Score 22775

CLUSTAL-Alignment file created  [clustalw.aln]
```

```
clustalw.aln

CLUSTAL 2.1 multiple sequence alignment


sp|P41143|OPRD_HUMAN       --------------------MEPAPSAGAELQPPLFANASDAYPSACP-
sp|P35372|OPRM_HUMAN       MDSSAAPTNASNCTDALAYSSCSPAPSPGSWVNLSHLDGNLSDPCGPNR-
sp|P41146|OPRX_HUMAN       ------------------MEPLFPAP-FWEVIYGSHLQGNLSLLSPNHS-
sp|P08588|ADRB1_HUMAN      ------------------------MGAGVLVLGASEPGNLSSAAPLPDG
sp|P07550|ADRB2_HUMAN      ---------------------------------MGQPGNGSAFLLAPN-
sp|P11229|ACM1_HUMAN       -------------------------------------MNTSAPPAVSPN-
sp|P20309|ACM3_HUMAN       ----------------------MTLHNNSTTSPLFPNISSSWIHSPS-
sp|P08172|ACM2_HUMAN       ---------------------------------------MNNSTNSSNNS-
sp|P28222|5HT1B_HUMAN      ------------------------MEEPGAQCAPPPPAGSETWVPQAN-
sp|P35367|HRH1_HUMAN       ------------------------------MSLPNSSCLLEDKMCEGN-
sp|P35462|DRD3_HUMAN       -----------------------------MASLSQLSSHLNYTCGAEN-
sp|P41595|5HT2B_HUMAN      -------------------MALSYRVSELQSTIPEHILQSTFVHVISSN-
sp|P61073|CXCR4_HUMAN      ----------------------------MEGISIYTSDNYTEEMGSGD-
sp|P51681|CCR5_HUMAN       -------------------------MDYQVSSPIYDINYYTSEPCQK-
sp|P29274|AA2AR_HUMAN      -------------------------------------------------
sp|Q9H244|P2Y12_HUMAN      ----------------------------------MQAVDNLTSAPGNTS-
sp|P25116|PAR1_HUMAN       -----MGPRRLLLVAACFSLCGPLLSARTRARRPESKATNATLDPRSFLL
sp|P30989|NTR1_HUMAN       ----------MRLNSSAPGTPGTPAADPFQRAQAGLEEALLAPGFGNASG


sp|P41143|OPRD_HUMAN       -----------SAGANASGPPG------------------------ARS
sp|P35372|OPRM_HUMAN       -----------TDLGGRDSLCP------------------------PTG
sp|P41146|OPRX_HUMAN       -----------LLPPHLLLNAS------------------------HGA
sp|P08588|ADRB1_HUMAN      AATAARLLVPASPPASLLPPAS------------------------ESPE
sp|P07550|ADRB2_HUMAN      ---------------GSHAPDH------------------------DVTQ
sp|P11229|ACM1_HUMAN       --------------ITVLAP------------------------------
sp|P20309|ACM3_HUMAN       --------------DAGLPPGTVTHFGSYNVSRAAGNFSSPDGTTDDPLG
sp|P08172|ACM2_HUMAN       --------------LALTSP------------------------------
sp|P28222|5HT1B_HUMAN      ---------------LSSAPSQ-----------------NCSAKDYIYQD
sp|P35367|HRH1_HUMAN       --------------------------------------------------KT
sp|P35462|DRD3_HUMAN       --------------------------------------------------STG
sp|P41595|5HT2B_HUMAN      ----------WSGLQTESIPEEMK----------------------QIV
sp|P61073|CXCR4_HUMAN      -------------YDSMKEP--------------------------CFRE
sp|P51681|CCR5_HUMAN       --------------------------------------------------I
sp|P29274|AA2AR_HUMAN      -------------------------------------------------
sp|Q9H244|P2Y12_HUMAN      ------------------------------------------------LCT
sp|P25116|PAR1_HUMAN       RNPNDKYEPFWEDEEKNESGLTEYRLVSINKSSPLQKQLPAFISEDASGY
sp|P30989|NTR1_HUMAN       ----------NASERVLAAPSS-----------------------ELDV

                                  _____                          _____
sp|P41143|OPRD_HUMAN       ASSLALAIAITALYSAVCAVGLLGNVLVMFGIV---RYTKMKTATNIYIF
sp|P35372|OPRM_HUMAN       SPSMITAITIMALYSIVCVVGLFGNFLVMYVIV---RYTKMKTATNIYIF
sp|P41146|OPRX_HUMAN       FLPLGLKVTIVGLYLAVCVGGLLGNCLVMYVIL---RHTKMKTATNIYIF
sp|P08588|ADRB1_HUMAN      PLSQQWTAGMGLLMALIVLLIVAGNVLVIVAIA---KTPRLQTLTNLFIM
sp|P07550|ADRB2_HUMAN      ERDEVWVVGMGIVMSLIVLAIVFGNVLVITAIA---KFERLQTVTNYFIT
sp|P11229|ACM1_HUMAN       GKGPWQVAFIGITTGLLSLATVTGNLLVLISFK---VNTELKTVNNYFLL
sp|P20309|ACM3_HUMAN       GHTVWQVVFIAFLTGILALVTIIGNILVIVSFK--VNKQLKTVNNYFLL
sp|P08172|ACM2_HUMAN       -YKTFEVVFIVLVAGSLSLVTIIGNILVMVSIK---VNRHLQTVNNYFLF
sp|P28222|5HT1B_HUMAN      SISLPWKVLLVMLLALITLATTLSNAFVIATVY---RTRKLHTPANYLIA
sp|P35367|HRH1_HUMAN       TMASPQLMPLVVVLSTICLVTVGLNLLVLYAVR--SERKLHTVGNLYIV        .
sp|P35462|DRD3_HUMAN       ASQARPHAYYALSYCALILAIVFGNGLVCMAVL---KERALQTTTNYLVV
sp|P41595|5HT2B_HUMAN      EEQGNKLHWAALLILMVIIPTIGGNTLVILAVS---LEKKLQYATNYFLM
sp|P61073|CXCR4_HUMAN      ENANFNKIFLPTIYSIIFLTGIVGNGLVILVMG---YQKKLRSMTDKYRL
sp|P51681|CCR5_HUMAN       NVKQIAARLLPPLYSLVFIFGFVGNMLVILILI---NCKRLKSMTDIYLL
sp|P29274|AA2AR_HUMAN      -MPIMGSSVYITVELAIAVLAILGNVLVCWAVW---LNSNLQNVTNYFVV
sp|Q9H244|P2Y12_HUMAN      RDYKITQVLFPLLYTVLFFVGLITNGLAMRIFF---QIR--SKSNFIIFL
sp|P25116|PAR1_HUMAN       LTSSWLTLFVPSVYTGVFVVSLPLNIMAIVVFI---LKMKVKKPAVVYML
sp|P30989|NTR1_HUMAN       NTDIYSKVLVTAVYLALFVVGTVGNTVTAFTLARKKSLQSLQSTVHYHLG
                                            :        *  ..    .

sp|P41143|OPRD_HUMAN       NLALADALA-TSTLPFQSAKYLME--TWPFGELLCKAVLSIDYYNMFTSI
```

```
sp|P35372|OPRM_HUMAN    NLALADALA-TSTLPFQSVNYLMG--TWPFGTILCKIVISIDYYNMFTSI
sp|P41146|OPRX_HUMAN    NLALADTLV-LLTLPFQGTDILLG--FWPFGNALCKTVIAIDYYNMFTST
sp|P08588|ADRB1_HUMAN   SLASADLVMGLLVVPFGATIVVWG--RWEYGSFFCELWTSVDVLCVTASI
sp|P07550|ADRB2_HUMAN   SLACADLVMGLAVVPFGAAHILMK--MWTFGNFWCEFWTSIDVLCVTASI
sp|P11229|ACM1_HUMAN    SLACADLIIGTFSMNLYTTYLLMG--HWALGTLACDLWLALDYVASNASV
sp|P20309|ACM3_HUMAN    SLACADLIIGVISMNLFTTYIIMN--RWALGNLACDLWLAIDYVASNASV
sp|P08172|ACM2_HUMAN    SLACADLIIGVFSMNLYTLYTVIG--YWPLGPVVCDLWLALDYVVSNASV
sp|P28222|5HT1B_HUMAN   SLAVTDLLVSILVMPISTMYTVTG--RWTLGQVVCDFWLSSDITCCTASI
sp|P35367|HRH1_HUMAN    SLSVADLIVGAVVMPMNILYLLMS--KWSLGRPLCLFWLSMDYVASTASI
sp|P35462|DRD3_HUMAN    SLAVADLLVATLVMPWVVYLEVTGG-VWNFSRICCDVFVTLDVMMCTASI
sp|P41595|5HT2B_HUMAN   SLAVADLLVGLFVMPIALLTIMFEA-MWPLPLVLCPAWLFLDVLFSTASI
sp|P61073|CXCR4_HUMAN   HLSVADLLF-VITLPFWAVDAVAN---WYFGNFLCKAVHVIYTVNLYSSV
sp|P51681|CCR5_HUMAN    NLAISDLFF-LLTVPFWAHYAAAQ---WDFGNTMCQLLTGLYFIGFFSGI
sp|P29274|AA2AR_HUMAN   SLAAADIAVGVLAIPFAITISTGF----CAACHGCLFIACFVLVLTQSSI
sp|Q9H244|P2Y12_HUMAN   KNTVISDLLMILTFPFKILSDAKLG-TGPLRTFVCQVTSVIFYFTMYISI
sp|P25116|PAR1_HUMAN    HLATADVLF-VSVLPFKISYYFSGS-DWQFGSELCRFVTAAFYCNMYASI
sp|P30989|NTR1_HUMAN    SLALSDLLTLLLAMPVELYNFIWVHHPWAFGDAGCRGYYFLRDACTYATA
                          :  .       .                          *

sp|P41143|OPRD_HUMAN    FTLTMMSVDRYIAVCHPVKALDFRTPAKAKLINICIWVLA--SGVGVPIM
sp|P35372|OPRM_HUMAN    FTLCTMSVDRYIAVCHPVKALDFRTPRNAKIINVCNWILS--SAIGLPVM
sp|P41146|OPRX_HUMAN    FTLTAMSVDRYVAICHPIRALDVRTSSKAQAVNVAIWALA--SVVGVPVA
sp|P08588|ADRB1_HUMAN   ETLCVIALDRYLAITSPFRYQSLLTRARARGLVCTVWAISA-LVSFLPIL
sp|P07550|ADRB2_HUMAN   ETLCVIAVDRYFAITSPFKYQSLLTKNKARVIILMVWIVSG-LTSFLPIQ
sp|P11229|ACM1_HUMAN    MNLLLISFDRYFSVTRPLSYRAKRTPRRAALMIGLAWLVS--FVLWAPAI
sp|P20309|ACM3_HUMAN    MNLLVISFDRYFSITRPLTYRAKRTTKRAGVMIGLAWVIS--FVLWAPAI
sp|P08172|ACM2_HUMAN    MNLLIISFDRYFCVTKPLTYPVKRTTKMAGMMIAAAWVLS--FILWAPAI
sp|P28222|5HT1B_HUMAN   LHLCVIALDRYWAITDAVEYSAKRTPKRAAVMIALVWVFS--ISISLPPF
sp|P35367|HRH1_HUMAN    FSVFILCIDRYRSVQQPLRYLKYRTKTRASATILGAWFLS----FLWVIP
sp|P35462|DRD3_HUMAN    LNLCAISIDRYTAVVMPVHYQHGTGQSSCRRVALMITAVWV-LAFAVSCP
sp|P41595|5HT2B_HUMAN   MHLCAISVDRYIAIKKPIQANQYNSRATAFIKITVVWLISIGIAIPVPIK
sp|P61073|CXCR4_HUMAN   LILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPA--LLLTIPDF
sp|P51681|CCR5_HUMAN    FFIILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVA--VFASLPGI
sp|P29274|AA2AR_HUMAN   FSLLAIAIDRYIAIRIPLRYNGLVTGTRAKGIIAICWVLS--FAIGLTPM
sp|Q9H244|P2Y12_HUMAN   SFLGLITIDRYQKTTRPFKTSNPKNLLGAKILSVVIWAFM--FLLSLPNM
sp|P25116|PAR1_HUMAN    LLMTVISIDRFLAVVYPMQSLSWRTLGRASFTCLAIWALA--IAGVVPLL
sp|P30989|NTR1_HUMAN    LNVASLSVERYLAICHPFKAKTLMSRSRTKKFISAIWLASALLAVPMLFT
                          :  : .:*:.        .

sp|P41143|OPRD_HUMAN    VMAV-------TRPRDGAVVCMLQFPSPS-WYWDTVTKICVFLFAFVVPI
sp|P35372|OPRM_HUMAN    FMAT-------TKYRQGSIDCTLTFSHPT-WYWENLLKICVFIFAFIMPV
sp|P41146|OPRX_HUMAN    IMGS-------AQVEDEEIECLVEIPTPQ-DYWGPVFAICIFLFSFIVPV
sp|P08588|ADRB1_HUMAN   MHWW-------RAESDEARRCYNDPKCCD-FVTNRAYAIASSVVSFYVPL
sp|P07550|ADRB2_HUMAN   MHWY-------RATHQEAINCYANETCCD-FFTNQAYAIASSIVSFYVPL
sp|P11229|ACM1_HUMAN    LFWQ-------YLVG--ERTVLAGQCYIQ-FLSQPIITFGTAMAAFYLPV
sp|P20309|ACM3_HUMAN    LFWQ-------YFVG--KRTVPPGECFIQ-FLSEPTITFGTAIAAFYMPV
sp|P08172|ACM2_HUMAN    LFWQ-------FIVG--VRTVEDGECYIQ-FFSNAAVTFGTAIAAFYLPV
sp|P28222|5HT1B_HUMAN   FWRQ-------AKAEEEVSECVVN-------TDHILYTVYSTVGAFYFPT
sp|P35367|HRH1_HUMAN    ILGW-------NHFMQQTSVRREDKCETD-FYDVTWFKVMTAIINFYLPT
sp|P35462|DRD3_HUMAN    LLFG-------FNTTGDPTVCSIS-------NPDFVIYSSVVSFYLPF
sp|P41595|5HT2B_HUMAN   GIET-------DVDNPNNITCVLTKER------FGDFMLFGSLAAFFTPL
sp|P61073|CXCR4_HUMAN   IFAN-------VSEADDRYICDRFYPN---DLWVVVFQFQHIMVGLILPG
sp|P51681|CCR5_HUMAN    IFTR-------SQKEGLHYTCSSHFPYSQYQFWKNFQTLKIVILGLVLPL
sp|P29274|AA2AR_HUMAN   LGWNNCGQPKEGKNHSQGCGEGQVACLFEDVVPMNYMVYFNFFACVLVPL
sp|Q9H244|P2Y12_HUMAN   ILTNR------QPRDKNVKKCSFLKSEFG--LVWHEIVNYICQVIFWINF
sp|P25116|PAR1_HUMAN    LKEQT-----IQVPGLNITTCHDVLNETLLEGYYAYYFSAFSAVFFFVPL
sp|P30989|NTR1_HUMAN    MGEQ-------NRSADGQHAGGLVCTPTIHTATVKVVIQVNTFMSFIFPM

sp|P41143|OPRD_HUMAN    LIITVCYGLMLLRLRSVRLLSGS--------------------------
sp|P35372|OPRM_HUMAN    LIITVCYGLMILRLKSVRMLSGS--------------------------
sp|P41146|OPRX_HUMAN    LVISVCYSLMIRRLRGVRLLSGS--------------------------
sp|P08588|ADRB1_HUMAN   CIMAFVYLRVFREAQKQVKKIDSCERRFLGGPARPPSPSPSPVPAPAPPP
sp|P07550|ADRB2_HUMAN   VIMVFVYSRVFQEAKRQLQKIDKSEGRFH---------------------
sp|P11229|ACM1_HUMAN    TVMCTLYWRIYRETENRARELAALQGS---------ETPGKGGGSSSSSE
sp|P20309|ACM3_HUMAN    TIMTILYWRIYKETEKRTKELAGLQASGTEAETENFVHPTGSSRSCSSYE
sp|P08172|ACM2_HUMAN    IIMTVLYWHISRASKSRIKKDKKEPVANQDP----VSPSLVQGRIVKPNN
sp|P28222|5HT1B_HUMAN   LLLIALYGRIYVEARSRILKQTPNRTG----------------------
sp|P35367|HRH1_HUMAN    LLMLWFYAKIYKAVRQHCQHRELINRSLPSFSEIKLRPENPKGDAKKPGK
sp|P35462|DRD3_HUMAN    GVTVLVYARIYVVLKQRRRKRILTRQN----------------------
sp|P41595|5HT2B_HUMAN   AIMIVTYFLTIHALQKKAYLVKNKPPQRLTWLTVSTVFQRDETPCSSPEK
sp|P61073|CXCR4_HUMAN   IVILSCYCIIISKLSHS--------------------------------
```

```
sp|P11229|ACM1_HUMAN    ------TVKRPTKKGRDRAGKGQKPRGKEQLAKRKTFSLVKEKKAARTLS
sp|P20309|ACM3_HUMAN    SSVGKSTATLPLSFKEATLAKRFALKTRSQITKRKRMSLVKEKKAAQTLS
sp|P08172|ACM2_HUMAN    --------SSGQNGDEKQNIVARKIVKMTKQPAKKKPPPSREKKVTRTIL
sp|P28222|5HT1B_HUMAN   --------ERKATKTLGIILGAFIVCWLPFFIISLVMPICKDACWFHLAI
sp|P35367|HRH1_HUMAN    TAPGKGKLRSGSNTGLDYIKFTWKRLRSHSRQYVSGLHMNRERKAAKQLG
sp|P35462|DRD3_HUMAN    ------TIAPKLSLEVRKLSNGRLSTSLKLGPLQPRGVPLREKKATQMVA
sp|P41595|5HT2B_HUMAN   ----------YVSSGVNPLVYTLFNKTFRDAFGRYITCNYRATKSVKTLR
sp|P61073|CXCR4_HUMAN   ----ITEALAFFHCCLNPILYAFLGAKFKTSAQHALTSVSRGS-------
sp|P51681|CCR5_HUMAN    ----------MTHCCINPIIYAFVGEKFRNYLLVFFQKHIAKRFCKC---
sp|P29274|AA2AR_HUMAN   ----------HTNSVVNPFIYAYRIREFRQTFRKIIRSHVLRQQEPFKAA
sp|Q9H244|P2Y12_HUMAN   ----------SLNACLDPFIYFFLCKSFRNSLISMLKCP-----------
sp|P25116|PAR1_HUMAN    ----------SISCCIDPLIYYYASSECQRYVYSILCCKESSDPSSYN--
sp|P30989|NTR1_HUMAN    -------ALFYVSSTINPILYNLVSANFRHIFLATLACLCPVWRRRRK--


sp|P41143|OPRD_HUMAN    ---------------------------------------FSRAREATARE
sp|P35372|OPRM_HUMAN    ---------------------------------------STRIRQNTRDH
sp|P41146|OPRX_HUMAN    ---------------------------------------SDRVRSIAKDV
sp|P08588|ADRB1_HUMAN   DRPRASGCLARPGPPPSPGAASDDDDDDV-------VGATPPARLLEPWA
sp|P07550|ADRB2_HUMAN   SN----------GNTGEQSGYHVEQEKEN-------KLLCEDLPGTEDFV
sp|P11229|ACM1_HUMAN    AILLAFILTWTPYNIMVLVSTFCKDCVPETLWE-LGYWLCYVNSTINPMC
sp|P20309|ACM3_HUMAN    AILLAFIITWTPYNIMVLVNTFCDSCIPKTFWN-LGYWLCYINSTVNPVC
sp|P08172|ACM2_HUMAN    AILLAFIITWAPYNVMVLINTFCAPCIPNTVWT-IGYWLCYINSTINPAC
sp|P28222|5HT1B_HUMAN   FDFFT--------------------------------WLGYLNSLINPII
sp|P35367|HRH1_HUMAN    FIMAAFILCWIPYFIFFMVIAFCKNCCNEHLHM-FTIWLGYINSTLNPLI
sp|P35462|DRD3_HUMAN    IVLGAFIVCWLPFFLTHVLNTHCQTCHVSPELYSATTWLGYVNSALNPVI
sp|P41595|5HT2B_HUMAN   KRSSKIYFRNPMAENSKFFKKHGIRNGINPAMYQSPMRLRSSTIQSSSII
sp|P61073|CXCR4_HUMAN   -----------------------------------SLKILSKGKRG
sp|P51681|CCR5_HUMAN    -----------------------------------CSIFQQEAPER
sp|P29274|AA2AR_HUMAN   GTSARVLAAHGSDGEQVSLRLNGHPPGVWANGSAPHPERRPNGYALGLVS
sp|Q9H244|P2Y12_HUMAN   ------------------------------------NSATSLSQDN
sp|P25116|PAR1_HUMAN    ------------------------------------SSGQLMASKM
sp|P30989|NTR1_HUMAN    ------------------------------------RPAFSRKADSV


sp|P41143|OPRD_HUMAN    RVTACTPSDGPGGGAAA---------------------------------
sp|P35372|OPRM_HUMAN    PSTANTVDRTNHQLENLEAETAPLP-------------------------
sp|P41146|OPRX_HUMAN    ALACKTSETVPRPA------------------------------------
sp|P08588|ADRB1_HUMAN   GCNGGAAADSDSSLDEPCRPGFASESKV----------------------
sp|P07550|ADRB2_HUMAN   GHQGTVPSDNIDSQGRNCSTNDSLL-------------------------
sp|P11229|ACM1_HUMAN    YALCNKAFRDTFRLLLLCRWDKRRWRKIPKRPGSVHRTPSRQC----
sp|P20309|ACM3_HUMAN    YALCNKTFRTTFKMLLLCQCDKKKRRKQQYQQRQSVIFHKRAPEQAL
sp|P08172|ACM2_HUMAN    YALCNATFKKTFKHLLMCHYKNIGATR-----------------------
sp|P28222|5HT1B_HUMAN   YTMSNEDFKQAFHKLIRFKCTS----------------------------
sp|P35367|HRH1_HUMAN    YPLCNENFKKTFKRILHIRS------------------------------
sp|P35462|DRD3_HUMAN    YTTFNIEFRKAFLKILSC--------------------------------
sp|P41595|5HT2B_HUMAN   LLDTLLLTENEGDKTEEQVSYV----------------------------
sp|P61073|CXCR4_HUMAN   GHSSVSTESESSSFHSS---------------------------------
sp|P51681|CCR5_HUMAN    ASSVYTRSTGEQEISVGL--------------------------------
sp|P29274|AA2AR_HUMAN   GGSAQESQGNTGLPDVELLSHELKGVCPEPPGLDDPLAQDGAGVS--
sp|Q9H244|P2Y12_HUMAN   RKKEQDGGDPNEETPM----------------------------------
sp|P25116|PAR1_HUMAN    DTCSSNLNNSIYKKLLT---------------------------------
sp|P30989|NTR1_HUMAN    SSNHTLSSNATRETLY----------------------------------
```

Figure 3. 2 *Multiple sequence alignment of the rhodopsin-like family*

Table3.1 Target and template alignment score

HELIX	ALIGNMENT		SCORE

I	`3RZE_A|PDBID|CHAIN|SEQUENCE`
	`sp|P20309|ACM3_HUMAN`
		`-PLVVVLSTICLVTVGLNLLVLYAVRS`
		`VFIAFLTGILALVTIIGNILVIVSFKV`
		`:..: . :.***: *:**: :.:`
			23.0769

II	`3RZE_A|PDBID|CHAIN|SEQUENCE`
	`sp|P20309|ACM3_HUMAN`
		`---VGNLYIVSLSVADLIVGAVVMPMNILYLL`
		`LKTVNNYFLLSLACADLIIGVISMNLFTTYII`
		`*.* :::**: ****:*.: * : *::`
			37.931

III	`3RZE_A|PDBID|CHAIN|SEQUENCE`
	`sp|P20309|ACM3_HUMAN`
		`GRPLCLFWLSMDYVASTASIFSVFILCIDRYRSVQ`
		`GNLACDLWLAIDYVASNASVMNLLVISFDRYFSIT`
		`*. * :**::*****.**::.::::.:*** *:`
			42.8571

IV	`3RZE_A|PDBID|CHAIN|SEQUENCE`
	`sp|P20309|ACM3_HUMAN`
		`KTRASATILGAWFLSF-LWVIPIL`
		`TKRAGVMIGLAWVISFVLW-----`
		`..**.. * **.:** **`
			36.8421

V	`3RZE_A|PDBID|CHAIN|SEQUENCE`
	`sp|P20309|ACM3_HUMAN`
		`TWFKVMTAIINFYLPTLLMLWFYAKIYKA`
		`-TITFGTAIAAFYMPVTIMTILYWRIYK-`
		`:.. *** **:*. :* :* :***`
			40.7407

VI	`3RZE_A|PDBID|CHAIN|SEQUENCE`
	`sp|P20309|ACM3_HUMAN`
		`NRERKAAKQLGFIMAAFILCWIPYFIFFMVIAFC-`
		`VKEKKAAQTLSAILLAFIITWTPYNIMVLVNTFCD`
		`:*:***: *. *: ***: * ** *:.:* :**`
			47.0588

VII	`3RZE_A|PDBID|CHAIN|SEQUENCE`
	`sp|P20309|ACM3_HUMAN`
		`EHLHMFTIWLGYINSTLNPLIYPLCNENFKKTFKRILH`
		`-TFWNLGYWLCYINSTVNPVCYALCNKTFRTTFKMLLL`
		`: : ** *****:**: *.***:.*:.*** :*`
			48.6486

Helix	M3 CRYSTAL STRUCTURE	M3-Tio from HRH1 receptor
TM1	W66-V95	F69-V95
TM2	V102-M130	V103-I129
TM3	N137-T170	G136-T170
TM4	T181-V210	T181-I204
TM5	P228-E258	T229-K255
TM6	Q489-F515	V484-C516
TM7	K255-K555	T523-L558

An important drawback of the crude models constructed by homology modeling regards transmembrane helix lengths. Residues involved in the diverse TMs, both in the crude models and in the M3 crystallographic structure are listed in Table 3.2. As can be seen, for some of the helices differences between the crude model and the M3 target structure are notorious. Thus, in the crude model generated using M2 as template there are not large differences. Important exceptions are TM3 that is seven residues longer than the corresponding helix in the M3 crystallographic structure and TM6 that is five residues longer. In contrast, crude models generated using the histamine H1 and rhodopsin as templates show larger deviations. In the case of the histamine H1 receptor TM1 is three residues shorter; TM4 is six residues shorter; TM5 is four residues shorter; and TM6 as well as TM7 are seven and three residues longer, respectively. In the case of rhodopsin TM4 is six residues shorter, whereas TM5 is five residue longer and TM2 and TM6 exhibit a three residue difference. Taking into account that a α-helix contains 3.6 residues per turn, differences in helix length can sum up to two turns.

3.3.2 Refinement Process
3.3.2.1 Study of the Equilibration of the System

The refinement process was aimed at relaxing the crude models to capture specific structural features of the target protein, not present in the template. As described in the methods section, refinement was carried out using MD simulations of a system consisting of the target protein embedded in a POCP lipid bilayer. The final model of the target receptors were produced from the average structures computed of the last 50 ns of the respective production runs, followed by energy minimization of the

protein structure using an effective dielectric constant of 2 to mimic the protein environment. This process was carried out simply to eliminate possible crashes between atoms found in the average structure.

In order for the refinement process to be robust, simulation times require to be long enough to ensure that the system is equilibrated. System equilibration can be monitored by the time evolution of the rmsd of the successive trajectory snapshots in regard to the starting structure. Figure 3.4 shows the rmsd time evolution using the α-carbons of all residues for each of the simulations performed in the present work. As can be seen, equilibration takes more than 200 ns of simulation time. Comparison of the rmsd time evolution between the whole protein and the transmembrane domain suggests that there is a large contribution of the loops to the high rmsd values observed. Moreover, analysis of Figures 3.3a and 3.3b clearly shows that during the refinement process, the initial structures suffer a reorganization that is larger for those templates that are more distant from the target receptor in the phylogenetic tree, as expected. Also, figure 5 confirm this result. Thus, models of M3 constructed from the M2 muscarinic receptor exhibit a rmsd of ~2.3 Å, the model constructed from the histamine H1 receptor exhibit a rmsd of ~2.5Å and larger for the model constructed from rhodopsin that reaches ~3.0 Å. Another interesting point inferred from the analysis of Figures 3a and 3b is that those models refined with a ligand bound exhibit shorter equilibration times. This can be clearly seen by looking at the rmsd time evolution of the three models constructed using the M2 receptor as template. The faster equilibration must be due to the smaller flexibility of the system expected when a ligand is bound to the protein[48]. These results agree with the use of ligands to model active sites for in silico screening[49].
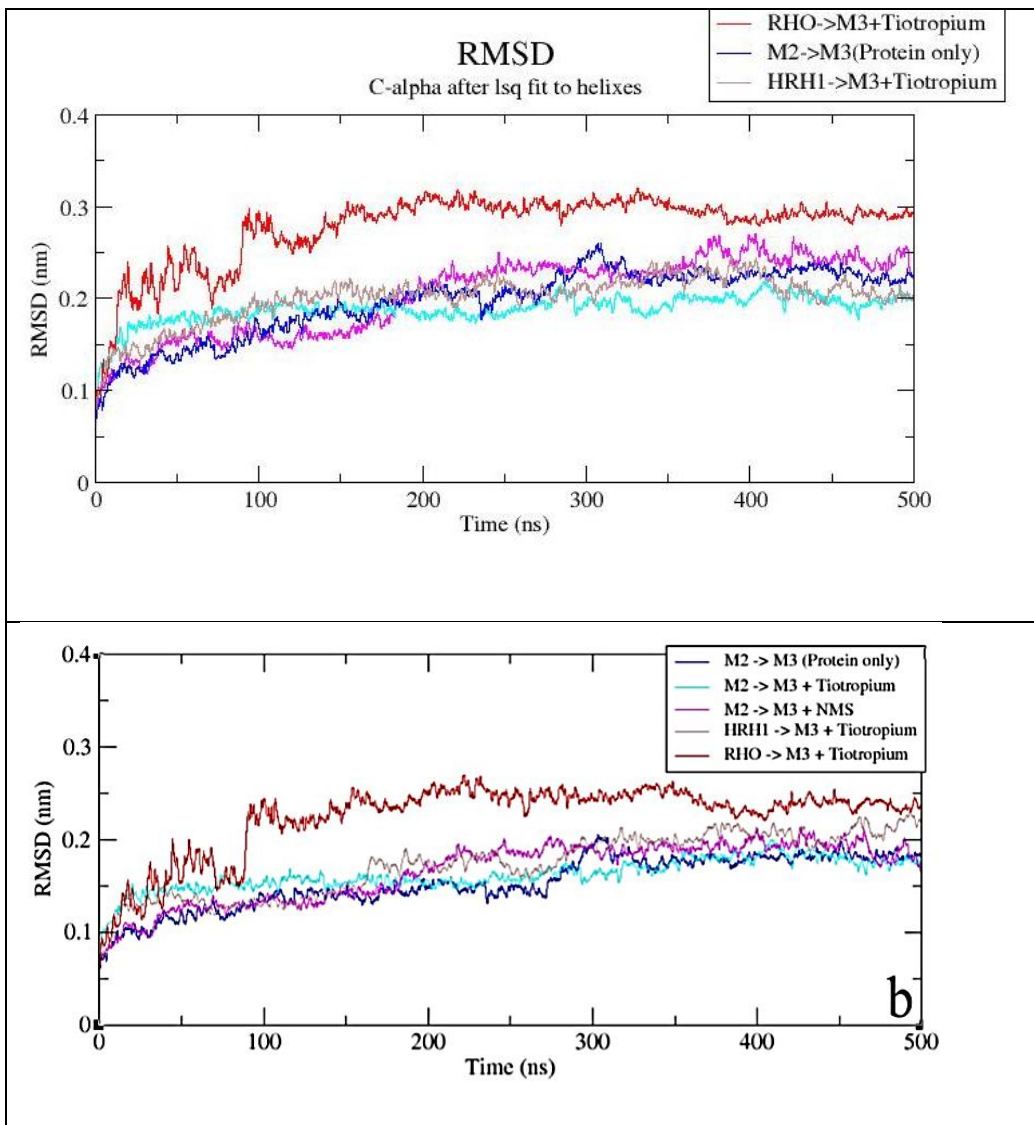
evolution of the equilibration using cα of the helices trough rmsd calculation a) the constructed model in this study b) the previously constructed models.
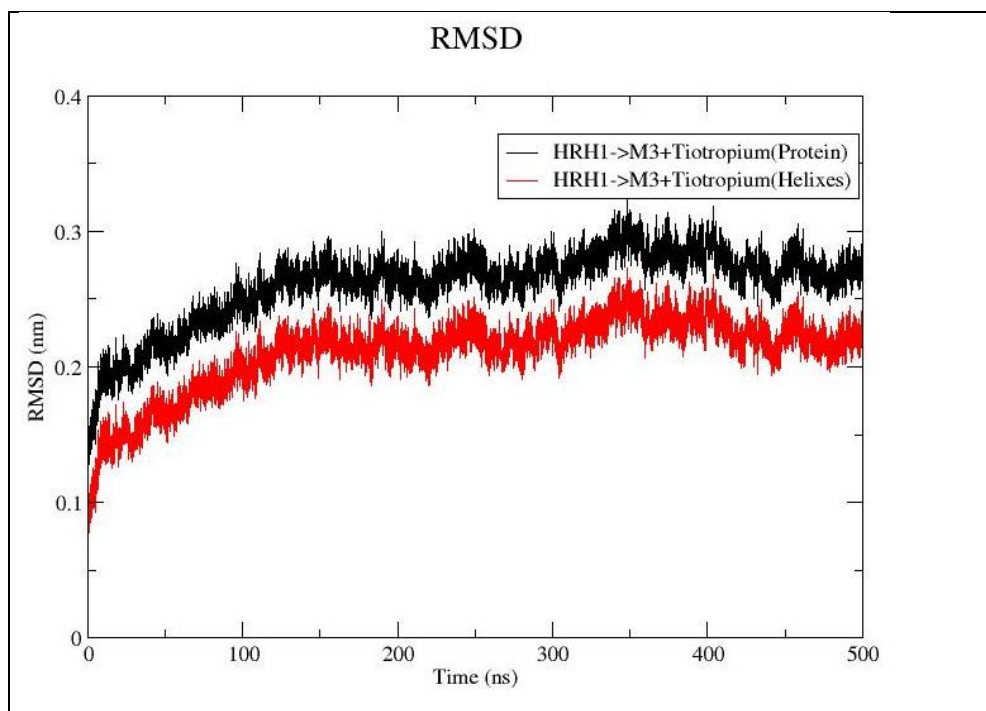
Figure 3. 4 *Time evolution of the equilibration using cα of the helices and protein of the recent study*

3.3.2.2 Analysis of the Refinement Process

As expected, the models constructed using the M2 muscarinic receptor as template, are closer to the target structure. Values of the rmsd of the diverse models constructed using the human muscarinic M2 receptor (using backbone Cα) are: 1.9 Å for the model refined with no ligand bound; 2.4 Å for the model constructed with tiotropium bound and 2.5 Å for the model constructed with NMS bound. However, the relative high rmsd values are due to the loops. The models are able to capture small rearrangements of the TM1 N-terminus segment; TM5 C-terminal segment, as well as TM7 middle segment towards the interior of the helix bundle, shown in the comparison of the M2 and M3 crystallographic structures[38]. Despite these rearrangements, there are certain features associated with the loops that are present in the starting models and are lost during the refinement process. Specifically, the pronounced outward bend at the extracellular end in TM4, found in the crystallographic structures of the M2 and M3 muscarinic receptors[50]. is preserved in the models constructed with tiotropium or NMS, but lost in the model constructed without any ligand.

The model of the muscarinic M3 constructed using the histamine H1 receptor bound to tiotropium yields an rmsd of 3.2 Å compared to the reference crystallographic structure. Comparison of the two structures reveals differences between the modelled and the crystallographic structure, being the largest difference found at the C-terminus of TM4. In this model both the ECL2 and ICL2 loops do not accommodate well to the space shown by the models constructed using the muscarinic M2 receptor,

although it can be seen certain improvement in regard the initial structures (see Figures 6 and 7). Regarding the length of the TM segments, the refinement process reduces slightly the differences found in the starting structure (see Table 3.1), but are not completely resolved. Specifically, TM1 is six residues shorter; TM2 and TM3 are one residue shorter; TM4 is in the final model four residues shorter catching up partly to the eight residues shorter of the starting model. Contrastingly, TM5 remains one residue longer while TM6 is seven residues longer correcting in part the eleven residues of the starting structure and TM8 remains one residue shorter.

The model of the muscarinic M3 receptor constructed using the rhodopsin bound to tiotropium yields a rmsd with the crystallographic structure of 4.79 Å similar to the starting model, suggesting that little improvement has been achieved during the refinement process. Comparison of the two structures reveals large differences between the modelled and the crystallographic structure. On the one hand, like in the model constructed using the histamine H1 receptor both the ECL2 and ICL2 loops do not accommodate well to the space shown by the models constructed using the muscarinic M2 receptor and only a small improvement can be seen in regard the initial structures (see Figures 3.5). In regard to the length of the TM segments the refinement process reduces slightly the differences found in the staring structure (see Table 3.1), but differences are still remarkable. Specifically, TM2 is one residue shorter; TM3 is four residues shorter; TM4 is thirteen residues shorter worsening the starting model; TM5 is remains one residue longer and TM6 is nine residues longer.

Finally, in order to understand the utility of the models generated for docking studies, we docked tiotropium onto the diverse models before and after being subjected to the refinement process. The crystallographic structure of tiotropium bound to the muscarinic M3 receptor shows the ligand sitting in a pocket conformed by helices TM3, MT5, TM6 and TM7. It is oriented in such a way that the ligand shows the tiophene groups in the proximity of TM5, whereas the quaternary nitrogen is flanked by TM3 and TM7. The ligand exhibits diverse interactions with the neighboring residues: the quaternary nitrogen sits close to Asp147; Asn507 forms two polar interactions with the carbonyl and with the hydroxyl groups of tiotropium, respectively; the epoxy group exhibits an interaction with the sidechain of Ser152 and one of the tiophene groups sits at interacting distance with Trp504[39]. All the models described in the present work are capable to reproduce the pose tiotropium adopts when bound to the M3 receptor. However, in regard to the diverse interactions with the sidechains of the neighboring residues, most of the initial structures fulfil only part the set of ligand-receptor interactions found in the crystal, however after refinement all the structures fulfil all the features of the crystallographic structure.
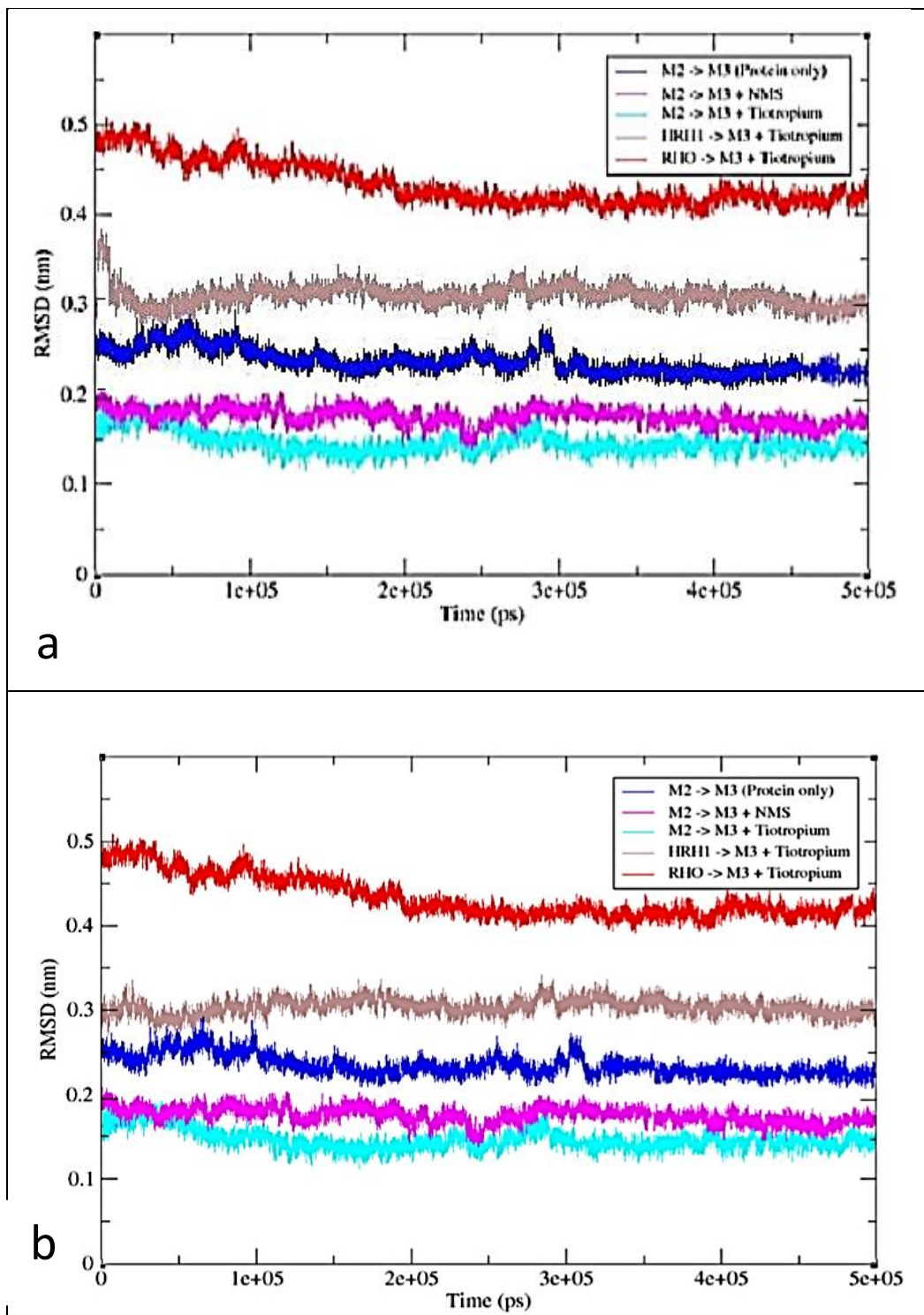
Figure 3. 5 *rmsd evaluation of the trajectories between TM regions of the refined model and the crystal structure and compare it with the previously results*

3.4 Conclusions

The goal of the present study was to assess the effect of selecting different templates as well as the conditions of the refinement process on the accuracy of the atomic resolution models produced of GPCRs by homology modeling. For this purpose, we constructed diverse models of the muscarinic M3 receptor which crystallographic structure is available with tiotropium bound, using different templates and refinement conditions. Specifically, we used as templates: the muscarinic M2 receptor (close in the phylogenetic tree), the histamine H1 receptor (at intermediate distance in the phylogenetic tree) and rhodopsin (distant in the phylogenetic tree). Refinement of the structures was carried out using molecular dynamics of the models constructed embedded in a lipid bilayer for about 500 ns. In order to understand the effect on the refinement of having a ligand bound to the receptor, we performed five different refinements. In the case of using the muscarinic M2 receptor as template, a model was refined without ligand, another one with tiotropium bound (the ligand bound to the muscarinic M3 crystallographic structure), and another one with N-methylscopolamine (a non-selective muscarinic antagonist) bound. The refinements using the histidine H1 and rhodopsin as templates were carried out with tiotropium bound.

The study reveals that these systems achieve equilibration after more than 250 ns of simulation, although if only the transmembrane region is considered equilibration is achieved much faster. Present results show that the use of molecular dynamics improves the quality of the homology models mostly on the transmembrane region. Specifically, the refinement process permits the correction of the length of the helices and improves the accuracy of the helix bundle. Despite the constructed models capture most of the features of the M3 receptor, the distance in the phylogenetic tree affects their quality, being the most accurate models those constructed using a template close to the target receptor. So that, the refinement process is not powerful enough to correct the problems associated when choosing a template distant from the target. Moreover, inclusion of a ligand on the modeling makes the refinement more robust since equilibration is observed faster.

3.5 References

1. Filmore, D. Cell-based screening assays and structural studies are fueling G-protein coupled receptors as one of the most popular classes of investigational drug targets. *Modern Drug Discovery.* **7**, 24-28 (2004).

2. Jacobson, K. A. & Costanzi, S. New Insights for Drug Design from the X-Ray Crystallographic Structures of G-Protein-Coupled Receptors. *Mol. Pharmacol.* **82**, 361–371 (2012).

3. Mason, J. S., Bortolato, A., Congreve, M. & Marshall, F. H. New insights from structural biology into the druggability of G protein-coupled receptors. *Trends Pharmacol. Sci.* **33**, 249–260 (2012).

4. Hauser, A. S., Attwood, M. M., Rask-Andersen, M., Schiöth, H. B. & Gloriam, D. E. Trends in GPCR drug discovery: new agents, targets and indications. *Nat. Rev. Drug Discov.* **16**, 829–842 (2017).

5. Lee, Y., Basith, S. & Choi, S. Recent Advances in Structure-Based Drug Design Targeting Class A G Protein-Coupled Receptors Utilizing Crystal Structures and Computational Simulations. *J. Med. Chem.* **61**, 1–46 (2018).

6. Portoghese, P. S. Relationships between stereostructure and pharmacological activities. *Annu. Rev. Pharmacol.* **10**, 51–76 (1970).

7. Latorraca, N. R., Venkatakrishnan, A. J. & Dror, R. O. GPCR Dynamics: Structures in Motion. *Chem. Rev.* **117**, 139–155 (2017).

8. The Nobel Prize in Chemistry 2012. Available at: https://www.nobelprize.org/prizes/chemistry/2012/summary/. (Accessed: 23rd September 2019)

9. GPCR-EXP for Experimental-solved and Predicted GPCR Structure. Available at: https://zhanglab.ccmb.med.umich.edu/GPCR-EXP/.

10. Ostermeier, C. & Michel, H. Crystallization of membrane proteins. *Curr. Opin. Struct. Biol.* **7**, 697–701 (1997).

11. Wakefield, A. E., Mason, J. S., Vajda, S. & Keserű, G. M. Analysis of tractable allosteric sites in G protein-coupled receptors. *Sci. Rep.* **9**, 6180 (2019).

12. Gentry, P. R., Sexton, P. M. & Christopoulos, A. Novel Allosteric Modulators of G Protein-coupled

Receptors. *J. Biol. Chem.* **290**, 19478–88 (2015).

13. Rajagopal, S., Rajagopal, K. & Lefkowitz, R. J. Teaching old receptors new tricks: biasing seven-transmembrane receptors. *Nat. Rev. Drug Discov.* **9**, 373–386 (2010).

14. Costanzi, S. Modeling G Protein-Coupled Receptors: a Concrete Possibility. *Chim. Oggi* **28**, 26–31 (2010).

15. Carlsson, J. *et al.* Ligand discovery from a dopamine D3 receptor homology model and crystal structure. *Nat. Chem. Biol.* **7**, 769–78 (2011).

16. Bissantz, C., Bernard, P., Hibert, M. & Rognan, D. Protein-based virtual screening of chemical databases. II. Are homology models of g-protein coupled receptors suitable targets? *Proteins Struct. Funct. Bioinforma.* **50**, 5–25 (2002).

17. Costanzi, S. Homology Modeling of Class A G Protein-Coupled Receptors. in *Methods in molecular biology (Clifton, N.J.)* **857**, 259–279 (2011).

18. Koehler Leman, J., Ulmschneider, M. B. & Gray, J. J. Computational modeling of membrane proteins. *Proteins Struct. Funct. Bioinforma.* **83**, 1–24 (2015).

19. Costanzi, S., Siegel, J., Tikhonova, I. & Jacobson, K. Rhodopsin and the Others: A Historical Perspective on Structural Studies of G Protein-Coupled Receptors. *Curr. Pharm. Des.* **15**, 3994–4002 (2009).

20. Kufareva, I. *et al.* Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. *Structure* **19**, 1108–26 (2011).

21. Katritch, V. & Abagyan, R. GPCR agonist binding revealed by modeling and crystallography. *Trends Pharmacol. Sci.* **32**, 637–643 (2011).

22. Simms, J. *et al.* Homology Modeling of GPCRs. *Methods Mol. Biol.* **552**, 97–113 (2009).

23. Beuming, T. & Sherman, W. Current Assessment of Docking into GPCR Crystal Structures and Homology Models: Successes, Challenges, and Guidelines. *J. Chem. Inf. Model.* **52**, 3263–3277 (2012).

24. Kufareva, I., Katritch, V., Stevens, R. C., Abagyan, R. & Abagyan, R. Advances in GPCR Modeling Evaluated by the GPCR Dock 2013 Assessment: Meeting New Challenges. *Structure* **22**, 1120–

1139 (2014).

25. Michino, M. *et al.* Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nat. Rev. Drug Discov.* **8**, 455–63 (2009).

26. Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739–45 (2000).

27. Martinez-Archundia, M., Cordomi, A., Garriga, P. & Perez, J. J. Molecular Modeling of the M3 Acetylcholine Muscarinic Receptor and Its Binding Site. *J. Biomed. Biotechnol.* **2012**, 1–12 (2012).

28. Becker, O. M., Shacham, S., Marantz, Y. & Noiman, S. Modeling the 3D structure of GPCRs: advances and application to drug discovery. *Curr. Opin. Drug Discov. Devel.* **6**, 353–61 (2003).

29. Archer, E., Maigret, B., Escrieut, C., Pradayrol, L. & Fourmy, D. Rhodopsin crystal: new template yielding realistic models of G-protein-coupled receptors? *Trends Pharmacol. Sci.* **24**, 36–40 (2003).

30. Bywater, R. P. Location and nature of the residues important for ligand recognition in G-protein coupled receptors. *J. Mol. Recognit.* **18**, 60–72 (2005).

31. Forrest, L. R., Tang, C. L. & Honig, B. On the Accuracy of Homology Modeling and Sequence Alignment Methods Applied to Membrane Proteins. *Biophys. J.* **91**, 508–517 (2006).

32. Liu, T., Tang, G. W. & Capriotti, E. Comparative modeling: the state of the art and protein drug target structure prediction. *Comb. Chem. High Throughput Screen.* **14**, 532–47 (2011).

33. Worth, C. L., Kleinau, G. & Krause, G. Comparative Sequence and Structural Analyses of G-Protein-Coupled Receptor Crystal Structures and Implications for Molecular Models. *PLoS One* **4**, e7011 (2009).

34. Latek, D., Pasznik, P., Carlomagno, T. & Filipek, S. Towards Improved Quality of GPCR Models by Usage of Multiple Templates and Profile-Profile Comparison. *PLoS One* **8**, e56742 (2013).

35. Goldfeld, D. A., Zhu, K., Beuming, T. & Friesner, R. A. Loop prediction for a GPCR homology model: Algorithms and results. *Proteins Struct. Funct. Bioinforma.* **81**, 214–228 (2013).

36. Kimura, S. R., Tebben, A. J. & Langley, D. R. Expanding GPCR homology model binding sites via a balloon potential: A molecular dynamics refinement approach. *Proteins Struct. Funct. Bioinforma.*

71, 1919–1929 (2008).

37.   Zhang, Y., Sham, Y. Y., Rajamani, R., Gao, J. & Portoghese, P. S. Homology Modeling and Molecular Dynamics Simulations of the Mu Opioid Receptor in a Membrane-Aqueous System. *ChemBioChem* **6**, 853–859 (2005).

38.   Platania, C. B. M., Salomone, S., Leggio, G. M., Drago, F. & Bucolo, C. Homology Modeling of Dopamine D2 and D3 Receptors: Molecular Dynamics Refinement and Docking Evaluation. *PLoS One* **7**, e44316 (2012).

39.   Kruse, A. C. *et al.* Structure and dynamics of the M3 muscarinic acetylcholine receptor. *Nature* **482**, 552–556 (2012).

40.   Lupala, C. S., Rasaeifar, B., Gomez-Gutierrez, P. & Perez, J. J. Using molecular dynamics for the refinement of atomistic models of GPCRs by homology modeling. *J. Biomol. Struct. Dyn.* **36**, 2436–2448 (2018).

41.   Lupala, C. S., Rasaeifar, B., Gomez-Gutierrez, P. & Perez, J. J. 193 Effect of template selection on the construction of atomistic models of GPCRs by homology modeling. *J. Biomol. Struct. Dyn.* **33**, 127–128 (2015).

42.   Nayeem, A., Sitkoff, D. & Krystek, S. A comparative study of available software for high-accuracy homology modeling: From sequence alignments to structural models. *Protein Sci.* **15**, 808–824 (2006).

43.   Cavasotto, C. N. & Palomba, D. Expanding the horizons of G protein-coupled receptor structure-based ligand discovery and optimization using homology models. *Chem. Commun.* **51**, 13576–13594 (2015).

44.   Molecular Operating Environment (MOE). Chemical Computing Group Inc. 1010 Sherbooke St. West Suite #910 Montreal QC Canada H3A 2R7.

45.   Haga, K. *et al.* Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature* **482**, 547–551 (2012).

46.   Shimamura, T. *et al.* Structure of the human histamine H1 receptor complex with doxepin. *Nature* **475**, 65–70 (2011).

47.   Friesner, R. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method

and Assessment of Docking Accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).

48.   Lin, S., Gether, U. & Kobilka, B. K. Ligand stabilization of the beta 2 adrenergic receptor: effect of DTT on receptor conformation monitored by circular dichroism and fluorescence spectroscopy. *Biochemistry* **35**, 14445–51 (1996).

49.   Evers, A., Gohlke, H. & Klebe, G. Ligand-supported Homology Modelling of Protein Binding-sites using Knowledge-based Potentials. *J. Mol. Biol.* **334**, 327–345 (2003).

50.   Kow, R. L. & Nathanson, N. M. Muscarinic receptors become crystal clear. *Nature* **482**, 480–481 (2012).

CHAPTER 4: Molecular Features Characterizing Non-peptide Selectivity to the Human B1 and B2 Bradykinin Receptors

4.1 Introduction

Bradykinin (BK) is a nonapeptide, member of the Kinin family with sequence Arg1-Pro2-Pro3-Gly4-Phe5-Ser6-Pro7-Phe8-Arg9. The peptide plays a key role in many pathophysiological situations such as anaphylaxis, arthritis, septic, and hemorrhagic shock inflammatory bowel disease, rhinitis, asthma [1,2]. Due to the importance of bradykinin in these pathologies, a great effort has been invested in finding BK antagonists in the past.

Bradykinin can activate two G-protein coupled receptors known as B1 and B2. Expression of B1 occurs during inflammation episodes and tissue trauma, while the B2 receptor is expressed constitutively in multiple types of cells, so that physiological actions of bradykinin are mostly mediated through the B2 bradykinin receptor[1–3]. Accordingly, most of the previous research carried out on bradykinin has been focused on finding new antagonists of the B2 receptor [4]. Fruit of these investigations, Icatibant was the first peptide antagonist released in the market[5,6]. Other small molecule B2 antagonists have also been reported since then[7]. On the other hand, although the B1 receptor is expressed during chronic and inflammatory pain[2,8,9], efforts of finding small molecule B1 antagonist has been increased in the last years, however despite these efforts still there is no drug in the market that can be used as B1 antagonist.

BK was synthesized for the first time in the 60s [10] together with diverse analogs with agonistic activity. These studies permitted identification of the structure of the ligand with 9-amino acid peptide chain(H-Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg-OH) . The first analog with antagonistic activity was reported in 1985, when Pro7 was replaced by an aromatic D-amino acid[12]. This discovery led to the production of a first generation of the BK antagonists such as D-Arg-[Hyp^3,D-Phe^7]-BK (NPC-567) (Hyp=hydroxyproline); D-Arg-[Hyp^3,Thi^5,8, D-Phe^7]-BK (NPC-349) (Thi=thienylala-nine); or D-Arg-[Hyp^3, D-Phe^7, Leu^8]-BK [13,14]. This first generation of antagonists provided an opportunity to understand the potential of BK antagonism for therapeutic intervention. However, these compounds exhibited low affinity for the B2 receptor in comparison to BK itself and were not selective for the B2. so that some refinement was necessary to get molecules to be used as therapeutic agents. Furthermore, it was also interesting to realize that removal of the C-terminal arginine decreased the affinity of the analogs for B2, producing more selective B1 receptor ligands[11].

Refinement of the pharmacological profile of BK antagonists was achieved by means of NMR spectroscopy combined with molecular modeling studies[15,16] . These studies showed the tendency of the C-terminus of bradykinin and active analogs to adopt a β-turn. Subsequently, it was assumed that this conformation was the conformation the peptide adopts when binds to its receptors. Accordingly, new

analogs conformationally constrained at the C-terminus were designed to adopt a β-turn secondary structure, giving rise to a second generation of BK antagonists. Potent analogs of this generation are Icatibant (HOE140) with sequence: D-Arg0-[Hyp3, Thi5, D-Tic7,Oic8]-BK(Tic = tetrahydroisoquinoline; Oic = octahydroindolecarboxylic acid)[5] or NPC17731 with the sequence: D-Arg-[Hyp3, D-HypE(trans-proyl)7, Oic8]-BK)[17]. Subsequent studies carried out to find shorter peptide sequences preserving activity, showed that in addition to the β-turn conformation at the C-terminus, other features of the molecule were also important for high-affinity antagonism. Specifically, the Arg at the N-terminus was shown to be necessary to interact with negatively charged residues such as Asp266 and Asp284 located at the mouth of the receptor[18]. A potent ligand containing all these features is the high-affinity ligand MEN11270 with the sequence D-Arg0-Arg1-Pro2-Hyp3-Gly4-Thi5-cyclo[Dab6-D-Tic7-Oic8-Arg9] (Dab = diaminobutyric acid) which a cyclic structure at the C-terminus designed to force the β-turn structure while the N-terminal sequence is that of Icatibant[19].

The second generation of B2 antagonists showed a better pharmacological profile in comparison to the first generation. Specifically, these analogs are high-affinity and selective for the B2 receptor and exhibit a better pharmacokinetic profile than the first generation compounds because they have higher resistance to enzymatic degradation. In contrast, the drawback they have is their limited oral bioavailability. For example, Icatibant from this generation should be used via injection which is a drug for acute attacks of hereditary angioedema treatment of the adults with deficiency of C1-esterase inhibitor[20].

In the 90s and the first years of the twenty-one century, aimed at improving their oral bioavailability a third generation of antagonists was designed[7,21]. Examples of this generation include WIN64338[22], FR173657[23], bradyzide[24], anatibant[25] or fastibant[26]. These antagonists exhibit high-affinity and selectivity for the B2 receptor, although their bioavailability is still limited due to their high molecular mass (500-600 Da). In order to overcome this drawback, in a step forward, some efforts were put forward to find molecules with lower molecular mass, leading to the discovery of JSM10299 as a potent selective antagonist with high affinity for the B2 receptor which had less molecular mass in comparison to the other analogs of the third generation[27].

4.2 Aim of the study

The features that small molecules must fulfill to be BK antagonists was a topic undertaken in a previous report of this group[28,29]. This can be summarized in the definition of pharmacophore models for B1 and B2 antagonism, shown in Figure 4.1. These models were derived from a modeling study that involved docking of known small molecule BK selective antagonists to the B1 and B2 receptors. However,

since there is no crystal structure of these receptors yet available, atomistic models of human B1 and B2 receptors were first constructed by homology modeling[28,29]. In a second step, these models were used to perform docking studies with some of the small molecules described in the literature [7]. Analysis of the results showed the binding manner of selective and non-selective small molecule ligands to each one of these two receptors confirming the results of previous studies that their binding site can be different from that of peptide ligands[30]. In a step forward, pharmacophore models of human B1 and B2 antagonism were validated by a *in silico* screening study that led to discovery of some novel small molecules with antagonist activity[31,32].

Comparison of the two pharmacophores for the B1 and B2 receptors can provide useful information about specific molecular features of the small molecules that make them selective for each one of these receptors. Accordingly, the aim of the present study is to describe the features that make bradykinin ligands to be selective. The results of this study provided us a hypothesis that was later validated through a *in silico* screening study.

4.3 Results and discussion

Figure 4.1 shows the pharmacophores of the BK B1 and B2 receptor antagonism together with the information regarding the residues involved deduced from the receptors constructed in a previous study by homology modeling[31,32].
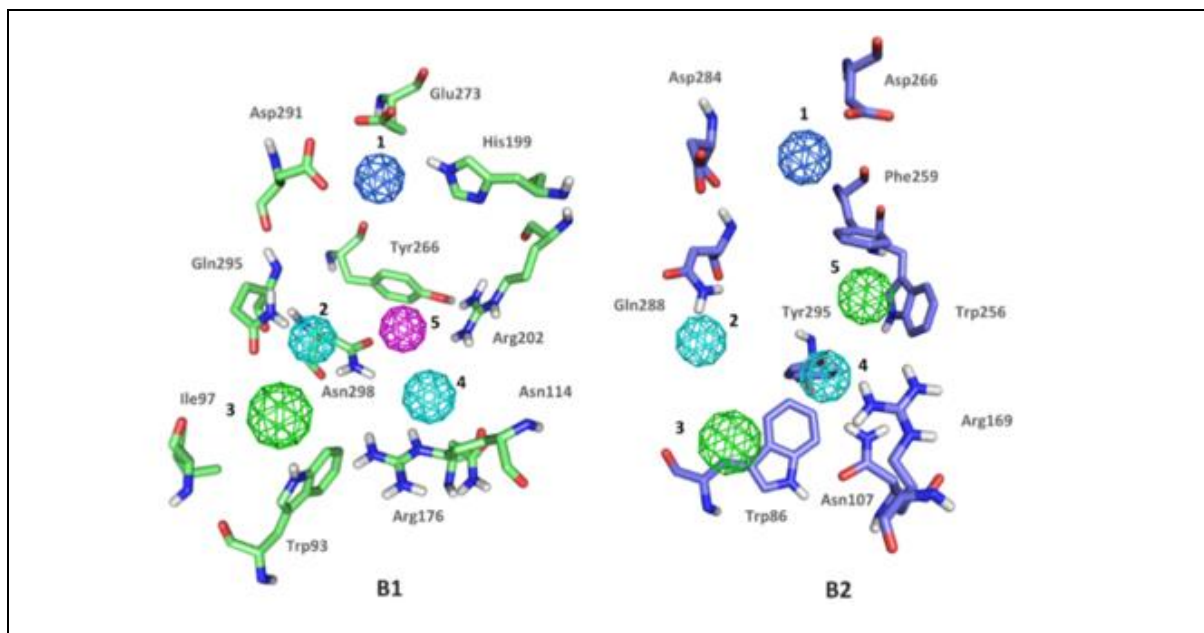
Figure 4. 1 *Pharmacophore hypothesis of the BK B1 and B2 receptors and residues defining each pharmacophore points according to their 3D structures.*

Simple inspection of the pharmacophore points and residues suggests that there is high residue identity in the binding pocket of both receptors, as anticipated from the high sequence identity of the aligned sequences (Figure 4.2). Moreover, comparison of the two pharmacophores suggests that pharmacophoric points #1-4 are common in the two pharmacophores, while the only noticeable difference regards pharmacophoric point #5. This difference arises from the difference in nature and chemical features of the residues defining point #5 in each of the two receptors. Thus, as pharmacophore points#1-4 are common in both of the receptors, the selectivity of the bradykinin antagonist is responsibility of the pharmacophore point#5. Furthermore, selective ligands should fulfill at least four pharmacophore points, being one of them point#5[31,32]. Furthermore, analysis of the Figure 4.1 permits to identify that the pharmacophore point #5 in the human B1 receptor is a polar moiety, defined by residues Arg202(R5.38), Tyr266(Y6.51) and Asn298(N7.39) (the Ballesteros-Weinstein notation[33] is written in parenthesis). In contrast, pharmacophoric point#5 in B2 is defined by an aromatic/hydrophobic moiety where ligands interact with hydrophobic residues such as Trp256(W648), Phe259(F6.51) and Tyr295(Y7.43). This information is also deisplayed in Table 4.1.
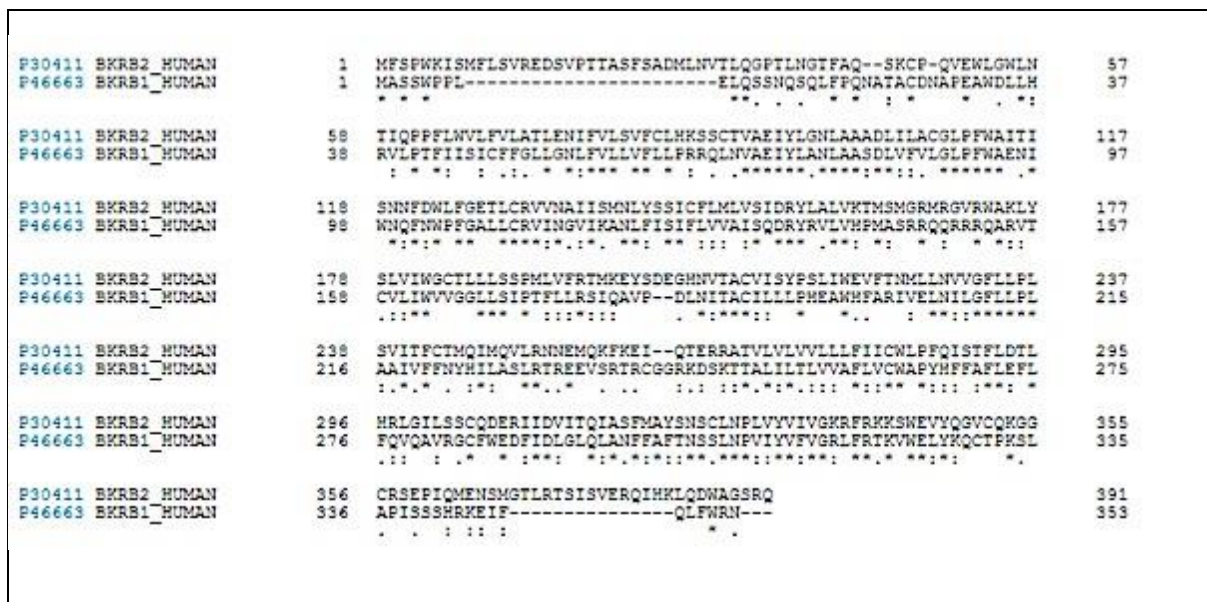
Figure 4. 2 *Sequence alignment of the human B1 and B2 receptors.*

We then proceeded to the superposition of the refined models of B2 and B1 and displayed the solvent accessible surface defining the orthosteric site as shown in Figure 4.3. Figure 4.3a shows the binding pocket of the B1 receptor with a ligand/compound bound and the pharmacophoric points are also shown explicitly. These surfaces permit to understand the relatives sizes of the binding pockets and different position of pharmacophoric point #5 in each of the two receptors.
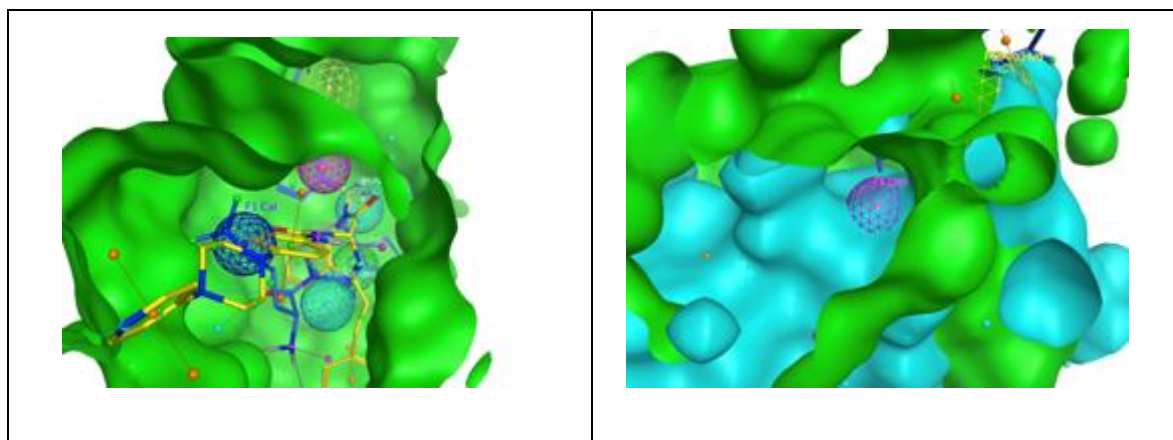
Figure 4. 3 *a) surface and pharmacophore points of B1 and B2   b) point 5 pharmacophores of both ligands*

As can be seen in Figure 4.3b, the pocket of the receptor B1 (in green) is smaller than the corresponding pocket of B2 (blue). Accordingly, the smaller volume of the B1 binding pocket prevents ligands selective to the B2 receptor to bind to the B1 receptor. (See Table 3.1 and Figure 4.3).

Table 4. 1 *Description of the pharmacophoric points of the B1 and B2 receptors together with the residues involved*

Pharmacophore	Type	B1	B2
#1	Cationic Atom	ASP291	ASP284
#2	Donor and Acceptor	GLN295	GLN288
#3	Hydrophobic Centroid	TRP93	TRP86
#4	Donor and Acceptor	ASN114 ARG176	ASN107 ARG169
#5 B1	Donor	ARG202 ASN298 TYR266	
#5 B2	Aromatic/ Hydrophobic		TRP256 PHE259 TYR259

A more detailed analysis of the two receptors permits to understand that a key residue establishing the differential nature and location of the pharmacophoric point#5 is residue Arg202(R5.38) in B1, which has Thr197(T5.38) in B2 as counterpart. In BK B1 receptor Arg202(R5.38) points towards the center of the binding pocket and interacts with Tyr266(Y6.51) via a hydrogen bond interaction that together with Asn289(N7.39) define pharmacophoric point #5. Moreover, this interaction prevents ligands to explore deeper parts of the receptor. In contrast, in B2 residues Thr197(T5.38) and Phe259(F6.51) (counterpart of

Tyr266(Y6.51)) are not able to produce such an interaction, making the binding pocket deeper. In this situation ligands have access to the deeper region surrounded with the aromatic residues such as TRP25(W6.48), Phe259(F6.51) and Tyr295(Y7.43) that define pharmacophoric point #5 in the B2 receptor.

Another difference between two receptors corresponds to different features of Lys118 (K3.33) in B1 receptor in regard to its counterpart, Ser111 (S3.33) in B2, both located in the same region in transmembrane helix 3 (TM3). In the B1 receptor, the side chain of Lys118 (K3.33) forms a hydrogen bond with the side chain of Glu205 (E5.41). In contrast, in the B2 receptor the counterpart of Glu205 (E5.41) is Leu207 (L5.41). Thus, in B1 there is a charge-charge interaction between TM3 and TM5 not observed in B2. The lack of this interaction is responsible for TM5 to appear in B1 displaced in regard to the corresponding helix, reducing the solvent-accessible surface area of the receptor. Interestingly, Lys118 (K3.33) in B1 is key to explain the selectivity profile of the diverse kinins[34]. Recent Solid State NMR studies show evidence that Lys118 (K3.33) interacts with the carboxyl C-terminal group of the B1 selective analogs desArg9-BK and desArg9-KD and forces the C-terminus of BK and KD to adopt a distinct orientation to avoid a repulsive interaction with the Arg9 side chain of the peptides[35](Figure 4.4).
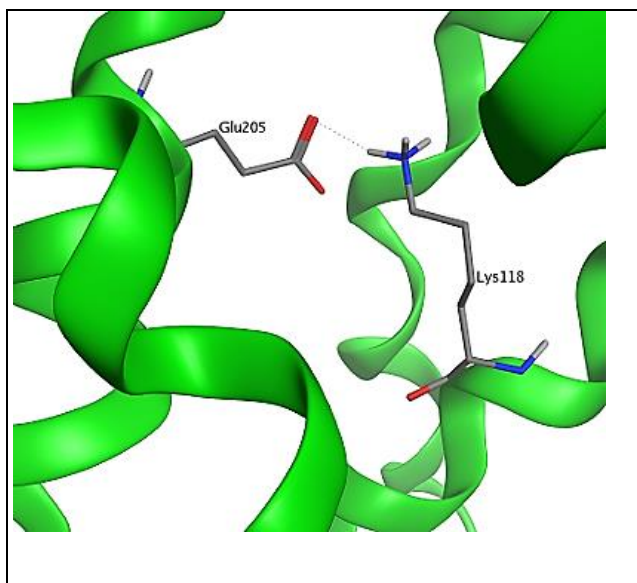
Figure 4. 4 Hydrogen bond interaction of Lys118(K3.33) with G205(E5.41) in B1 receptor

To investigate the robustness of this hypothesis we analyzed the selectivity profile of a set of diverse hits previously identified in a *in silico* screening[31,32]. The chemical structure of the six compounds selected is shown in Figure 4.5 and their antagonist profile listed in Table 4.2.

Table 4. 2 pharmacologic antagonistic profile and fulfillment of the pharmacophore B1 and B2 points by a set of previously in silico discovered molecules[31,32].

compound	B1 inhibition	B2 inhibition	Point#1	Point#2	Point#3	Point#4	Point#5 @B1	Point#5 @B2
#1	31%@50uM	36%@50uM	✓		✓			
#2	46%@50uM	65%@50uM	✓		✓			
#3	31%@10uM	41%@50uM	✓	✓	✓			
#4	32%@10uM	34%@10uM	✓	✓		✓		
#5	14%@10uM	33%@10uM			✓	✓	✓	✓
#6	12%@50uM	45%@50uM	✓		✓			

Figure 4. 5 *chemical structure of discovered molecules by in silico screening* [31,32].

Hits were docked onto the constructed models of the human B1 and B2 receptors, to evaluate the pharmacophoric points that each of them fulfill. The results are also shown in Table 4.2. Analysis of the results suggest that the antagonist activity of these hits correlates well with the number of pharmacophoric points fulfilled. The lack of selectivity of most of the compounds can be explaining duw to fulfilling pharmacophoric points #1-4. However, there is an interesting result regarding compound #5. This compound is non-selective, but fulfills point#5 of both receptors. Evaluation of the binding manner of compound #5 by docking studies shows that this compound is small enough to fulfil pharmacophore point#5 of each receptor because it is capable to bind in a different pose in each of the two receptors (Figure 4.6). Specifically, the carboxyl moiety of this ligand interacts with Arg202(R5.38) (Figure 4.6a) when bound to the B1 receptor, fulfilling pharmacophore point#5, while in B2 the carboxyl moiety faces to Arg169(R3.57) (Figure 4.6b) of the receptor fulfilling point #4. Moreover, in B1 it also fulfills point #3 and #4 of the pharmacophore, while in B2 it also fulfills points #3 and #5. This result suggests that to define selectivity in addition to fulfill pharmacophore point #5, is it necessary that the ligand is bulky enough to prevent binding in different poses to each one of the bradykinin receptors. This result suggests that

fulfilling pharmacophore point#5 is a necessary condition, but is not enough to define selectivity of the ligand.
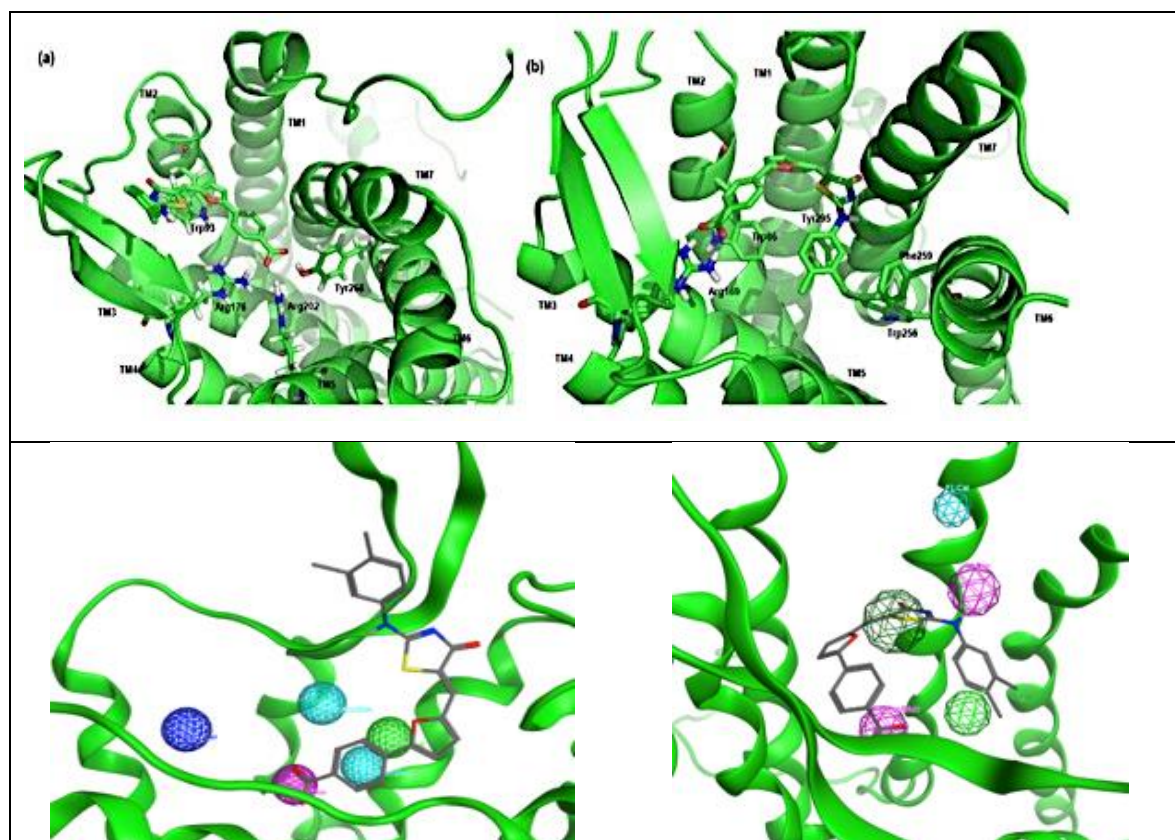
Figure 4. 6 *a) prospective bound conform and fulfilled pharmacophore point of compound #5 for B1 receptor b) prospective bound conform and fulfilled pharmacophore point of compound #5 for B2 receptor*

Moreover, the previous docking results[31,32] were compared with the results of this study. This shows that in the previous study most of the active compounds fulfil at least four points of the pharmacophores while docking of the novel hits in this study shows that the most active hits just fulfil three points of the pharmacophores. This result suggests that selectivity of compounds depends on fulfilling at least four points of the pharmacophores which one of them should be pharmacophore point #5[36].

4.4 Conclusion

In this study, a comparison was done between constructed models of human B1 and B2 receptors. Analysis of the results provides some structural features corresponding to the selectivity of the ligands for these receptors. comparison of the pharmacophores shows that point #1-4 are in common in both of the receptors, but point#5 was different in each one of them which comes from the difference in nature and

location that define selectivity of the ligand. This difference in nature and location of point #5 corresponded to the different position of Arg202(R5.38) in B1 and its counterpart Thr197(T5.38) in B2. which this hypothesis was confirmed by the analysis of the set of non-selective small molecules defined previously by *in silico* study using pharmacophore points of both of the receptors[31,32]. Among all of the compounds, none of them fulfills pharmacophore points #1-4 that can be describes why they are non-selective. However, compound #5 showed an interesting result .it fulfills pharmacophore point #5 in both of the receptors because it is as small as can binds in different modes to the receptors. This result helps us to conclude that fulfillment of the point#5 pharmacophore is a necessary condition, but is not enough. more than this condition, ligand should be bulky enough to prevent binding of that in different modes and it can be provided by fulfillment at least four points of the pharmacophore.

4.5 References

1. Regoli, D. & Barabé, J. Pharmacology of bradykinin and related kinins. *Pharmacol. Rev.* **32**, 1–46 (1980).

2. Leeb-Lundberg, L. M. F., Marceau, F., Müller-Esterl, W., Pettibone, D. J. & Zuraw, B. L. International Union of Pharmacology. XLV. Classification of the Kinin Receptor Family: from Molecular Mechanisms to Pathophysiological Consequences. *Pharmacol. Rev.* **57**, 27–77 (2005).

3. Hall, J. M. Bradykinin receptors: pharmacological properties and biological roles. *Pharmacol. Ther.* **56**, 131–90 (1992).

4. Stewart, J. M. Bradykinin antagonists: Development and applications. *Biopolymers* **37**, 143–155 (1995).

5. Hock, F. J. *et al.* Hoe 140 a new potent and long acting bradykinin-antagonist: in vitro studies. *Br. J. Pharmacol.* **102**, 769–73 (1991).

6. Cole, S. W. & Lundquist, L. M. Icatibant for the Treatment of Hereditary Angioedema. *Ann. Pharmacother.* **47**, 49–55 (2013).

7. Dziadulewicz, E. K. Non-peptide ligands for bradykinin receptors 1995 – 2004. *Expert Opin. Ther. Pat.* **15**, 829–859 (2005).

8. Marceau, F. Kinin B1 receptors: a review. *Immunopharmacology* **30**, 1–26 (1995).

9. Ahluwalia, A. & Perretti, M. B1 receptors as a new inflammatory target. Could this B the 1? *Trends Pharmacol. Sci.* **20**, 100–4 (1999).

10. Boissonnas, R. A., Guttmann, S. & Jaquenoud, P.-A. Synthèse de la L -arginyl- L -prolyl- L -prolyl- glycyl- L -phénylalanyl- L -Séryl- L -prolyl- L -phénylalanyl- L -arginine, un nonapeptide présentant les propriétés de la bradykinine. *Helv. Chim. Acta* **43**, 1349–1358 (1960).

11. Skidgel, R. A., and Erdös, E.G. Enzymatic breakdown of bradykinin. *Pro-inflammatory Anti-inflamm. Pept.* Marcel Dekker. New York, NY. 1998. pp. 459–476.

12. Vavrek, R. J. & Stewart, J. M. Competitive antagonists of bradykinin. *Peptides* **6**, 161–4(1985).

13. Steranka, L. R., Farmer, S. G. & Burch, R. M. Antagonists of B2 bradykinin receptors. *FASEB J.* **3**, 2019–25 (1989).

14.   Stewart, J. M. Bradykinin antagonists: discovery and development. *Peptides* **25**, 527–532 (2004).

15.   Kyle, D. J. *et al.* NMR and computational evidence that high-affinity bradykinin receptor antagonists adopt C-terminal β-turns. *J. Med. Chem.* **36**, 1450–1460 (1993).

16.   Lopez, J. J. *et al.* The Structure of the Neuropeptide Bradykinin Bound to the Human G-Protein Coupled Receptor Bradykinin B2 as Determined by Solid-State NMR Spectroscopy. *Angew. Chemie Int. Ed.* **47**, 1668–1671 (2008).

17.   Kyle D J, Burch R. M. Recent advances toward novel bradykinin antagonists. *Drugs Future.* **17**, 305–312 (1992).

18.   Kyle, D. J., Chakravarty, S., Sinsko, J. A. & Stormann, T. M. A Proposed Model of Bradykinin Bound to the Rat B2 Receptor and Its Utility for Drug Design. *J. Med. Chem.* **37**, 1347–1354 (1994).

19.   Meini, S. *et al.* MEN 11270, A novel selective constrained peptide antagonist with high affinity at the human B2 kinin receptor. *J. Pharmacol. Exp. Ther.* **289**, 1250–6 (1999).

20.   Dubois, E. A. & Cohen, A. F. Icatibant. *Br. J. Clin. Pharmacol.* **69**, 425–6 (2010).

21.   Heitsch, H. Non-peptide Antagonists and Agonists of the Bradykinin B2 Receptor. *Curr. Med. Chem.* **9**, 913–928 (2005).

22.   Sawutz, D. G. *et al.* The nonpeptide WIN 64338 is a bradykinin B2 receptor antagonist. *Proc. Natl. Acad. Sci. USA.* **91**, 4693–7 (1994).

23.   Asano, M. *et al.* The identification of an orally active, nonpeptide bradykinin B_2 receptor antagonist, FR173657. *Br. J. Pharmacol.* **120**, 617–624 (1997).

24.   Burgess, G. M. *et al.* Bradyzide, a potent non-peptide B_2 bradykinin receptor antagonist with long-lasting oral activity in animal models of inflammatory hyperalgesia. *Br. J. Pharmacol.* **129**, 77–86 (2000).

25.   Pruneau, D. *et al.* Pharmacological profile of LF 16-0687, a new potent non-peptide bradykinin B2 receptor antagonist. *Immunopharmacol.* **43**, 187–94 (1999).

26.   Cucchi, P. *et al.* MEN16132, a novel potent and selective nonpeptide antagonist for the human bradykinin B2 receptor. In vitro pharmacology and molecular characterization. *Eur. J. Pharmacol.* **528**, 7–16 (2005).

27.   Dziadulewicz, E. K. *et al.* The design of non-peptide human bradykinin B2 receptor antagonists employing the benzodiazepine peptidomimetic scaffold. *Bioorg. Med. Chem. Lett.* **9**, 463–8 (1999).

28.   Cordomí, A., Edholm, O. & Perez, J. J. Effect of different treatments of long-range interactions and sampling conditions in molecular dynamic simulations of rhodopsin embedded in a dipalmitoyl phosphatidylcholine bilayer. *J. Comput. Chem.* **28**, 1017–1030 (2007).

29.   Lupala, C. S., Rasaeifar, B., Gomez-Gutierrez, P. & Perez, J. J. Using molecular dynamics for the refinement of atomistic models of GPCRs by homology modeling. *J. Biomol. Struct. Dyn.* **36**, 2436–2448 (2018).

30.   Meini, S. *et al.* Site-directed mutagenesis at the human B2 receptor and molecular modelling to define the pharmacophore of non-peptide bradykinin receptor antagonists. *Biochem. Pharmacol.* **67**, 601–9 (2004).

31.   Lupala, C. S., Gomez-Gutierrez, P. & Perez, J. J. New insights into the stereochemical requirements of the bradykinin B1 receptor antagonists binding. *J. Mol. Graph. Model.* **68**, 184–196 (2016).

32.   Lupala, C. S., Gomez-Gutierrez, P. & Perez, J. J. New insights into the stereochemical requirements of the bradykinin B2 receptor antagonists binding. *J. Comput. Aided. Mol. Des.* **30**, 85–101 (2016).

33.   Ballesteros, J. A. & Weinstein, H. [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* **25**, 366–428 (1995).

34.   Fathy, D. B., Mathis, S. A., Leeb, T. & Leeb-Lundberg, L. M. A single position in the third transmembrane domains of the human B1 and B2 bradykinin receptors is adjacent to and discriminates between the C-terminal residues of subtype-selective ligands. *J. Biol. Chem.* **273**, 12210–8 (1998).

35.   Joedicke, L. *et al.* The molecular basis of subtype selectivity of human kinin G-protein-coupled receptors. *Nat. Chem. Biol.* **14**, 284–290 (2018).

36.   Rasaeifar, B., Lupala, C. S., Gomez-Gutierrez, P. & Perez, J. J. Molecular features characterizing

non-peptide selectivity to the human B1 and B2 bradykinin receptors. *Bioorganic Med. Chem. Lett.* **29**, 11–14 (2019).

CHAPTER 5: New Insights into the Sterochemical Requirements of the Bombesin Receptors Antagonists Binding

5.1 Introduction

Bombesin is a tetradecapeptide with the sequence: $Glp^1-Gln^2-Arg^3-Leu^4-Gly^5-Asn^6-Gln^7-Trp^8-Ala^9-Val^{10}-Gly^{11}-His^{12}-Leu^{13}-Met^{14}-NH_2$ (where Glp = pyroglutamic acid), originally isolated from the skin of the European frog *Bombina bombina*[1]. After its discovery, other peptides with high sequence identity like ranatensin, alytensin, phyllolitorin, or litorin among others, were also characterized, constituting the bombesin-like family of peptides[2]. Despite these peptides were originally isolated from the skin of diverse amphibians, it was later found that they are also widely distributed in mammals[3]. Thus, two specific bombesin-like peptides have so far been isolated in mammals: Neuromedin B (NMB) with a sequence: $Gly^1-Asn^2-Leu^3-Trp^4-Ala^5-Thr^6-Gly^7-His^8-Phe^9-Met^{10}-NH_2$ [4] and the Gastrin-releasing peptide (GRP)[5], a 27 residue long peptide or its short version, the gastrin-releasing peptide(18-27), also known as Neuromedin C (NMC) with the sequence: $Gly^1-Asn^2-His^3-Trp^4-Ala^5-Val^6-Gly^7-His^8-Leu^9-Met^{10}-NH_2$.

Members of the Bombesin-like family of peptides are involved in a wide spectrum of biological activities in the central nervous system including satiety, control of circadian rhythm, thermoregulation and in peripheral tissues, stimulation of gastrointestinal hormone release, activation of macrophages, and effects on development[6]. Moreover, they play a role in the control of cellular proliferation[7]. Actions of this family of peptides are mediated through three GPCRs: The neuromedin B receptor (BB1R), the gastrin-releasing peptide receptor (BB2R), and the orphan, bombesin receptor subtype 3 (BB3R)[6]. Natural occurring members of the bombesin-like family bind to the first two receptors, being no endogenous ligand identified of BB3R identified so far, so the receptor is classified as orphan. NMB and NMC are the mammal endogenous ligands for BB1R and BB2R respectively[8]. Specifically, they are selective agonists for both receptor binding to them with high affinity: NMB binds to BB1R receptor with the affinity of K_i=0.052nM and exhibits 100 times higher affinity for BB2R, whereas GRP exhibits a Ki=0.19nM to the BB2R and approximately 1000 times higher for the BB1R[9]. However, a few peptides and non-peptide selective agonist has been discovered for orphan BB3R[9].

Due to the wide spectrum of biological activities mediated by bombesin receptors, there is considerable interest in the clinical potential of novel agonists and/or antagonists, particularly in the fight against cancer[10,11]. Specifically, BB3R agonists could be used in the treatment of obesity/diabetes mellitus[12] and BB1R or BB2R antagonists for the treatment of itching in atopic dermatitis[13]. Hence, they are interesting targets for drug discovery[14,15]. However, to develop new drugs, a deeper understanding of its structure-activity relationships is necessary. Despite the knowledge accumulated from previous studies about the potential use of bombesin antagonists as therapeutic agents, progress has been hampered by

the lack of diversity in the available bombesin ligands. A few peptide antagonists with diverse selectivity have been discovered for the three receptors[9,16,17]. However, it is more desirable to make non-peptide molecules because of the low absorption, poor oral bioavailability, fast degeneration by proteolytic enzymes, and immunogenic profile of peptides[18,19]. Efforts in this direction led to discover of second-generation peptoids PD168368 and PD176252, along with a set of analogs with diverse substitutions[20] exhibiting antagonistic profile for BB1R and BB2R, with a diverse degree of selectivity. Furthermore, the same scaffold was used later to design ML18, a BB3R selective antagonist, and more recently, compounds AM-37 and ST-36 with diverse bombesin receptor pharmacological profiles have been disclosed[21,22].

All these compounds are analogs of the same chemical structure and it would be interesting to have compounds with a different chemical scaffold. Recently, it was reported that both PD168368 and PD176252 act as potent antagonists of the human formyl-peptide receptors, questioning if the action observed by these compounds is only due to their BB1/BB2 antagonists profile[23]. Until now only a few compounds without peptoid scaffold have been reported: a dibenzodiazepine scaffold for BB1R[24], NSC-77427[25], and Bentag-1 for BB2R. The latter is a peptidomimetic designed by an isosteric replacement that also shows antagonism for BB3R[26]. In order to have a better understanding of the therapeutic effects of bombesin antagonists, it is necessary to discover novel compounds involving diverse chemical structures.

The work reported in the present chapter has been designed with the aim to discover novel chemical scaffolds for the three bombesin receptors, suitable for developing novel antagonists. For this purpose, 3D models of three bombesin receptors were constructed by homology modeling and refined using molecular dynamics. The modeling of the BB1 and BB2 receptors was carried out with the antagonist PD176252 bound on the orthosteric site, and in regards to BB3 the refinement process was carried out in complex to AM-37. Analyses of ligand-receptor complexes together with known structure-activity relationship studies were used to define pharmacophores for each three receptors. Unfortunately, with the experimental information available we only could define unambiguously a pharmacophore for the BB1R. For the other two receptors, we suggest diverse possibilities. In the case of the BB1R, the pharmacophore defined was used to carry out a virtual screening that led to the identification of a set of small molecules that were purchased and tested at 50 µM for its capacity antagonized NMB at the BB1R.

5.2 Construction of the 3D models of the BB1R, BB2R, and BB3R

Crude models of human BB1R, BB2R, and BB3R were constructed by homology modeling using the rat neurotensin receptor NTS1 as a template (PDB ID: 4GRV)[27]. The selection of this template comes from being one of the few receptors located in some branch of bombesin in GPCRs phylogenetic three with observed crystallographic structures[28,29]. Since the 4GRV crystal structure is a fusion protein of the NTS1 receptor and the T4 lysozyme, the template was modified by removing the coordinates of the T4 lysozyme and joining the segment of the ECL3 left at both sides. In the next step, the sequences of BB1, BB2, and BB3 receptors were aligned with the NTS1 sequence separately.

The alignment of the template and target is a critical step for being sure about the accuracy of the constructed models by homology modeling[30]. Sequence identity between the target receptors and the template is low: 20% identity and 34% homology for BB1R; 20% identity and 37% homology for BB2R and 19% identity and 36% homology for BB3R. Accordingly, multiple sequence alignment was performed to improve the quality of the alignment. For this purpose, we chose a set of 20 sequences of various GPCRs from class A with known crystallographic structure and their sequences were aligned with the sequences of the template and each bombesin receptor subtypes using the CLUSTALW software[31].

Figure 5.1 shows the results of multiple sequence alignment. As can be seen, all the conserved residues in the family are well aligned (using the Ballesteros-Weinstein numbering scheme[32]): N1.50 in TM1; L2.46, A2.47, D2.50 in TM2; D/E3.49, R3.50, Y3.51 in TM3; W4.50 in TM4; F5.47, P5.50, Y5.58 in TM5; F6.44, W6.48, P6.50 in TM6; N7.49, P7.50, Y7.53 in TM7. If we focus on the sequence of the NTS1 and the bombesin receptors, we can see the conserve residue Y5.58 in the 5th transmembrane domain of the bombesin receptor is corresponding to N5.58 in the NTS1 receptor, however, this difference does not produce any structural consequence. In contrast, Y7.53 in all the rhodopsin family corresponds to Y7.54 in bombesin receptors. If we do the closer inspection to the alignment, it can be suggested that this is corresponding to displacement but not an insertion and it can affect inducing a buldge in TM7. The other well-aligned conserved motifs in the alignment are like D(E)RY in TM3, CWxP(Y/F/L) in TM6, or the NPxxY in TM7.
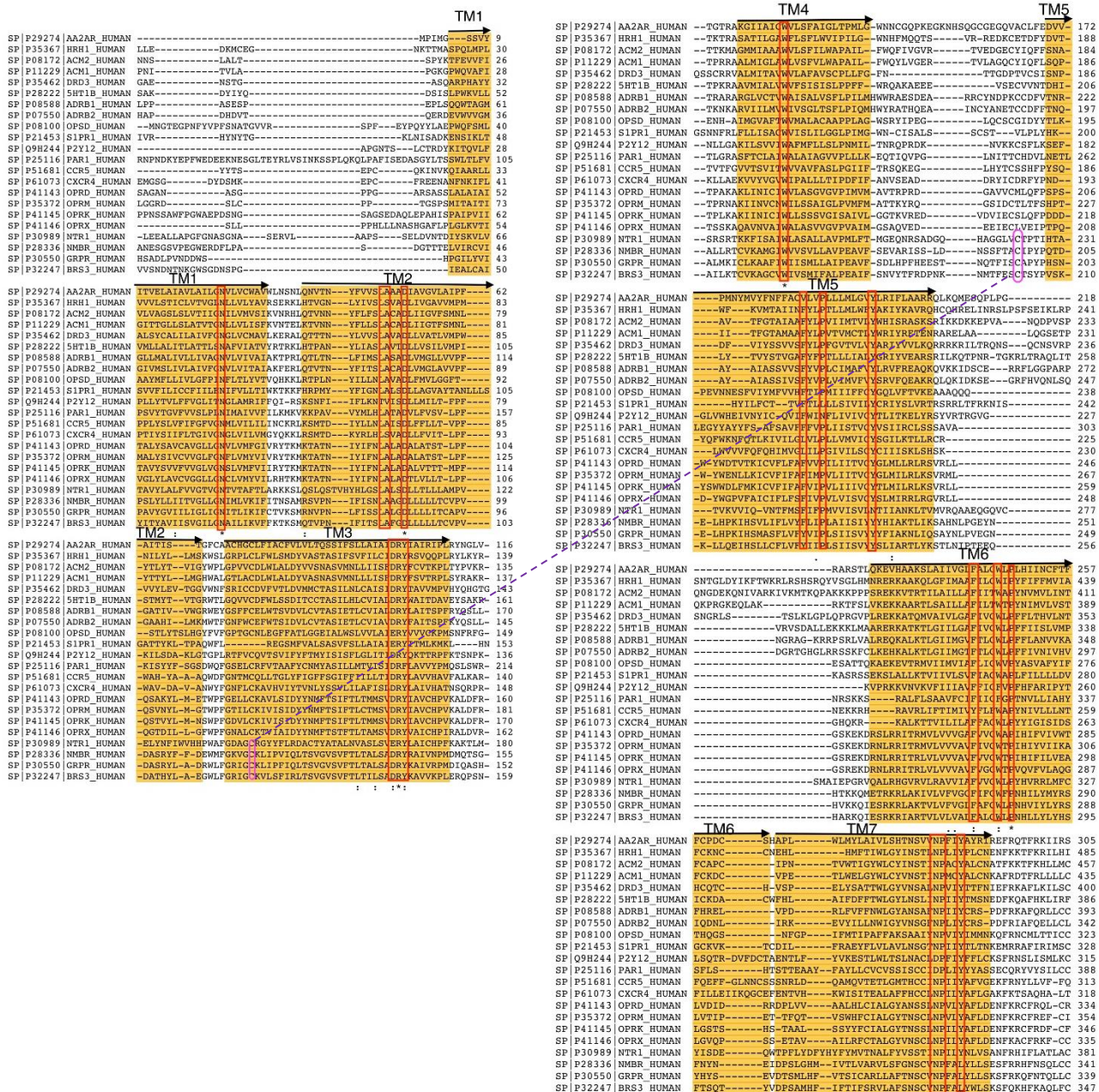
Figure 5. 1 *Multiple sequence alignment of diverse GPCRs used in the present work (see text). Transmembrane segments are colored in orange and conserved residues are in red boxes. There is also a purple line indicating a disulfide bridge.*

In a subsequent step, the alignment was used to thread the sequence of each of the bombesin receptor onto the backbone of the template structure, using the molecular operating environment (MOE) program[33]. After performing the homology modeling process, thirty models were produced for each of the bombesin receptors. The diverse models correspond to the incorporation of alternative side-chain

116

conformations by using a rotamer library produced from a high-resolution structural database, embedded in MOE. Then, missing hydrogens were added using protonate3D method[34] and subsequently, energy minimization of each of the models was carried out using a contact energy function to eliminate serious steric strains. Models were checked based on inter-residue interaction as well as backbone conformation using Ramachandran Plot and scores. Then, models were selected for further refinement process taking into consideration the lower root-mean-square-deviation (rmsd) in regard the average structure and the highest score in comparison to the other models.

The refinement process was performed through a 500ns molecular dynamics simulation for BB1R, BB2R, and BB3R in complex with the ligand-bound in the orthosteric binding site. The presence of the ligand makes the refinement process to be more efficient[29]. Accordingly, the antagonist PD176252 was docked in the orthosteric site of crude models of the BB1R and BB2R separately and Am-37 to BB3R using the GLIDE software[35]. After performing the docking process, multiple orientations of receptor-ligand complexes were produced and ranked ordered using the XP scoring function. In the case of the BB1R and BB2R, the poses with the highest score were selected for energy minimization in vacuo with a distance-dependent dielectric constant of 2. In the case of the BB3R minimization was carried out without any ligand bound. Minimization was carried out using the steepest descent method to eliminate any possible contacts in the structures. After that, the ligand-receptor complex of each receptor was embedded in an equilibrated box consisting of a lipid bilayer of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and water molecules, as described elsewhere[36]. The initial dimension of the box was 8.9 ×8.3 ×10.5 nm3 (XYZ) on the bilayer plan orientated on the XY plane of this box. The protein was positioned in the center of the box and the overlapping molecules were deleted. Specifically, the removed overlapping molecules including all water molecules with oxygen atoms closer than 0.40nm to the non-hydrogen atom of proteins as well as all lipid molecules with at least one atom closer than 0.25nm to a non-hydrogen atom of protein. This process results to have the final system with different amounts of lipids and water atoms for each system. These amounts of lipids and water molecules for each system are 193 and 12000 for BB1R, 193, and 14317 for BB2R and 192 and 14304 for BB3R, respectively. Because of removing the overlapped molecules, a small voids were produced between protein and water or lipid molecules that were removed during the first part of MD simulation trough adjustment of the lipid bilayer and water molecules to the protein. In a next step, some molecules of water molecules were selected randomly and replaced by sodium and chloride ions to have a system with an approximate concentration of 0.2 M on sodium chloride. Chloride and sodium ions were added in such a way to make the systems neutral. Specifically, the number of sodium and chloride ions for each system were: 47 and 54 for BB1R, 46 and 56

for BB2R and 44 and 57 for BB3R, respectively. This concentration is as same as a biological organism, although they showed various intra- and extra- cellular ion concentration. Then the energy minimization of each system was carried out to prevent steric clashes using the steepest descent method. The MD simulation proceeds to 500ns with constant pressure using GROMAX package 4.6[37]. Description of molecules during MD simulation was based on the OPLS-AA force field[38] embedded in the GROMACS package and the only exception corresponded to the water molecules that were based on the TIP3 model[39]. The systems were defined with periodic boundary conditions in the three coordinate directions. The temperature was kept constant at 300K using different thermostats for each component of the system such as protein, water, ions, and lipid molecules. The time constant was set to 0.1ps for the thermostat, except for water with a smaller value of 0.01ps. The pressure constant in three coordinate directions was set at 0.1MPa by independent Berendsen barostat using a time constant of 1.0ps. The equation of the motion was carried out using Leapfrog Algorithm with a time step of 2fs. Freezing of all bonds containing hydrogen atoms within the protein and lipid molecules was done using the LINCS algorithm[39]. Fixing of the water molecules was done using the analytical SETTLE method[40]Computing of Lennard-Jones interactions was done using a cut-off of 1.0nm. Also, the electrostatic interactions were computed using the PME technique[40].

Figure 5.2 shows the root-mean-square-deviation (rmsd) of Cα of the three bombesin receptors as well as root mean square of fluctuation of them during the MD trajectory. As can be observed the structure of the BB1 and BB2 receptors are equilibrated after 100ns, which this result has been observed previously for other MD simulations of some GPCRS[29]. However, the equilibration of the BB3 receptor took more time due to the absence of ligand. Evaluation of the rmsd permits us to construct the final model of the receptors from the average structure from the last 100ns of the molecular dynamics simulations trajectory. Then, the average structure was minimized in a two-step procedure using the steepest descent method with a distant dielectric constant of 2. In the first step, side chains of the structure were optimized with the backbone atoms constrains and then they proceed to the second minimization step. Also, the evaluating of the root-mean-square-fluctuations (rmsf) of the receptors demonstrates that the fluctuation of the loops is more than the transmembrane regions of the proteins.
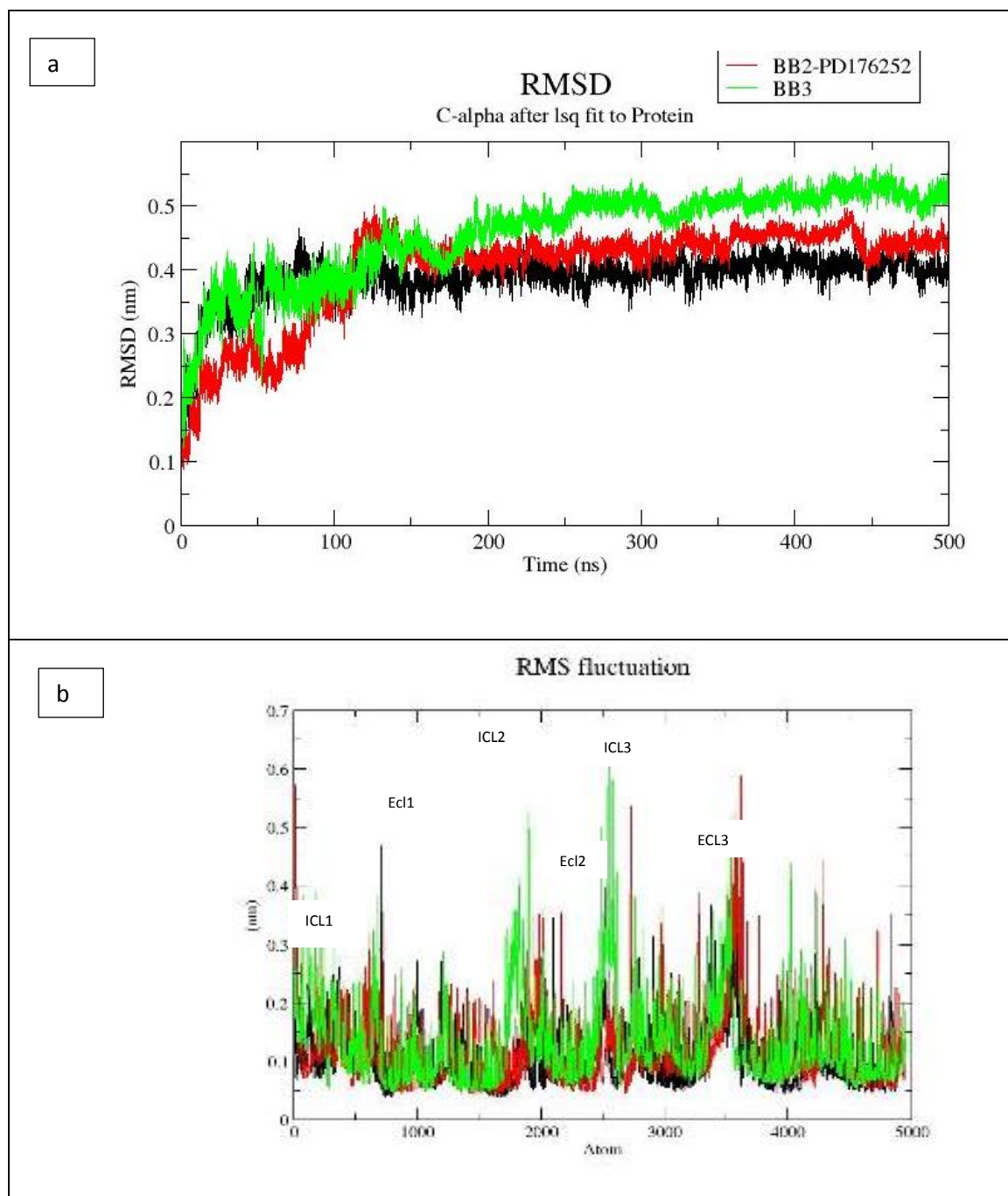
Figure 5. 2 *a) root-mean-square-deviation (rmsd) of Cα of the three bombesin receptors structures b) root mean square of fluctuation of three bombesin receptors*

5.3 Pharmacophore for BB1R antagonism

Since we were interested in understanding the molecular features of ligands binding to the BB1 receptor, we compared the binding modes of both PD176252 and PD168368 when bound onto the BB1 receptor. For this aim, we used all the available information such as structure-activity relationship[41,42],

119

available mutagenesis studies[43], as well as residue differences between the orthosteric binding site of BB1 and BB2 receptors. Accordingly, the constructed model of BB1R was used for docking of various conformation and orientation of the antagonist ligand. For this aim, the same protocol was applied as described previously. After the docking process, multiple orientations were produced and ranked based on the score using the XP scoring function of GLIDE[35]. Energy minimization of the produced complexes was performed in vacuo with the distance dielectric constant of 2, using the steepest descent method to allow the ligand to have the best orientation in the environment.

For the comparison of the two ligands, first, we evaluated the structures of both ligands by taking the asymmetric carbon as reference. It can be observed that the structure of both ligands contains three branches involving the including the nitrophenylurea moiety, the indole moiety of the tryptophan side chain, and the 2-pryridinecyclohexane moiety. Evaluation of diverse conformation and orientation of the ligands shows that each branch of the ligand can occupy the different site of the orthosteric binding site of the BB1 receptor and it let us analyze produced interactions between ligand and receptor. As can be expected it is possible to compare orientations of both PD168368 and PD176252 ligands because they have similar size and there are small differences between their structures. Specifically, the most important difference corresponds to having extra methoxyl moiety at the Pyridine part in PD176252 as can be observed in figure 5.3. However, it should be taken into account that this extra moiety provides more affinity of PD176252 to BB2R which can be proposed that there is an additional favorable interaction with this receptor.
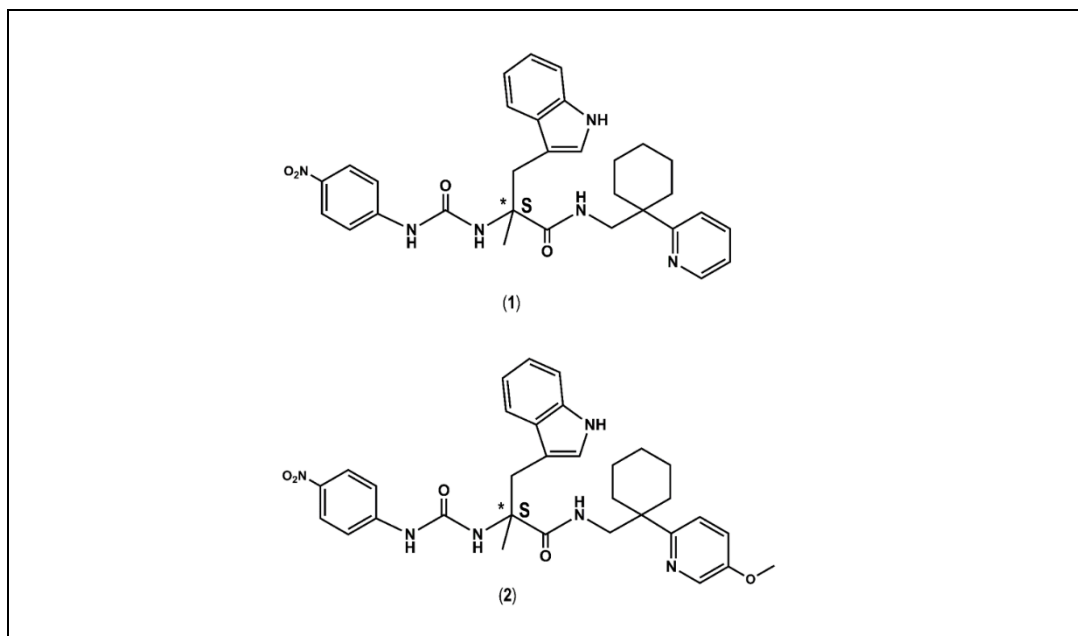
Figure 5. 3 *Chemical structures of the bombesin antagonists studied in the present work. PD168368 (1) and PD176252 (2). The chirality of the asymmetric center is also specified.*

Therefore, for evaluating most likely binding mode of the ligand to the receptor, diverse ligand-receptor interactions were analyzed by the aim of identifying the key residues and checking conservation of them in BB2R, to be sure that there is a specific interaction responsible for the differential binding affinity of PD176252 antagonist for both receptors. As mentioned in the previous study, Tyr[220] of the BB1 receptor is one of the key residues that involve in the binding of both ligands. Actually, mutation of Tyr[220] to phenylalanine decreases the affinity of both compounds for BB2R[43]. These results proposed that Tyr[220] in BB1 receptors plays a critical role in the interaction with ligand and this interaction can disappear by mutation of Tyr[220] to Phe[220]. In addition, this mutation suggests that the hydroxyl group of Tyrosine involve in hydrogen bond interaction with the ligand. This result was evaluated in the selected receptor-ligand docking poses. As can be seen in Figure 4, Analyses of the various poses demonstrated that the hydroxyl group of the Tyr[220] side chain points toward the center of the aromatic ring of the ligand nitrobenzene group. Interestingly, there is a quadruple-quadruple interaction between the aromatic ring of the ligand and the aromatic ring of Tyrosine sidechain that produces a T-shape relative orientation. All this information can explain the difference in the affinity of PD168368 and Pd176252 for the BB2 receptor. While there is a quadruple- quadruple interaction between aromatic rings of ligand and BB2 receptor, so the decrease of affinity is corresponding to the loss of hydrogen bond[44]. Also, the nitrophenyl part of both of the ligands involved in the hydrogen bond interaction with His[283] (this residue is conserved in BB2R) as shown in Figure 5.4.
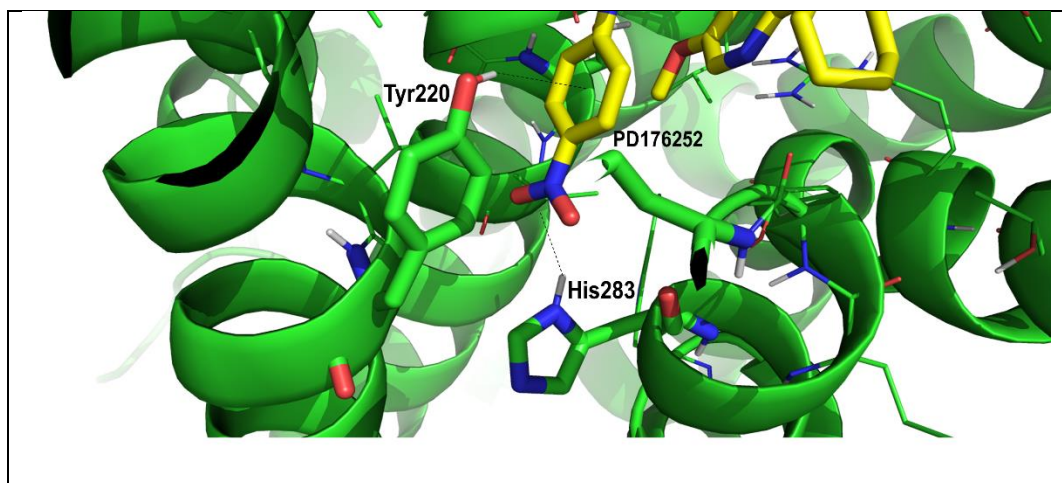
Figure 5. 4 *Close up of interaction between the nitrophenyl moiety of PD176252 and diverse residues BB1R including Tyr220 and His286(See Text)*

The indole moiety of both ligand position in a hydrophobic pocket surrounding by the residues Phe[181], Pro[120,] and Leu[215] (Figure 5.5). In addition, the orientation of the indole moiety permits a hydrogen bond interaction with Glu[187]. All the mentioned residues are conserved in BB2 receptor, so these interactions can justify part of the affinity of the ligand, however, it cannot justify the affinity differences of PD176252 and PD168368 for the BB1R.
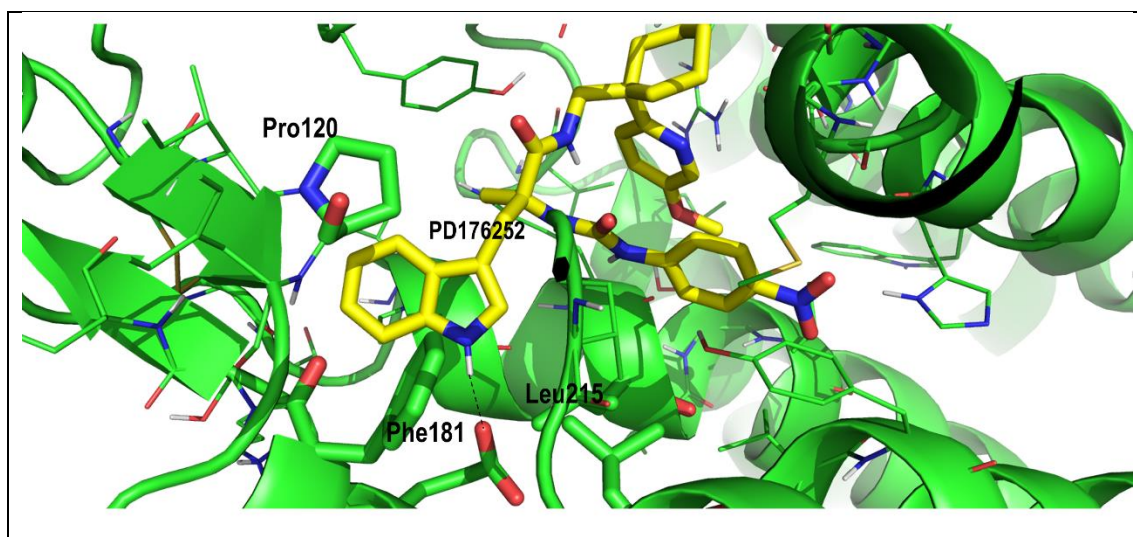
Figure 5. 5 *Close-up interaction between indole moiety of PD176252 and residues Phe181, Pro120, and Leu215(See text)*

To understand what is the difference between the binding affinity of the two ligands, we should focus on the binding mode of the 2-pyridinecyclohexane moiety. Since the size of both molecules is the

same, it can be hypothesized that the methoxy group attached to the pyridine ring is responsible of increase the magnitude affinity to the BB2 receptor. Evaluation of the set of poses that had interaction with Tyr[220], released that the methoxy group positioned near to the sidechain of Ser[126] (conserve residue in BB2R) of the BB1 receptor that in this case, oxygen of Methoxy moiety interacts with sidechain of Ser[126](Figure5.6), Moreover, Gln[123] (also is conserved in BB2R) has a hydrogen bond with the nitrogen of the heterocycle. And also there is a hydrogen bond between the aromatic pyridine ring of PD176252 ligand and the sidechain of the Arg[310] that interacts tightly with Asp[100] (both are conserved in BB2R). Among all the mentioned interactions, the exhibition of the hydrogen bond between Ser[126] of the ligand can prove the difference of the affinity between PD168368 and PD176252 to BB2 receptors. However, this information cannot explain why the affinity of both ligands remains the same as the BB1 receptor. This can be explained by the different binding conformation of the 2-pyridine moiety in both ligands. In PD168368, the pyridine ring tends to be near to the Arg[289] in the way that side chain of Arg[289] has a hydrogen bond with the nitrogen of the heterocycle (Figure5.7). Since Arg[289] is a conserved residue in both BB1 and BB2 receptors, it cannot explain the difference of the compound affinity for these two receptors. However, the heterocycle moiety of the ligand has a quadruple- quadruple interaction with Phe[105], which this interaction cannot be seen in the BB2 receptor since this position corresponds to Leu[102] in BB2 receptor. Moreover, the cyclohexane group of the ligand is positioned in a site surrounded by Met [287](Leu[285] in BB2R), the conserved Ile[296], and the aliphatic chain of the conserved Lys[210]. All these explanations in addition to Tyr[220] interaction can explain the difference of compounds magnitude affinity for both receptors. Accordingly, the reason of why PD176252 shows the same affinity for both ligands can be explained as a compensation of the interactions, by the means of having a hydrogen bond with Ser[126] can compensate the loss of the quadruple- quadruple interaction with Phe[105].
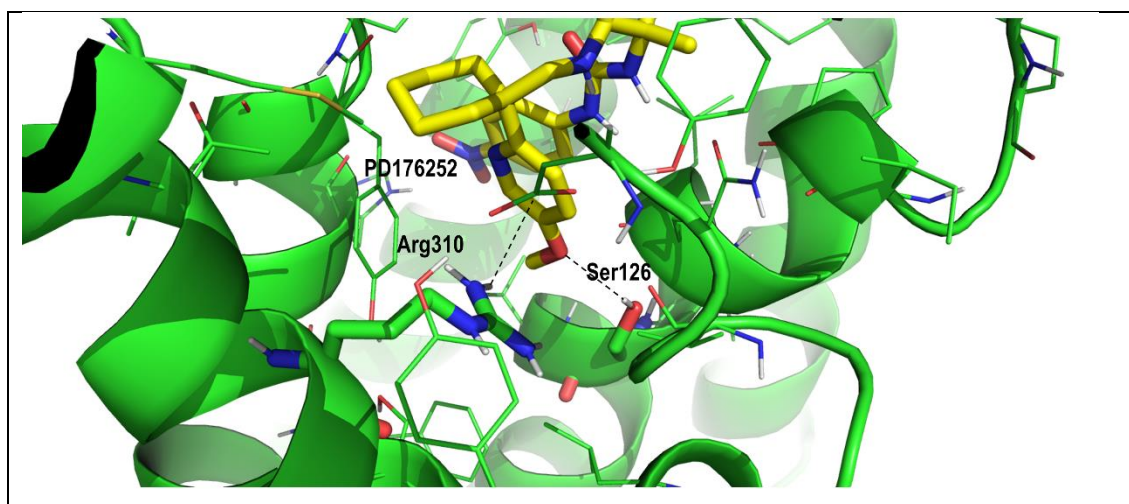
Figure 5. 7 Showing the

interaction between its phenoxyl-2-pyridine moiety of PD176252 and diverse residues OF BB1 receptor including Arg289 and Phe105(see text)

According to the previous analysis, we can distinguish binding differences of PD168368 and PD176252 ligands bound at the BB1R. Figure 5.8 shows the superimposition of the two compounds binding in the orthosteric binding site of BB1R. As can be seen, the nitrophenyl part of both ligands is close to Tyr[220] and His[283] and the indole moiety positioned in a hydrophobic pocket surrounded by residues Phe[178], Pro[120] and Leu[215]. All these three residues are conserved in BB2R. Finally, it can be observed that the 2-pyridinecyclohexane moiety shows different binding conformation for both ligands. In the case of PD168368, an extended pyridine ring can be obtained, while the pyridine ring of the PD176252 has twice a dihedral angle about 60° which led to having a hydrogen bond with Ser[126]. In contrast, the special conformation of PD168368 lets the pyridine ring has hydrogen bond interaction with Arg[289] and a quadruple-quadruple interaction with Phe[105].
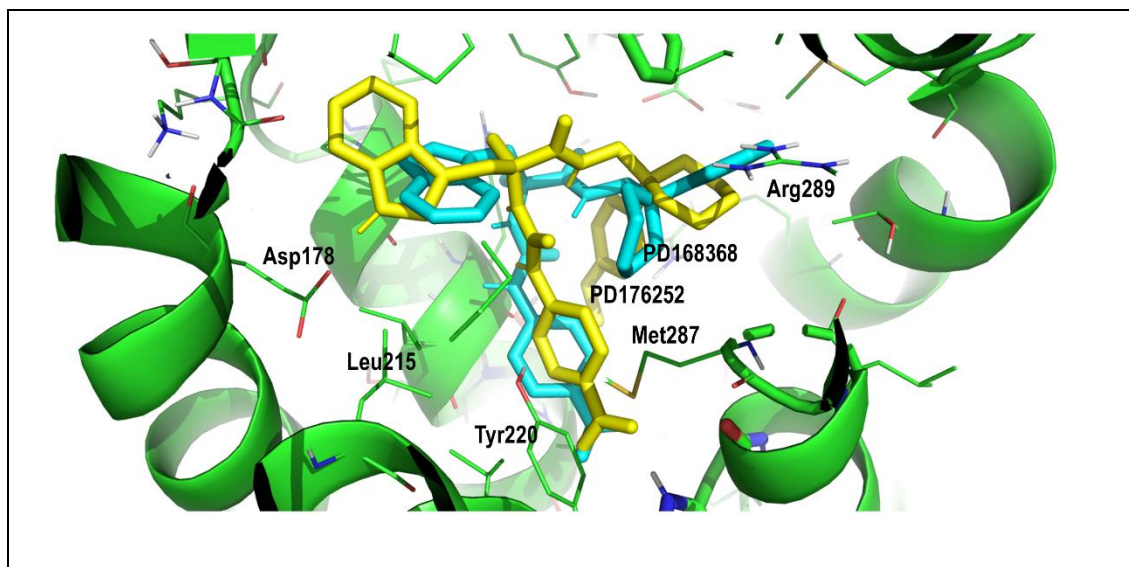
Figure 5. 8 *Superposition of PD168368 (cyan) and PD176252 (yellow) in their prospective bound conformations, respectively to the BB1R.*

To further assess the feasibility of both prospective bound conformations, we checked the extend that these models explain the available structure-activity results of this family of compounds[41]. Analysis of the activity of the diverse compounds shows that when the nitro group of the nitrophenyl moiety is attached in position 3 or when it is substituted by a proton accepting chemical group like nitrile, the compound preserves its activity. However, when the nitro group is placed in position 2 or substituted by a group with lower proton accepting capability, the affinity drops at least one order of magnitude. The same trends are observed regarding the binding affinity for the BB2R. On the other hand, substitutions on the 2-pyridine moiety that preserve a proton accepting center in position 4 have similar behavior.

5.3 1 Proof of concept

We proceed to identify the structure-activity relationship (SAR) of ligand-receptor by the meaning of finding the key resides and using selected bound conformation of PD176252 onto the BB1 receptor. Subsequently, a pharmacophore was defined from SAR study that this pharmacophore applied as a query for an *in silico* virtual screening. Accordingly, pharmacophore is defined with three pharmacophore points (Figure 5.9). The simplicity of the pharmacophore is because of finding easily hits with chemical scaffolds of diverse profiles. Also, the decision to define the pharmacophore is based on the involvement of Tyr[220] as proton accepting point, to discover selective compounds for BB1R whereas BB2R misses this capability because of the lack of this residue. The three defined pharmacophore points are i) a proton accepting center in the direction of the OH bond of the hydroxyl moiety of Tyr[220] side chain, located at 2.5 Å of the hydroxyl hydrogen; ii) a hydrophobic center located at a center defined by the side chains of Pro[200], Phe[181],

and Ile[211]; iii) a proton donor centrally located in the plane defined by the atoms of the carboxyl group of Asp[100] at 2.5 Å from the center of the two oxygens. These three pharmacophore points were applied as a query for in silico screening using various databases embedded in the molecular operating environment (MOE)[33]. More precisely, the radii of each of the spheres defined in the form that gives more tolerance for searching of hits in each point of pharmacophore points. Specifically, the radii of proton acceptor and donor spheres were defined as 1.2 Å, while the radii of the hydrophobic point were 1.8 Å. The two databases used for screening of 3D structure of the small molecules in databases embedding in MOE software. These databases including a Lead-like database containing 650,000 commercially available compounds[33] and the lead-now subset of zinc database with around 4,200,000 unique small molecules that this database downloaded in 2015[45]. Each database more than the 3D structure of compounds contain a set of generated conformations from a build-up procedure of systematic conformational searches from molecular fragments.
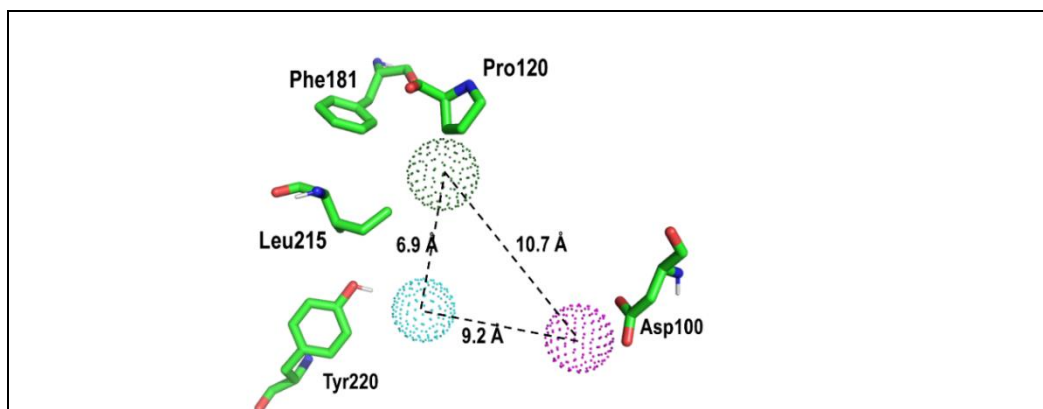
Figure 5. 9 *Point pharmacophore defined by the geometries of a few residues characterizing the sterochemical features involved in BB1R binding (see text)*

When the screening was carried out for 3D structures, a few hundred hits were recognized and subjected for further analysis. For this purpose, the encoding of molecules was done as a bit string using a typed atom triangle (TAT) which is one atom type method that atoms are grouped in trio based on their chemical nature and mutual distance[46]. In the next step, computing of the distance of two-bit strings performed using Tanimoto coefficient[47]. Finally, the molecules were clustered using the Jarvis-Patrick algorithm[48]. This procedure let us select a subset of compounds that were different from the initial set and had a diversity of chemical scaffold. The selection of representative molecules was done from diverse clusters based on their suitability of chemical groups responsible for each pharmacophore point and check the molecules not having any steric clashes. In this way, fifty compounds were selected and docked onto the BB1 receptor using the GLIDE software[35] and the docking complexes rank in the table based on the XP

scoring function. From all of the compounds, thirteen with the nest sores were purchased and tested at 50Mμ for its capacity to displace the radioligand to the BB1R, as explained in the binding assays part[3]. Among the tested compounds, the chemical structure of six of them finds to displace the radioligand used in the binding assays more than 15% which are summarized in Table 5.1. This result shows 50% success, as observed in the other studies[49,50]. Also, the result of the radioligand displacement experiment on BB1Rand BB2R shows that the ligands are selective for BB1R.

Radioligand displacement experiments were carried out by Eurofins using the following protocol: BB1 antagonism assays were carried out following a protocol described elsewhere[3]. Specifically, human recombinant bombesin BB1 receptors expressed in CHO-K1 cells were used in modified HEPESK OH buffer pH 7.4. A 0.2 μg protein aliquot was incubated with 0.1 nM $[^{125}I][Tyr^4]$-bombesin for 60 minutes at 25°C. Non-specific binding was estimated in the presence of 1 μM neuromedin B. Membranes were filtered and washed, the filters were then counted to determine $[^{125}I][Tyr^4]$-bombesin (Kd= 0.13nM) specifically bound. The compound potency was calculated as the percentage of displacement of radioligand by our molecules at 50 μM.
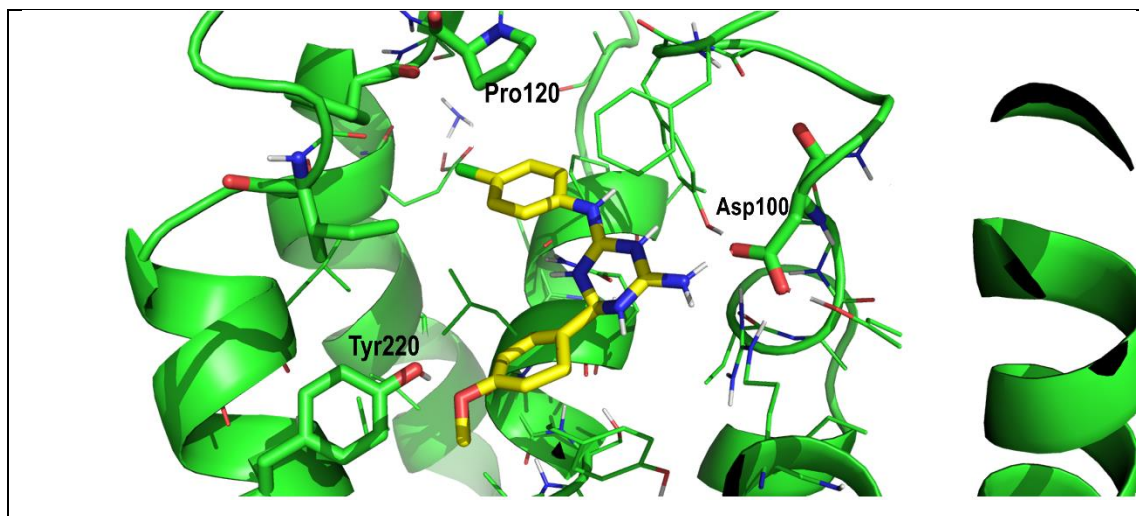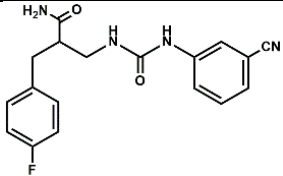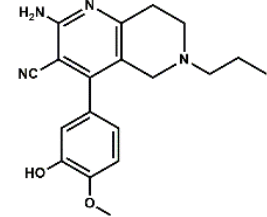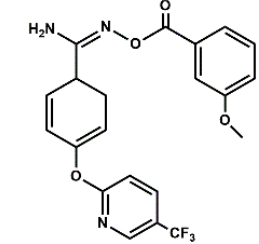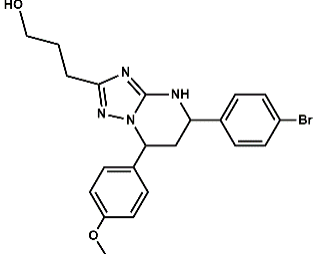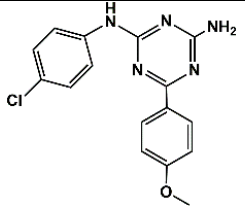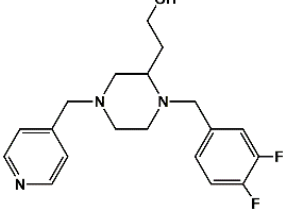
Figure 5. 10 *Pictorial view of the proposed binding mode of Compound#5 to the bombesin BB1 receptor.*

In order to explain further present results, the novel disclosed hits were docked to the BB1 receptor in the same protocol as described in the method section previously. Among all the compounds, compound#5 shows the highest inhibition with 38% BB1R radioligand displacement. Figure 5.10 shows its bound conformation to BB1R. As can be seen, the triazine ring works as a scaffold with three branches: a phenoxy moiety, a chlorophenyl moiety, and an amine. Further analysis of the prospective bound conformation demonstrated that the oxygen of the methoxy group fulfills pharmacophore point#1 by having a hydrogen bond interaction with the hydroxyl part of Tyr[220]. Also, the chlorine atom position in

the hydrophobic point by the means of fulfilling pharmacophore point#2. Finally, amine can fulfill pharmacophore point#3 by having a hydrogen bond with Asp[100]. However, in the case of BB2R, since the ligand cannot have a hydrogen bond interaction with Tyr[220], it can only fulfill two points of the pharmacophore and because of that shows a lower affinity for BB2R in comparison to BB1R. Finally, although the data in the table demonstrated that the affinities of the other molecules are not very high for BB1R, however, they represent a set of small molecules with high diversity that can be the starting point for discovering novel selective antagonists for BB1R.

Table 5. 1 *Listing of small molecules identified in the in silico screening (see text). Column 2 shows their chemical structure and columns 3 and 4 the displacement of the corresponding BB1R and BB2R radioligands, respectively(in percentage)at 50μM(N=2)*

Compound#	Structure	BB1 Radioligand Displacement (%)	BB2 Radioligand Displacement (%)
1		19.7	0.0
2		24.3	0.0
3		28.1	0.0
4		30.5	0.0
5		38.0	10.5
6		16.1	0.0

5.4 Pharmacophore for the BB2R antagonism

Docking of PD176252 in several conformations onto the BB2R was carried out with the side chains of the receptor allowed to move. 196 poses of the ligand were rank-ordered according to the MOE scoring function and subsequently minimized using a dielectric constant of 2. The results of the docking study were clustered to have a limited number of poses so that the bound conformations could be explained in qualitative terms. Considering PD176252 as a molecule with three legs pending on a chiral carbon, the bound structures can be rationalized by focusing on the anchoring point of the nitro group of the p-nitrophenyl moiety and then analyzing the positions that the indole and the methoxylpyridine moieties occupy.

The nitro group is a potent proton acceptor moiety and sits in four different locations in the docking study. In some poses is located in the vicinity of Arg[287]; in other cases, the nitro is located close to Arg[308] and Ser[126]. Moreover, the nitro group is also located between the side chains of Arg[308] and Arg[287] and in the rest of the cases is located close to His281, as found in the prospect bound structure to the BB1R. Let us analyze these different poses in detail.

Figure 5.11 shows pose#1 once energy minimized as the representative bound conformation with the nitro group exhibiting a hydrogen bond with Arg[287]. In these poses, the indole moiety sits in the neighborhood of Pro[198], Tyr[101], and Phe[178] attaining quadrupole-quadruple interactions. In addition, the amine nitrogen of the indole ring is involved in a hydrogen bond interaction with the carbonyl of the Phe[178] backbone (Figure 5.12).
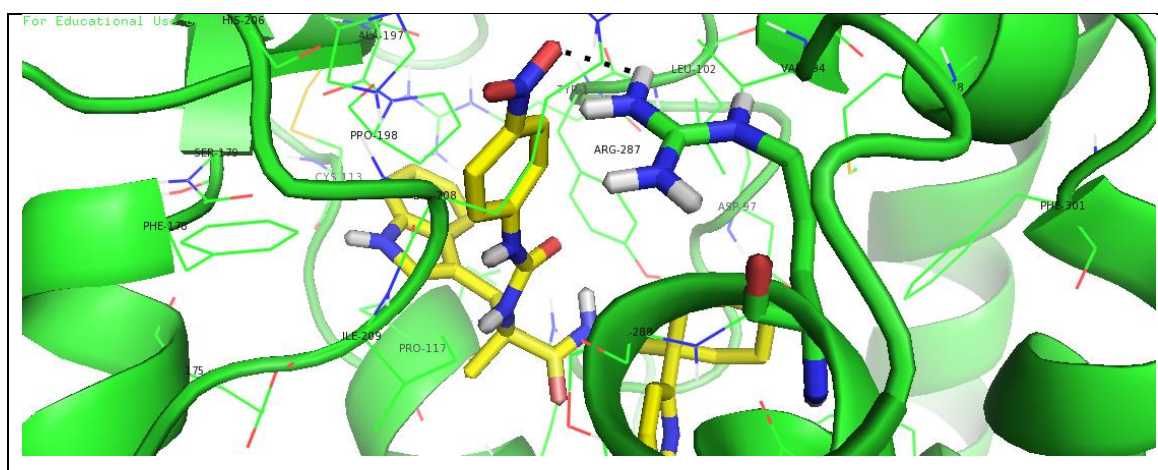

Figure 5. 11 *Close-up of PD176252 in its prospective bound conformation to the bombesin BB2 receptor, showing the interaction between its nitrophenyl moiety and Arg287 of BB2R.*
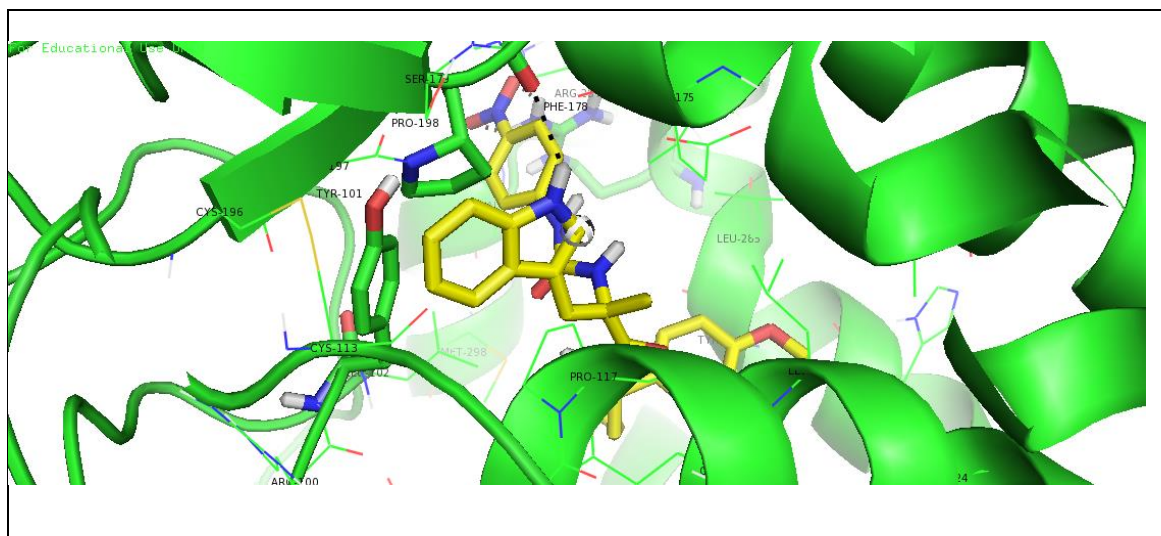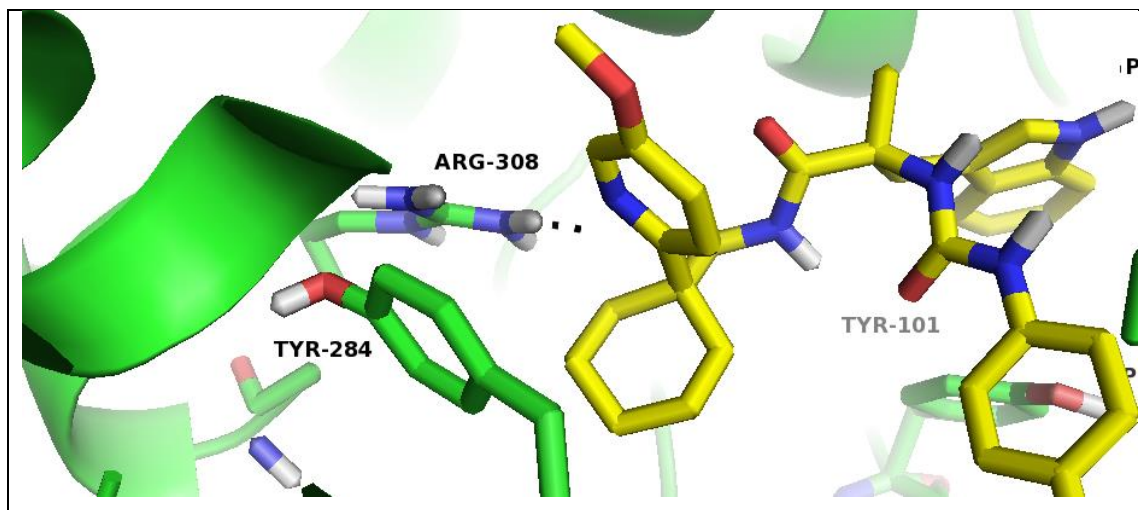
Figure 5. 12 *Close-up of PD176252 in its prospective bound conformation to the bombesin BB2 receptor, showing the interaction between its indole moiety and Phe178.*

On the other hand, the methoxylpyridine moiety sits in a hydrophobic pocket in such a way that the pyridine ring is involved in a π-π stacking interaction with the aromatic ring of Tyr[284]. Furthermore, the amine nitrogen of the pyridine ring forms a hydrogen bond with the Arg[308] side chain (Figure 5.13). Interestingly, the methoxyl group does not seem to play a specific function in this conformation and accordingly, this pose cannot explain the difference in affinity between PD176252 and PD168368.
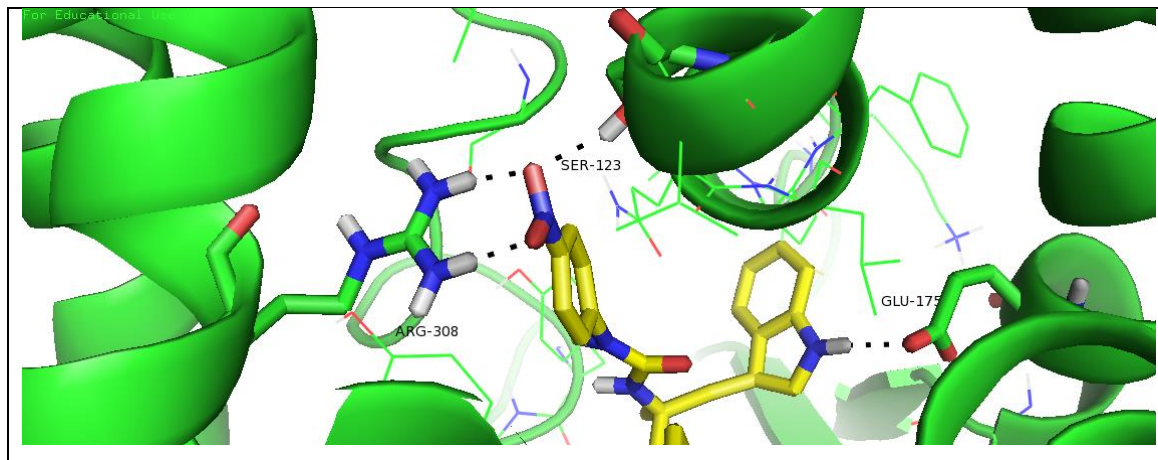

Figure 5. 13 *Close-up of PD176252 in its prospective bound conformation to the bombesin BB2 receptor, showing the interaction between its indole moiety and Phe178*

In the poses where the nitro group forms a hydrogen bond with Arg[308] and Ser[123] (Figure 5.14), despite several conformations available, the most likely is that the indole moiety occupies a pocket surrounded by the side chains of Pro[198], Tyr[101], Phe[178], and Glu[175], showing a hydrogen bond with the carboxyl group of Glu[175] (Figure 5.15).

In this pose, the methoxylpyridine moiety sits close to Arg[287] forming two hydrogen bonds: one with the heterocycle nitrogen and the other with the oxygen of the methoxyl group, whereas the cyclohexane moiety sits in a hydrophobic environment close to Leu[285] (Figure 5.16). In addition, there is a hydrogen bond between one of the NH groups of urea moiety and carbonyl of Gln[120]. This pose explains the difference in affinity observed between PD176252 and PD168368 and also, structure-activity studies showing that the amine of the pyridine increases the affinity of the methoxyl derivative.

Figure 5. 14 *Close-up of PD176252 in its prospective bound conformation to the bombesin BB2 receptor, showing the interaction between its nitro moiety and Arg308 and Ser123.*
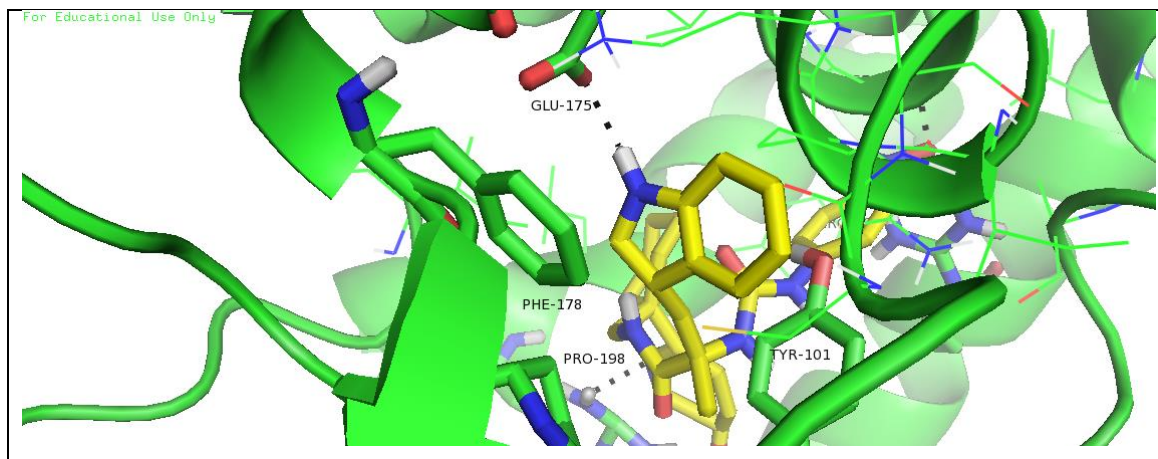
Figure 5. 15 *Pictorial view of the cavity of Indole moiety and its interaction with Glu175.*
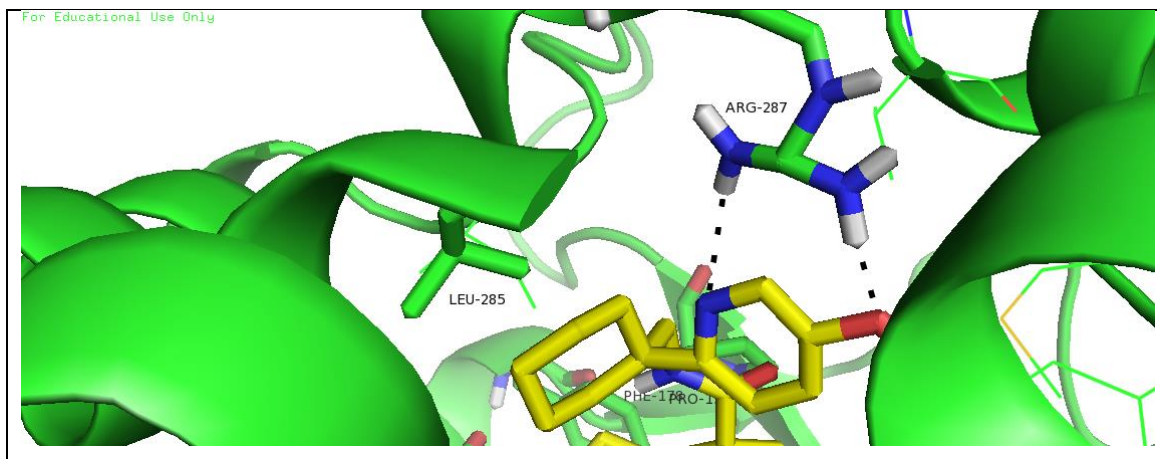
Figure 5. 16 *Pictorial view of the interaction between nitro and Methoxy group with Arg287 and sitting of cyclohexane moiety close to Lue285.*

In the case the nitro group forms a hydrogen bond with His[281] and Phe[218] (Figure 5.17), the ligand binds as explained in the case of the BB1R, although the fact that Phe[218] is a Tyr in BB1R makes the interaction to be weaker accounting by a substantial affinity difference. This pose, as explained above can explain the structure-activity studies.
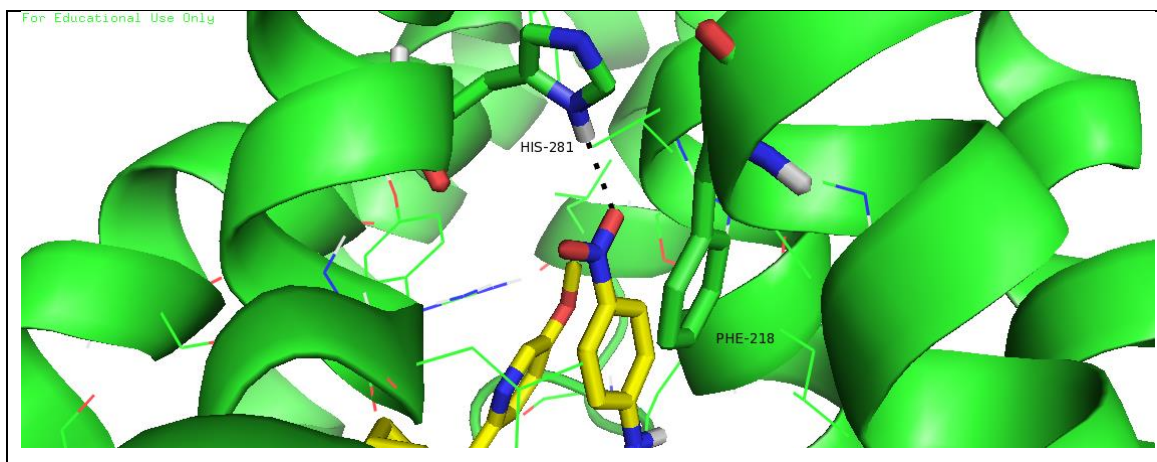
Figure 5. 17 *Close-up of PD176252 in its prospective bound conformation to the bombesin BB2 receptor, showing the interaction between its nitro moiety and His281 and Phe218.*

Finally, there is the case where the nitro group interacts simultaneously with Arg[308] and Arg[287] (Figure 5.18). Although they are interesting poses, the other legs of the molecule do not reach interesting spots for interaction. The most interesting pose is one where the ligand exhibits a conformation with the aromatic rings of the p-nitrobenzene and the methoxylpyridine forming a π-π stacking interaction. In this case, the nitro group exhibits a hydrogen bond with Arg[308] and the mehoxylpyridine with Arg[287] through the aromatic ring. In this position, the cyclohexane moiety is close to Leu[285]. Moreover, the indole moiety occupies the same pocket as in other cases, where several residues stabilize the indole moiety through

quadrupole-quadruple interactions. This pose can also explain the structure-activity studies available of these analogs. Present results confirm previous findings that residues Arg[287], Arg[308], Tyr[101] Tyr[284], Pro[198] and Gln[120] in BB2 receptor are involved in the ligand-receptor bound conformation of PD176252[20].
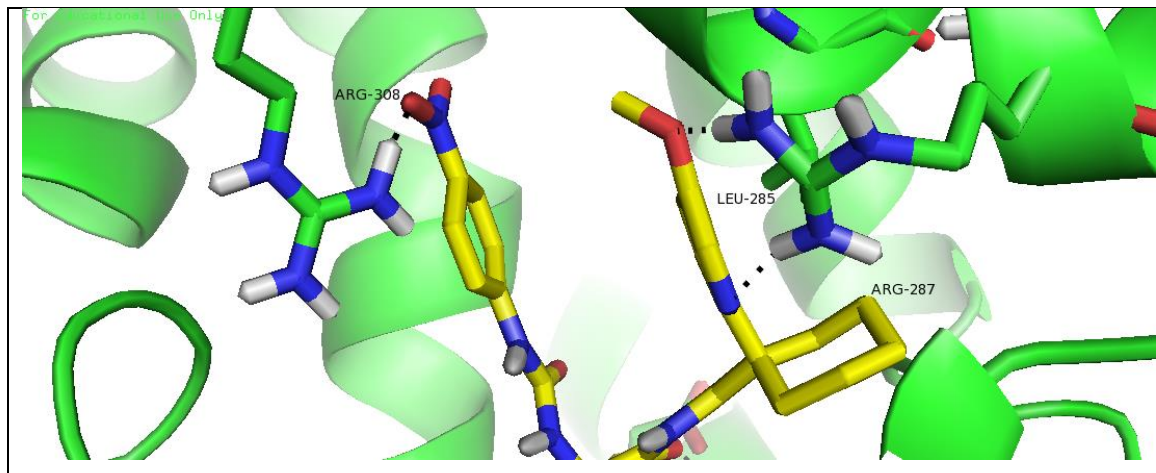

Figure 5. 18 *Pictorial view of the other possible nitro and Methoxy interactions of PD176252 in complex with BB2 receptor.*

In conclusion, according to the results of the docking studies, PD176252 may bind in three different forms onto the BB2R. The diverse poses identified locate the indole moiety in a hydrophobic pocket formed by Pro[198], Tyr[101], and Phe[178] suggesting a suitable site for the indole moiety. Moreover, Glu[175] or even the backbone carbonyl oxygen of Phe[178] might be involved in a hydrogen bond with the amine group of the indole moiety. However, the other legs of the ligand bind in different poses, making it difficult to define a unique pharmacophore. Site-directed mutagenesis studies could clarify the relevance of the diverse poses. Thus, the involvement of Arg[308] and Arg[287] could be key to understand the actual bound conformation of the compound onto the BB2R.

5.5. Pharmacophore for the BB3R antagonism

While diverse peptide antagonist have been disclosed for the BB3 receptor, the only reported small-molecule antagonists presently disclosed are the peptoides AM-37 and ST-36 –the R and S enantiomers of the same chemical structure- inspired in PD168368 and PD176252[22]. For this study, we decided to study AM-37 that shows a slight better pharmacological profile. Figure 5.19 shows its chemical structure together with those of PD168368 and PD176252. In addition, the affinity of the compound for the BB1R, BB2R and BB3R is also displayed. Understanding the pharmacological profile of this compound is very challenging. Despite the structural similarity between AM-37 and the PD-compounds, the former exhibits micromolar affinity for the three receptors. Specifically, AM-37 exhibits micromolar affinity for the BB3R, but also a decreased affinity for BB1R and BB2R in regard PD168368.
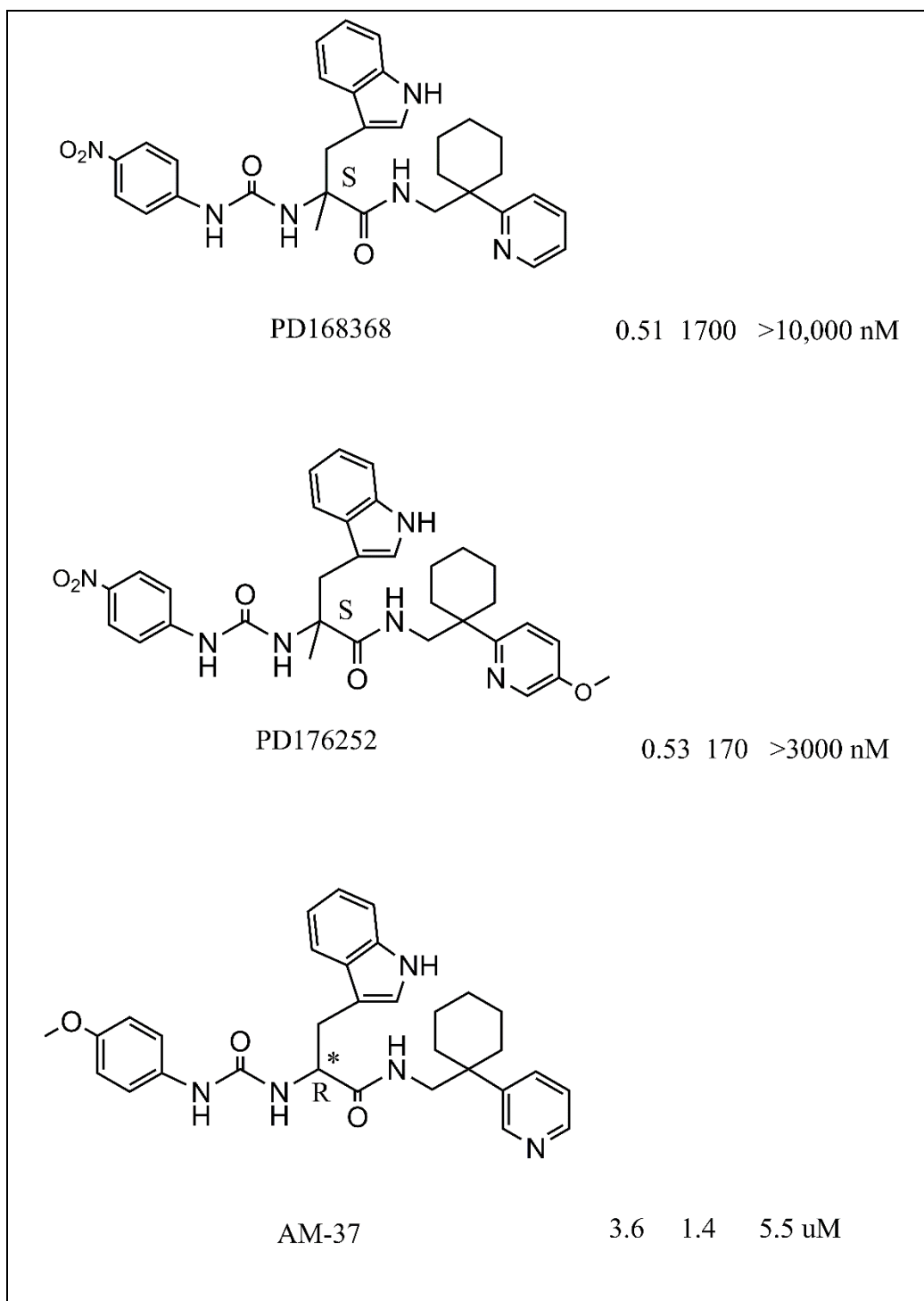
Figure 5. 19 *Chemical structures of PD168368, PD176252 and AM-37 as well as their affinity for the BB1R, BB2R and BB3R, respectively.*

The pharmacological profile of AM-37 tells us that the compound fulfills a common pharmacophore for the three receptors, but lacks of specific features, preventing it being a selective ligand for any of the three receptors. The second consideration to be taken into consideration is that the asymmetry provided

by the chiral carbon cannot be blamed for the observed pharmacological profile, since its enantiomer ST-36 exhibits a similar binding profile[22]. ST-36 exhibits three orders of magnitude lower affinity than PD168368 for the BB1R and similar affinity for the BB2R. According to the discussion outlined above, we consider that it is likely that the indole moiety of this family of compounds including ST-36 sits in the hydrophobic pocket in both BB1R and BB2R organized by the conserved residues Pro[120] (Pro[117] in BB2R), Pro[200] (Pro[198] in BB2R), Phe[181] (Phe[178] in BB2R) and Ile[211] (Ile[209] in BB2R) and in addition, forming a hydrogen bond with Glue[178] (Glu[175] in BB2R). In the case of the BB2R, binding affinity of PD168368 and ST-36 is similar and could be explained in the same terms as discussed above: lack of the methoxyl moiety attached to the pyridine ring, together with the replacement of Tyr[220] in BB1R for a Phe in BB2R that is conserved in BB3R. In contrast, it is harder to explain the low affinity towards the BB1R compared to PD168368. Our docking studies suggest that the ligand binds similarly to PD168368 but it lacks the hydrogen bond with His[283] since it is a bond shorter and also lacks the hydrogen bond of the pyridine nitrogen with Arg[289] side chain.

Now, in order to understand the micromolar affinity exhibited by ST-36 to the BB3R, we undertook a docking study. Analysis of the poses obtained reveals differential behavior in the way peptoid compounds bind to the BB3R in comparison to the BB1R and BB2R. Specifically, the indole moiety appears to change its favorite position in the receptor. Thus, the hydrophobic pocket found in BB1R and BB2R and described above, now becomes more hydrophilic since the flanking prolines (Pro[120] and Pro[200] in BB1R) are replaced in BB3R by polar residues: Ser[205] and Ser[124]. Despite these differences, we find poses in the docking study with the indole moiety bound in this cavity, although they are not as favorable as other poses (Figure 5.20).
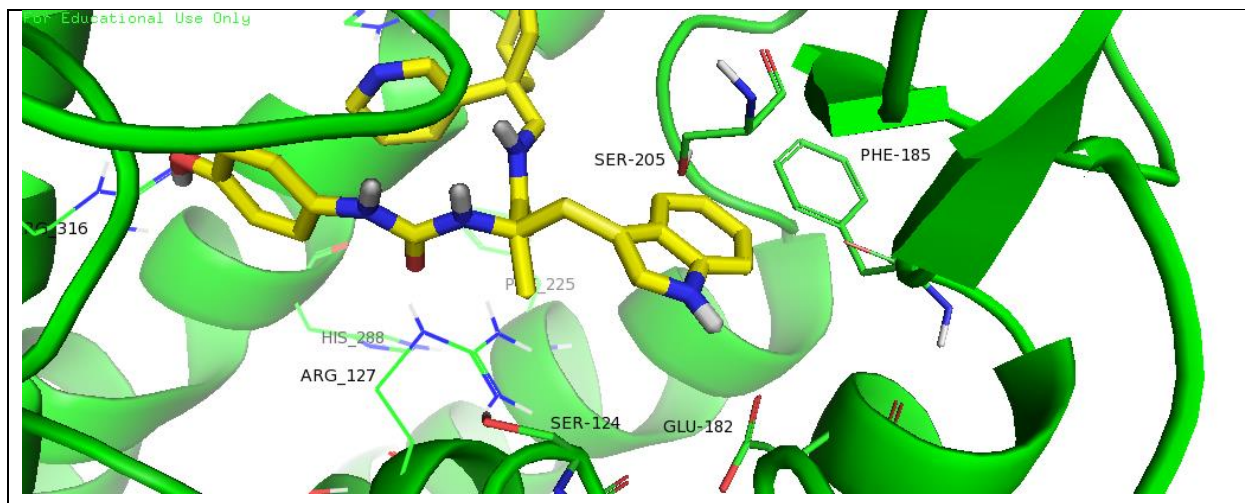
Analysis of the diverse poses found in the docking study, once rankordered suggest that the ligand can bind in very diverse forms. In one of the best poses, the indole moiety sits in a pocket in the neighborhood of Asp[104], Tyr[108] and Arg[316]. In this position, the indole amine nitrogen forms a hydrogen bond with the hydroxyl moiety of Asp[104] and Tyr[108]. Moreover, the cyclohexyl group sits in a hydrophobic pocket observed in diverse poses surrounded by Leu[292], Ile[216], Ile[69] and Tyr[291]. Finally, the pyridine nitrogen exhibits a hydrogen bond with Arg[127] and the phenoxyl moiety binds close to Ser[205] and Tyr[108] (Figure 5.21).
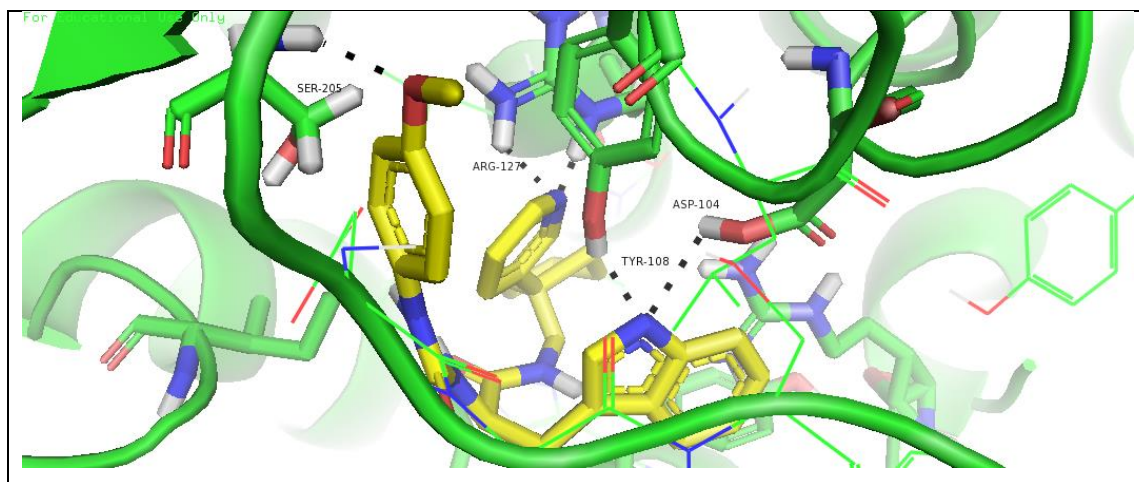

Figure 5. 21 Close-up showing interaction between Ser205 and phenoxyl moiety as well as Tyr108 and Asp104 with indole moiety and nitrogen of the pyrrole ring and Arg127.

In another interesting pose, the indole moiety binds in the pocket where the cyclohexyl group was before. Ser[130] may form a hydrogen bond with the nitrogen of the indole ring. In this case, Tyr[291] and His[288] may help to organize an environment suitable for an aromatic ring. In this pose, the cyclohexyl group sits in a pocket already observed with the PD-compounds close to Val[302]. In this position, the pyridine ring may form a hydrogen bond with His[294]. Finally, the phenoxyl group sits close to Ser[205] with the aromatic ring in the neighborhood of Ile[216] (Figure 5.22).
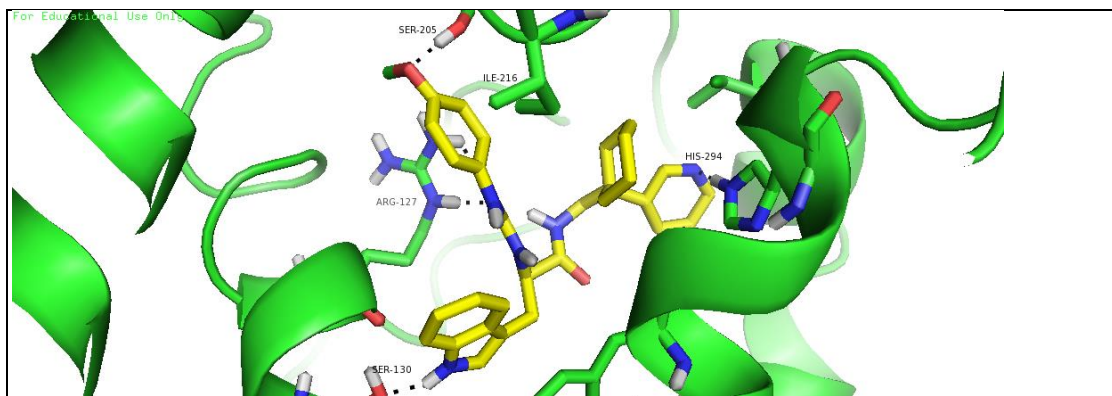
Figure 5. 22 *Pictorial view of the interaction between phenoxyl and Ser205,nitrogen of the indole and Ser130 as well as nitrogen of the pyridine ring and His294.*

In the other pose, the cyclohexyl group is again buried in a hydrophobic pocket as in first pose. In this case the pyridine ring may form a hydrogen bond with the hydroxyl group of Tyr[291] and Ser[130]. The methoxyl groups is close to Arg[316] and Tyr[108]. Finally, the indole moiety is facing the upper part of the orthosteric binding pocket forming a hydrogen bond with Glu[215] (Figure 5.23).
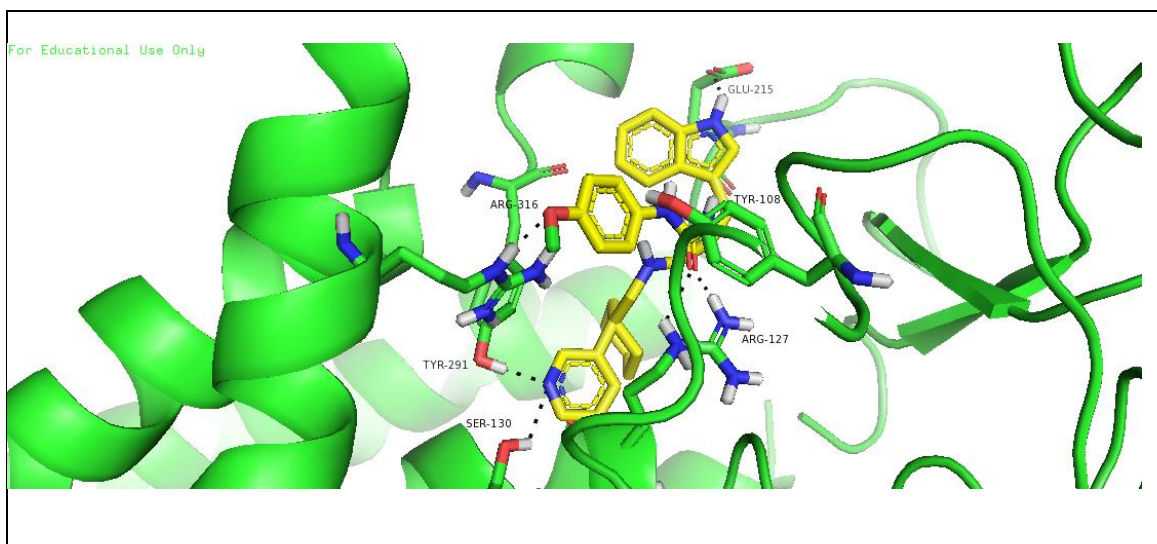

Figure 5. 23 *Pictorial view of the other possible nitro and Methoxy and interactions of ST-36 in complex with BB3 receptor.*

As can be seen, the lack of mutagenesis studies together with the lack a diversity of ligands prevent us to define a pharmacophore for the BB3R. Unfortunately, the paper where ST-36 was reported there are no studies of structure-activity that could be very useful to understand the effect of diverse substitutions on the affinity of the compound. However, present studies suggest the peptoid molecules bind in a different way that those observed in the BB1R and BB2R. Further structure-activity studies and/or mutagenesis studies could help to discard some of the hypothesis proposed in the present work.

5.6 References

1.   Anastasi, A., Erspamer, V. & Bucci, M. Isolation and structure of bombesin and alytesin, two analogous active peptides from the skin of the european amphibiansBombina andAlytes. *Experientia* **27**, 166–167 (1971).

2.   Erspamer, V. & Melchiorri, P. Active polypeptides: from amphibian skin to gastrointestinal tract and brain of mammals. *Trends Pharmacol. Sci.* **1**, 391–395 (1980).

3.   Ohki-Hamazaki, H., Iwabuchi, M. & Maekawa, F. Development and function of bombesin-like peptides and their receptors. *Int. J. Dev. Biol.* **49**, 293–300 (2005).

4.   Minamino, N., Kangawa, K. & Matsuo, H. Neuromedin B: A novel bombesin-like peptide identified in porcine spinal cord. *Biochem. Biophys. Res. Commun.* **114**, 541–548 (1983).

5.   McDonald, T. J. *et al.* Characterization of a gastrin releasing peptide from porcine non-antral gastric tissue. *Biochem. Biophys. Res. Commun.* **90**, 227–233 (1979).

6.   Jensen, R. T., Battey, J. F., Spindel, E. R. & Benya, R. V. International Union of Pharmacology. LXVIII. Mammalian bombesin receptors: nomenclature, distribution, pharmacology, signaling, and functions in normal and disease states. *Pharmacol. Rev.* **60**, 1–42 (2008).

7.   Flores, D. *et al.* Gastrin-Releasing Peptide Receptors Regulate Proliferation of C6 Glioma Cells through a Phosphatidylinositol 3-Kinase-Dependent Mechanism. *Curr. Neurovasc. Res.* **5**, 99–105 (2008).

8.   Jensen, R.T.; Moody, T.W. Bombesin peptides (cancer). In Hand-book of biologically active peptides; Kastin, A.J.; Ed.; Elsevier, Amsterdam, The Netherlands. 2013, 506-511.

9.   Ramos-Álvarez, I. *et al.* Insights into bombesin receptors and ligands: Highlighting recent advances. *Peptides* **72**, 128–144 (2015).

10.   Mahmoud, S. *et al.* [Psi 13,14] bombesin analogues inhibit growth of small cell lung cancerin vitro and in vivo. *Cancer Res.* **51**, 1798–802 (1991).

11.   Qin, Y., Ertl, T., Cai, R. Z., Halmos, G. & Schally, A. V. Inhibitory effect of bombesin receptor

antagonist RC-3095 on the growth of human pancreatic cancer cells in vivo and in vitro. *Cancer Res.* **54**, 1035–41 (1994).

12. González, N., Moreno, P. & Jensen, R. T. Bombesin receptor subtype 3 as a potential target for obesity and diabetes. *Expert Opin. Ther. Targets* **19**, 1153–1170 (2015).

13. Ehling, S., Fukuyama, T., Ko, M. C., Olivry, T. & Bäumer, W. Neuromedin B induces acute itch in mice via the activation of peripheral sensory neurons. *Acta Derm. Venereol.* **99**, 587–593 (2019).

14. Merali, Z., McIntosh, J. & Anisman, H. Role of bombesin-related peptides in the control of food intake. *Neuropeptides* **33**, 376–386 (1999).

15. Yamada, K., Wada, E., Santo-Yamada, Y. & Wada, K. Bombesin and its family of peptides: prospects for the treatment of obesity. *Eur. J. Pharmacol.* **440**, 281–90 (2002).

16. Cristau, M. *et al.* Synthesis and biological evaluation of bombesin constrained analogues. *J. Med. Chem.* **43**, 2356–2361 (2000).

17. González, N. *et al.* Characterization of putative GRP- and NMB-receptor antagonist's interaction with human receptors. *Peptides* **30**, 1473–86 (2009).

18. Perez, J., Corcho, F. & Llorens, O. Molecular Modeling in the Design of Peptidomimetics and Peptide Surrogates. *Curr. Med. Chem.* **9**, 2209–2229 (2002).

19. Perez, J. J. Designing Peptidomimetics. *Curr. Top. Med. Chem.* **18**, 566–590 (2018).

20. Carrieri, A. *et al.* Structural Determinants in the Binding of BB2 Receptor Ligands: In Silico, X-Ray and NMR Studies in PD176252 Analogues. *Curr. Top. Med. Chem.* **17**, 1599–1610 (2017).

21. Moody, T. W. *et al.* ML-18 is a non-peptide bombesin receptor subtype-3 antagonist which inhibits lung cancer growth. *Peptides* **64**, 55–61 (2015).

22. Moody, T. W. *et al.* AM-37 and ST-36 Are Small Molecule Bombesin Receptor Antagonists. *Front. Endocrinol. (Lausanne).* **8**, 176 (2017).

23. Schepetkin, I. A., Kirpotina, L. N., Khlebnikov, A. I., Jutila, M. A. & Quinn, M. T. Gastrin-Releasing Peptide/Neuromedin B Receptor Antagonists PD176252, PD168368, and Related Analogs Are

Potent Agonists of Human Formyl-Peptide Receptors. *Mol. Pharmacol.* **79**, 77–90 (2011).

24. Fu, J., Shuttleworth, S. J., Connors, R. V., Chai, A. & Coward, P. Discovery and optimization of a novel Neuromedin B receptor antagonist. *Bioorganic Med. Chem. Lett.* **19**, 4264–4267 (2009).

25. Martínez, A., Zudaire, E., Julián, M., Moody, T. W. & Cuttitta, F. Gastrin-releasing peptide (GRP) induces angiogenesis and the specific GRP blocker 77427 inhibits tumor growth in vitro and in vivo. *Oncogene* **24**, 4106–4113 (2005).

26. Moreno, P. *et al.* Comparative pharmacology of bombesin receptor subtype-3, nonpeptide agonist MK-5046, a universal peptide agonist, and peptide antagonist Bantag-1 for human bombesin receptors. *J. Pharmacol. Exp. Ther.* **347**, 100–116 (2013).

27. White, J. F. *et al.* Structure of the agonist-bound neurotensin receptor. *Nature* **490**, 508–513 (2012).

28. Costanzi, S. *et al.* Homology modeling of a Class A GPCR in the inactive conformation: a quantitative analysis of the correlation between model/template sequence identity and model accuracy. *J. Mol. Graph. Model.* **70**, 140 (2016).

29. Lupala, C. S., Rasaeifar, B., Gomez-Gutierrez, P. & Perez, J. J. Using molecular dynamics for the refinement of atomistic models of GPCRs by homology modeling. *J. Biomol. Struct. Dyn.* **36**, 2436–2448 (2018).

30. Nayeem, A. A comparative study of available software for high-accuracy homology modeling: From sequence alignments to structural models. *Protein Sci.* **15**, 808–824 (2006).

31. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; Thompson, J.D.; Gibson, T.J.; Higgins, D.G. Clustal W and Clustal X version 2.0. Bioinformatics 2007; 23; 2947-2948.32.   Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* **25**, 366–428 (1995).

33. Molecular Operating Environment (MOE). Chemical Computing Group Inc. 1010 Sherbooke St. West Suite #910 Montreal QC Canada H3A 2R7.

34. Labute, P. Protonate3D: Assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins Struct. Funct. Bioinforma.* **75**, 187–205 (2009).

35. Richard A. Friesner, *,† *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem*. **47**,1739-1749 (2004).

36. Cordomí, A., Edholm, O. & Perez, J. J. Effect of different treatments of long-range interactions and sampling conditions in molecular dynamic simulations of rhodopsin embedded in a dipalmitoyl phosphatidylcholine bilayer. *J. Comput. Chem.* **28**, 1017–1030 (2007).

37. Van Der Spoel, D. *et al.* GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).

38. Kaminski, G. A., Friesner, R. A., Tirado-Rives, J. & Jorgensen, W. L. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **105**, 6474–6487 (2001).

39. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

40. Miyamoto, S. & Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **13**, 952–962 (1992).

41. Ashwood, V. *et al.* PD 176252 — The first high affinity non-peptide gastrin-releasing peptide (BB2) receptor antagonist. *Bioorg. Med. Chem. Lett.* **8**, 2589–2594 (1998).

42. Carrieri, A. *et al.* Structural Determinants in the Binding of BB2 Receptor Ligands: In Silico, X-Ray and NMR Studies in PD176252 Analogues. *Curr. Top. Med. Chem.* **17**, 1599–1610 (2016).

43. Tokita, K. *et al.* Tyrosine 220 in the 5th transmembrane domain of the neuromedin B receptor is critical for the high selectivity of the peptoid antagonist PD168368. *J. Biol. Chem.* **276**, 495–504 (2001).

44. Bissantz, C., Kuhn, B. & Stahl, M. A medicinal chemist's guide to molecular interactions. *J. Med.l Chem.* **53**, 5061–5084 (2010).

45. Sterling, T. & Irwin, J. J. ZINC 15 - Ligand Discovery for Everyone. *J. Chem. Inf. Model.* **55**, 2324–2337 (2015).

46. Bender, A. *et al.* How similar are similarity searching methods? A principal component analysis of molecular descriptor space. *J. Chem. Inf. Model.* **49**, 108–119 (2009).

47. Lipkus, A. H. A proof of the triangle inequality for the Tanimoto distance. *J. Math. Chem.* **26**, 263–265 (1999).

48. Patrick, E. A. Clustering Using a Similarity Measure Based on Shared Near Neighbors. *IEEE Trans. Comput.* **C–22**, 1025–1034 (1973).

49. Krishna, S. *et al.* Pharmacophore-based screening and identification of novel human ligase i inhibitors with potential anticancer activity. *J. Chem. Inf. Model.* **54**, 781–792 (2014).

50. Gomez-Gutierrez, P., Campos, P. M., Vega, M. & Perez, J. J. Identification of a Novel Inhibitory Allosteric Site in p38α. *PLoS One* **11**, e0167379 (2016).

Epilogue

Present situation of the covid-19 pandemic has brought to the forefront bradykinin as an important mediator of the illness caused by the infection of SARS-Cov-2. In this framework, we carried a small research on the molecular features of non-selective small molecule antagonists of the bradykinin receptors and its utilization in drug repurposing, shown in this epilogue, based on the results derived in Chapter 4.

1. Introduction

Beginning in December 2019, a novel coronavirus designated SARS-CoV-2 was identified as the pathogen causing an international outbreak of respiratory illness termed Covid-19, originated in Wuhan, Hubei Province, China. Despite the virus has a fatality rate of only ~2-3% it exhibits a high transmission rates, resulting in a high overall death toll that forced the World Health Organization to declare SARS-CoV-2 as a pandemic infectious disease of international concern on March 11, 2020 [1]. Until July 18, 2020, there have been 14,108.240 confirmed cases of Covid-19 with 602.695 confirmed deaths [2]. Unfortunately, ~20% of the people infected are likely to develop pneumonia of varying severity that may evolve to acute respiratory distress syndrome (ARDS), sepsis, and death [3]. Presently, clinical treatment of Covid-19 is mainly symptomatic by using anti-inflammatories like dexamethasone [4] or cytokine inhibitors, combined with antibiotics to treat secondary infections. Knowledge of the mechanism behind SARS-CoV-2 infection will help to identify other therapeutic agents to be used for the treatment of patients with Covid-19. This report focuses in the mechanism of infection and the implication of the renin-angiotensin-aldosterone and the kallikrein/kinin systems in illness progression [5].

The renin-angiotensin-aldosterone system (RAAS) [6] and the kallikrein/kinin system (KKS) [7] are involved in the regulation of intravascular volume, blood pressure and tissue repair via inflammatory and proliferative mechanisms. The angiotensin converting enzyme (ACE) and the angiotensin converting enzyme 2 (ACE2) are key players in both systems. Cross-talk between the two systems is summarized in Figure 1. ACE is a carboxydipeptidase that produces the octapeptide angiotensin II from its inactive precursor angiotensin I, orchestrating a plethora of actions including sodium reabsorption and increase of blood pressure mediated through the AT1 receptor and vasodilation and natriuresis mediated through the AT2 receptor [8]. On the other hand, ACE2 is an integral membrane carboxypeptidase with its catalytic domain at the extracellular side that counterbalances the actions of ACE. Specifically, ACE2 degrades angiotensin II to produce angiotensin (1-7), a peptide that similarly to angiotensin II through the AT2 receptor, elicits vasodilation and natriuresis via activation of the Mas receptor [9]. Furthermore, ACE2 also

converts angiotensin I into angiotensin (1-9), a peptide that elicits vasodilation and anti-inflammatory effects through activation of the AT2 receptor [10]. Angiotensin (1-9) is further converted into angiotensin (1-7) by the action of ACE [8]. On the other hand, kallikreins are serine proteases that produce bradykinin (BK) and kallidin (Lys-BK) -two members of the kinin family- from kininogens in response to inflammation, trauma, burns, shock, allergy and some cardiovascular diseases [11]. Other members of the kinin family include the corresponding desArg^9 analogs: desArg^9-BK and desArg^9-Lys-BK.
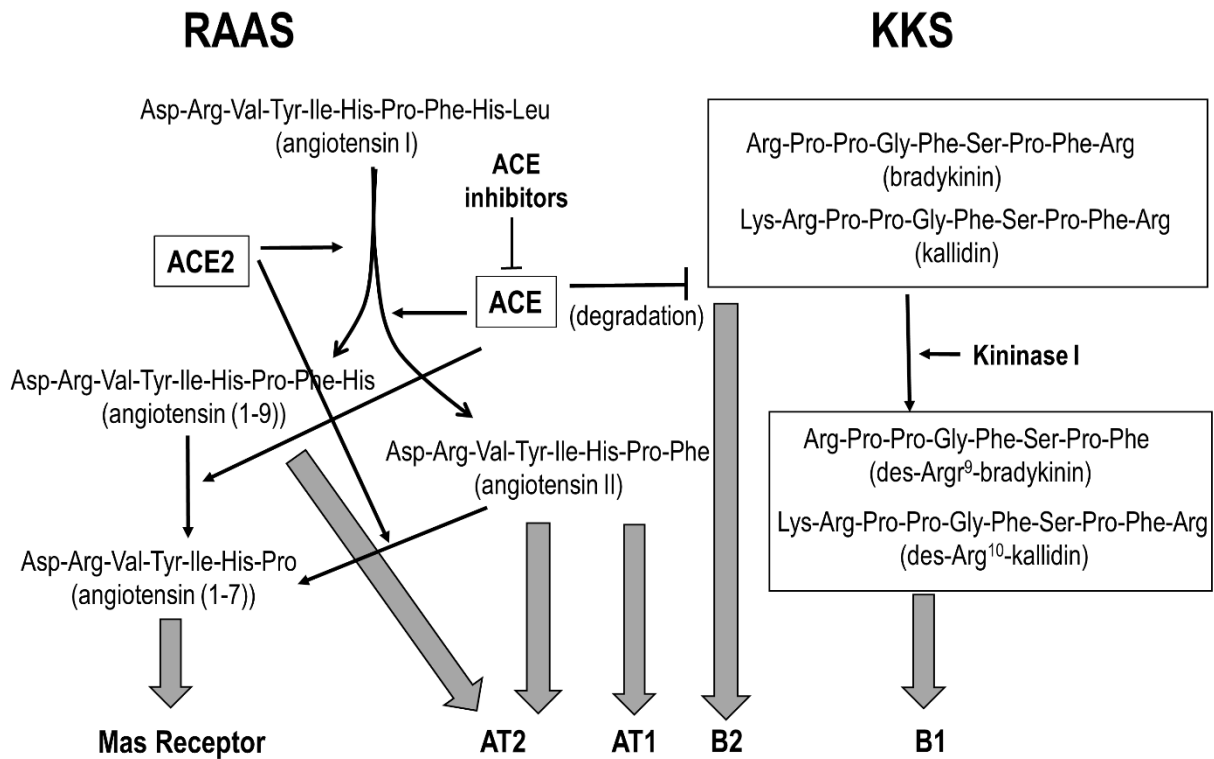

Figure 1. *Cross-talk between RAAS and KKS. The angiotensin converting enzymes ACE and ACE2 are key players of RAAS, regulating the production of diverse mediators (see text), producing a plethora of physiological actions through the activation of different receptors (solid arrows). Thus, activation of the angiotensin AT1 receptor produces vasoconstriction, hypertrophy and fibrosis; whereas activation of the AT2 and Mas receptors produce vasodilation, antihyperthophy and antifibrosis. On the other hand, ACE regulates the levels of kinins that produce vasodilatation and increased vascular permeability through the B1 and B2 receptors.*

Kinins produce a plethora of physiological actions including vasodilatation and increased vascular permeability via activation of the B1 and B2 receptors [12, 13]. The former is upregulated during inflammation episodes or tissue trauma, whereas the latter is constitutively expressed in a variety of cell types. BK binds with high affinity to the B2 receptor, Lys-BK to both B1 and B2, whereas the desArg[9] analogs bind preferably to the B1 receptor [14]. ACE and the ACE2 enzymes are actors involved in the cross-talk between RAAS and KKS. The former upregulates angiotensin II and downregulates BK, whereas the latter upregulates angiotensin (1-9) and downregulates Arg[9]-detached kinins, respectively [15].

SARS-CoV-2 binds with high affinity to ACE2 facilitating cell fusion and entry [16, 17]. Endocytosis of the SARS-CoV-2 spike protein-ACE2 complex into endosomes reduces surface ACE2 expression, being detrimental for its role in tissue protection; producing a clear interference on RAAS mediated homeostasis functions. Taking into account that ACE2 is more abundantly present in the epithelia of the lungs and on lymphocytes [18], its downregulation is a key negative factor for severity of lung edema and acute lung failure observed in patients infected by SARS-CoV. Actually, downregulation of ACE2 translates into altered levels of diverse mediators of the SAARS and KKS. Specifically, levels of angiotensin II and the desArg[9] kinins are increased, whereas levels of angiotensin (1-9) and angiotensin (1-7) are decreased. About fifteen years ago, it was hypothesized that the observed physiological effects produced in patients infected by SARS-CoV were due to the actions of angiotensin II on the AT1 and AT2 receptors [19, 20]. Presently, there is growing evidence that inflammation may be triggered through the desArg[9] peptides/B1 axis-mediated signaling pathway [21- 23]. This new perspective suggests that inhibition of BK signaling may be a suitable therapy to avoid the cytokine storm associated with the Covid-19 illness [24].

Based on this novel mechanistic hypothesis, selective and non-selective BK antagonists should be considered as therapeutic agents for the treatment of covid-19 [25, 26]. Despite the enormous effort devoted in the past to design peptide and non-peptide selective ligands targeting the BK receptors [13, 27], icatibant is the only BK antagonist presently approved as therapeutic agent for the symptomatic treatment of acute attacks of hereditary angioedema in adults with C1-esterase-inhibitor deficiency [28]. Despite being a B2 selective antagonist, the compound is presently involved in a clinical trial to assess its benefits for the treatment of the covid-19 illness [29]. Considering the lack of BK antagonists in the market and the urgency to have new treatments for the covid-19 available, drug repurposing is a valuable strategy for quickly discover novel therapeutic uses of already approved drugs. Specifically in this case, the discovery of approved therapeutic agents with a BK antagonist profile.

Virtual screening methods can be very valuable in drug repurposing, provided we count on specific structural knowledge of the therapeutic target of interest [30]. Specifically, for BK we recently reported the results of a modeling study addressed to analyze the stereochemical features required for non-peptide selective ligands to bind to the B1 and B2 receptors [31, 32]. Furthermore, the results of the study also permitted to identify the stereochemical features associated with selective binding to each of them [33]. As a complement to that work, we discuss in the present contribution the characterization of the molecular features that confer a non-selective binding profile to small molecule ligands targeting the B1 and B2 receptors and show preliminary results from a virtual screening aimed at the identification of approved drugs with a non-selective profile to the BK receptors.

2. Results and Discussion

2.1. Stereochemical features of non-selective small molecule ligands targeting the B1 and B2 receptors

Due to the lack of crystallographic structures of the BK receptors, the construction of 3D models at atomic resolution by homology modeling of the B1 and B2 receptors was recently performed and reported [31, 32]. Models were subsequently used to undertake a docking study that permitted the analysis of diverse ligand-receptor complexes of known selective small molecule compounds. From this study, the corresponding pharmacophores describing the stereochemical features that ligands must fulfill for binding to each of the two receptors were defined. The two receptors share a high sequence identity of 28% (sequence homology is 43%) that reaches ~50% when the orthosteric sites are compared. Accordingly, it is expected to find common residues in their respective orthosteric sites. Comparison of the pharmacophores shows that they exhibit four points in common (Figure 2): a positive charge (P1); a hydrogen bond donor/acceptor (P2); an aromatic ring (P3); and a hydrogen bond donor/acceptor (P4). Point P5 in the two pharmacophores discriminates binding to the two receptors: in B1 is hydrogen bond acceptor, whereas in B2 is hydrophobic/aromatic moiety. Accordingly, ligands fulfilling pharmacophore points P1-P4, common to the two receptors are expected to be non-selective bradykinin antagonists. As a proof of concept, we disclosed a short list of non-selective hits identified through the virtual screening of a large database of small molecules [33]. In contrast, the design of selective ligands is trickier. In addition of the differential chemical nature of P5 in the two receptors, analysis of the 3D models of the ligand-receptor complexes suggests that there is a steric hindrance that prevents selective B1 ligands to

bind B2 and, vice versa. The steric hindrance is produced by the differential nature of the side chains of the non-conserved residues Arg[202] in B1 compared to its counterpart, Thr[197] in B2 [33].
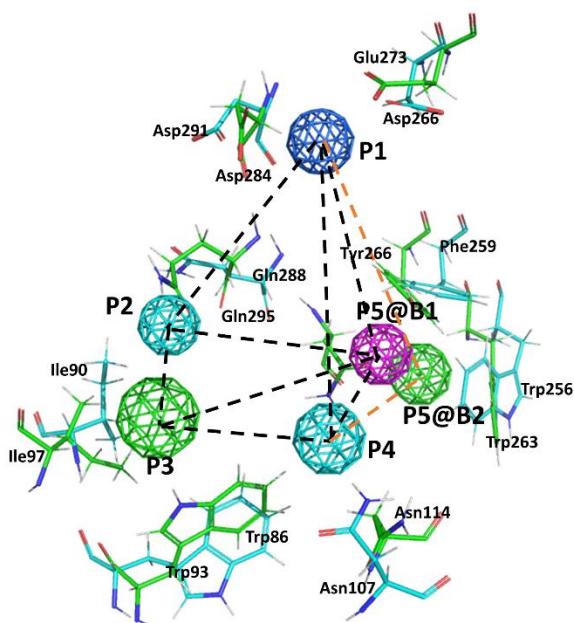

Figure 2. *Superposition of the B1 and B2 receptor pharmacophores. Colored spheres represent pharmacophore points (P1-P5) according to the following color code: dark blue represents a positive charge moiety; magenta a hydrogen bond accepting center; light blue a hydrogen bond donor/acceptor center; green an aromatic/lipophilic center. @B1 and @B2 is used to differentiate P5 for the B1 and B2 receptors, respectively. Consensus distances between common pharmacophore points are (black dotted lines): d(1,2)=9 Å; d(1,3)=14 Å; d(1,4)=10.5 Å; d(2,3)=6 Å; d(2,4)=7 Å; d(3,4)=7.5 Å. Specific distances for the B1 pharmacophore: d(1,5)=9.5 Å; d(2,5)=9.3 Å; d(3,5)=9.5 Å; d(4,5)=5.7 Å; whereas for the B2 pharmacophore are (orange dotted lines): d(1,5)=11 Å; d(2,5)=9 Å; d(3,5)=8.8 Å; d(4,5)=8.4 Å. Side chains of the main residues involved in defining the binding pocket for non-peptide ligands are explicitly depicted: green for the B1 and blue for the B2, respectively.*

2.2. Drug repurposing

Using the four common pharmacophore points P1-P4 as a query, we carried a *in silico* screening on the DrugBank using the Molecular Operating Environment (MOE) program [34]. The DrugBank is a database containing comprehensive information of all FDA approved drugs [35]. In order to carry out the virtual screening study, we first generated the corresponding 3D DrugBank database using the Database

Viewer module of MOE [34]. The database contains for each molecule its 3D structure together with a set of conformations, generated using a build-up procedure from systematic conformational searches of molecular fragments. Virtual screening was carried out on a subset of 1703 molecules selected according to their molecular weight between 200 and 600. Hits obtained were subsequently docked onto the 3D models of the B1 and B2 receptors, respectively to check for possible steric hindrance. Preliminary results of the virtual screening yielded eight drugs (Figure 3): raloxifene [36], a selective estrogen receptor modulator; sildenafil [37], a phosphodiesterase type 5 inhibitor; cefepime [38] and cefpirome [39], two β-lactam antibiotics; imatinib [40] and ponatinib [41], two bcr-Abl tyrosine kinase inhibitors; abemaciclib [42], a dual inhibitor of cyclin-dependent kinases 4 and 6; and entrectinib [43] a non-selective tyrosine kinase inhibitor. According to the results of this study, these compounds exhibit the characteristics to be non-selective BK ligands. Figure 3 shows the location of the common pharmacophore points on their 2D chemical structures. Evaluation of the ability of these compounds to act as B1 and B2 receptor antagonists is currently underway.
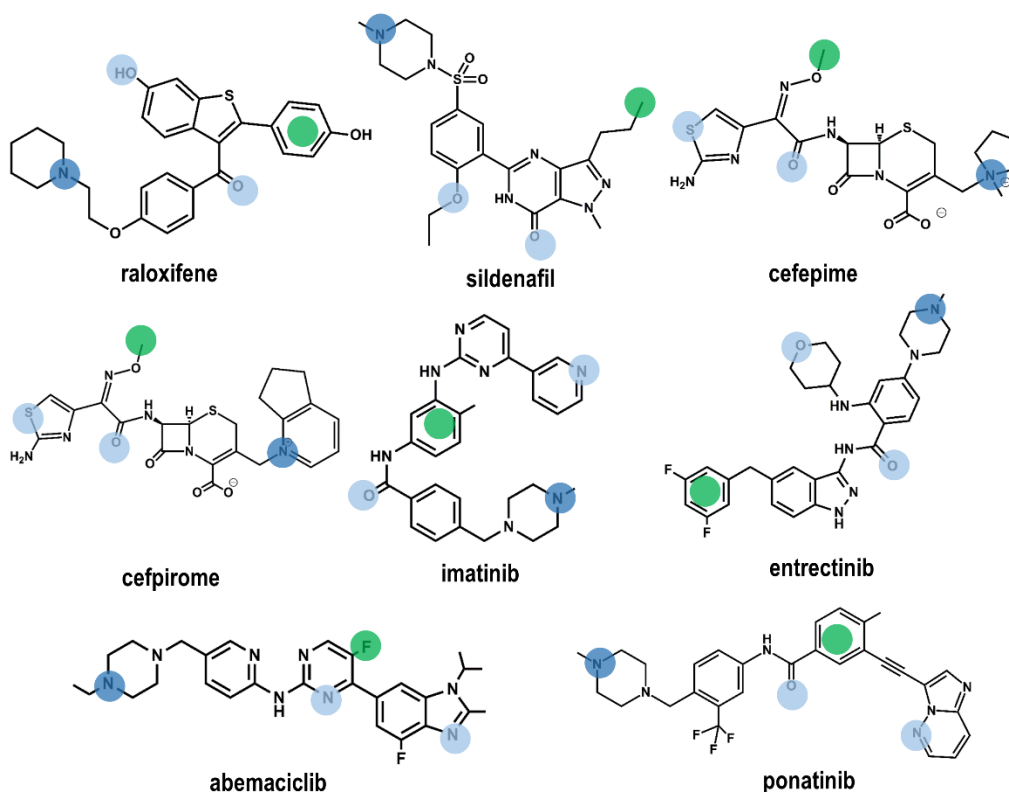

Figure 3. *Chemical structures of the eight compounds identified by virtual screening of the DrugBank database. Color dots are drawn on top of the moieties responsible for fulfilling pharmacophore points P1-P4 according to the color code described in Figure 1.*

Among the drugs identified, there are no studies reporting a direct BK antagonistic profile of any of them. However, in the case of raloxifene there are studies that show a synergistic action with bradykinin. Actually, rats treated with raloxifene show an increased reduction of systolic blood pressure on administration of bradykinin, suggesting an enhanced bioavailability of NO in these animals [44]. Also, sildenafil does not exhibit a synergetic action in the reduction of BK induced glucose uptake in humans when administered together with N^{ω}-monomethyl-L-arginine a nitric-oxide synthase inhibitor [45], indirectly suggesting that sildenafil may not interact with the BK receptors.

3. Concluding remaks

Around 20% of the people infected with SARS-CoV-2 are likely to develop pneumonia of varying severity that may evolve to acute respiratory distress syndrome (ARDS), sepsis, and death. SARS-CoV-2 binds with high affinity to ACE2, mediating cell fusion and entry. ACE2 downregulation was pointed as the origin of the observed inflammatory response of sever cases of Covid-19, mediated by angiotensin II through the AT1 receptor. However, there is an increasing evidence that inflammation is mediated through the bradykinin B1 receptor due to the increased levels of desArg^{9}BK. Accordingly, antagonists of the bradykinin receptors could be useful therapeutic agents to block the inflammatory signaling process. Presently, icatibant is the only bradykinin antagonist approved drug in the market and there are clinical studies in progress to assess its efficacy for the treatment of Covid-19. However, icatibant is a B2 selective antagonist and it is desirable to have also B1 selective or non-selective antagonists available.

In order to have novel therapeutic treatments available in a short time, drug repurposing is a valuable procedure. Repurposing of already approved drugs has several advantages like their known safety/tolerability profiles, availability and low cost. In order to speed up the identification of approved drugs for novel therapeutical uses, virtual screening can be a valuable tool, provided that structural information on the target is available.

As a continuation of a previous study devoted to identify the molecular features required by compounds to exhibit an antagonist profile to the bradykinin receptors, we discussed in the present manuscript those molecular features that provide a non-selective profile to them. These features were used as a query to carry out a virtual screening on the DrugBank, a database containing all approved drugs. The study yielded eight molecules that were subsequently docked onto the 3D models of the B1 and B2

receptors respectively, to check for possible steric hindrance. Evaluation of their profile as bradykinin antagonists is currently under investigation.

References

1. World Health Organization. General's Opening Remarks at the Media Briefing on COVID-19-18 March **2020**; World Health Organization: Geneva, Switzerland, **2020**.

2. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). Available on-line at: https://coronavirus.jhu.edu/map.html. (accessed on July, 18th **2020**).

3. Wu, Z.; McGoogan, J.M. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. *J. Am. Med. Assoc.* **2020**, 323, 1239-1242.

4. Horby, P.; Lim, W. S.; Emberson, J. R.; Mafham, M.; Bell, J. L.; Linsell, L.; Staplin, N.; Brightling, C.; Ustianowski, A.; Elmahi, E.; Prudon, B.; Green, C.; Felton, T.; Chadwick, D.; Rege, K.; Fegan, C.; Chappell, L. C.; Faust, S. N.; Jaki, T.; Jeffery, K.; Montgomery, A.; Rowan, K.; Juszczak, E.; Baillie, J. K.; Haynes, R.; Landray. M. J. Dexamethasone in Hospitalized Patients with Covid-19 - Preliminary Report. *N. Engl. J. Med.* **2020**; 10.1056/NEJMoa2021436. doi:10.1056/NEJMoa2021436. Online ahead of print.

5. van de Veerdonk, F.; Netea, M.G.; van Deuren, M.; van der Meer, J.W.; de Mast, Q.; Bruggemann, R.J.; van der Hoeven, H. Kinins and Cytokines in COVID-19: A Comprehensive Pathophysiological Approach. *Preprints* **2020**, 2020040023 (doi: 10.20944/preprints202004.0023.v1).

6. Crowley, S. D.; Coffman, T. M. Recent advances involving the renin–angiotensin system. *Exp. Cell Res.* **2012**, *318*, 1049-1056.

7. Schmaier, A. H. The plasma kallikrein-kinin system counterbalances the renin-angiotensin system. *J. Clin. Invest.* **2002**, *109*, 1007–1009.

8. Bernstein, K. E.; Ong, F. S.; Blackwell, W.-L.; Kandarp, B.; Shah, H.; Giani, J. F.; Gonzalez-Villalobos, R. A.; Shen, X. Z.; Fuchs, S. A Modern Understanding of the Traditional and Nontraditional Biological Functions of Angiotensin-Converting Enzyme. *Pharmacol Rev.* **2013**, *65*, 1-46.

9. Santos, R. A.; Simoes e Silva, A.C.; Maric, C.; Silva, D.M.; Machado, R.P.; de Buhr, I.; Heringer-Walther, S.; Pinheiro, S.V.; Lopes, M.T.; Bader, M.; Mendes, E.P.; Lemos, V.S.; Campagnole-Santos, M.J.;

Schultheiss, H.P.; Speth, R.; Walther, T. Angiotensin-(1-7) is an endogenous ligand for the G protein-coupled receptor Mas. *Proc. Natl. Acad. Sci. USA.* **2003**, *100*, 8258–8263.

10. Flores-Munoz, M.; Work, L.M.; Douglas, K.; Denby, L.; Dominiczak, A.F.; Graham, D.; Nicklin, S. Angiotensin-(1–9) attenuates cardiac fibrosis in the stroke-prone spontaneously hypertensive rat via the angiotensin type 2 receptor. *Hypertension*. **2012**, 59, 300–307.

11. Golias, C.; Charalabopoulos, A.; Stagikas, D.; Charalabopoulos, K.; Batistatou, A. The kinin system - bradykinin: biological effects and clinical implications. Multiple role of the kinin system-bradykinin. *Hippokratia*. **2007**, 11, 124-128.

12. Regoli, D.; Barabe, J. Pharmacology of bradykinin and related kinins. *Pharmacol. Rev*. **1980**, *32*, 1-46.

13. Marceau, F.; Bachelard, H.; Bouthillier, J.; Fortin, Morissette, J.-P. G.; Bawolak, M.-T.; Charest-Morin, X.; Gera, L. Bradykinin receptors: Agonists, antagonists, expression, signaling, and adaptation to sustained stimulation. Int. Immunopharmacol. **2020**, *82*, 106305.

14. Leeb-Lundberg, L. M. F.; Marceau, F.; Muller-Ester, W.; Pettibone, D. J.; Zuraw, B. L. International Union of Pharmacology. XLV. Classification of the Kinin Receptor Family: from Molecular Mechanisms to Pathophysiological Consequences. *Pharmacol. Rev.* **2005**, *57*, 27–77.

15. Su, B. J. Different cross-talk sites between the renin–angiotensin and the kallikrein–kinin systems *J. Renin-Angiotensin-Aldosterone Syst.* **2014**, *15*, 319–328.

16. Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N.-H.; Nitsche, A.; Müller, M. A.; Drosten, C.; Pöhlmann, S. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*. **2020**, *181*, 1–10.

17. Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K.-Y.; Wang, Q.g; Zhou, H.; Yan, J.; Qi, J. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell*. **2020**, *181*, 894-904.

18. Hamming, I.; Timens, W.; Bulthuis, M.; Lely, A.; Navis, G.; van Goor, H. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J. Pathol.* **2004**, *203*, 631–637.

19. Imai, Y.; Kuba, K.; Rao, S.; Huan, Y.; Guo, F.; Guan, B.; Yang, P.; Sarao, R.; Wada, T.; Leong-Poi, H.; Crackower, M. A.; Fukamizu, A.; Hui, C.-C.; Hein, L.; Uhlig, S.; Slutsky, A. S.; Jiang, C.; Penninger, J. M. Angiotensin-converting enzyme 2 protects from severe acute lung failure. *Nature*. **2005**, *436*, 112-116.

20. Ni, W.; Yang, X.; Yang, D.; Bao, J.; Li, R.; Xiao, Y.; Hou, C.; Wang, H.; Liu, J.; Yang, D.; Xu, Y.; Cao, Z.; Gao, Z. Role of angiotensin-converting enzyme 2 (ACE2) in COVID-19. *Crit. Care*. **2020**, *24*, 422.

21. Sodhi, C. P.; Wohlford-Lenane, C.; Yamaguchi, Y.; Prindle, T.; Fulton, W. B.; Wang, S.; McCray, Jr.; P. B.; Chappell, M.; Hackam, D. J.; Jia, H. Attenuation of pulmonary ACE2 activity impairs inactivation of des-Arg9 bradykinin/BKB1R axis and facilitates LPS-induced neutrophil infiltration. *Am. J. Physiol. Lung Cell Mol. Physiol.* **2018**, *314*, L17–L31.

22. van de Veerdonk, F. L.; Netea, M. G.; van Deuren, M.; van der Meer, J. WM.; de Mast, Q.; Brüggemann, R. J.; van der Hoeven, H. Kallikrein-kinin blockade in patients with COVID-19 to prevent acute respiratory distress syndrome. *eLife.* **2020**, *9*, e57555.

23. de Maat, S.; de Mast, Q.; Danser, A.H. J.; van de Veerdonk, F. L.; Maas, C. Impaired Breakdown of Bradykinin and Its Metabolites as a Possible Cause for Pulmonary Edema in COVID-19 Infection. *Semin. Thromb. Hemost.* **2020**, doi: 10.1055/s-0040-1712960. Online ahead of print.

24. Mahmudpour, M.; Roozbeh, J.; Keshavarz, M.; Farrokhi, S.; Nabipour, I. COVID-19 cytokine storm: The anger of inflammation. *Cytokine.* **2020**, *133*, 155151.

25. Rameshrad, M.; Ghafoori, M.; Mohammadpour, A. H.; Nayeri, M. J. D.; Hosseinzadeh, H. A comprehensive review on drug repositioning against coronavirus disease 2019 (COVID19). *Naunyn-Schmiedeberg's Archiv. Pharmacol.* **2020**, *393*, 1137–1152.

26. Tolouian, R.; Vahed, S. Z.; Ghiyasvand, S.; Tolouian, A.; Ardalan, M. COVID-19 interactions with angiotensin-converting enzyme 2 (ACE2) and the kinin system; looking at a potential treatment. *J. Renal Inj. Prev.* **2020**, *9*, e19.

27. Marceau, F.; Regoli, D. Bradykinin Receptor Ligands: Therapeutic Perspectives. *Nat. Drug Discov.* **2004**, *3*, 845-852.

28. Bork, K.; Yasothan, U.; Kirkpatrick, P. Icatibant. *Nat. Rev. Drug Discov*. **2008**, *7*, 801-802.

29. https://clinicaltrials.gov/ct2/show/NCT04488081?term=icatibant&draw=3&rank=19 (accessed on July, 18th **2020**).

30. Sree Gns, H.; GR, S.; Muraharia, M.; Krishnamurthy, M. An update on Drug Repurposing: Re-written saga of the drug's fate. *Biomed. Pharmacother.* **2019**, *110*, 700-716.

31. Lupala, C.L.; Gomez-Gutierrez, P.; Perez, J.J. New insights into the stereochemical requirements of the bradykinin B1 receptor antagonist binding. *J. Mol. Graphics Model.* **2016**, *68*,184–196.

32. Lupala, C.L.; Gomez-Gutierrez, P.; Perez, J.J. New insights into the stereochemical requirements of the bradykinin B2 receptor antagonist binding. *J. Comp.-Aided Mol. Design*, **2016**, *30*, 85-101.

33. Rasaeifar, B.; Lupala, C. S.; Gomez-Gutierrez, P.; Perez, J. J. Molecular Features Characterizing Non-peptide B1 and B2 Bradykinin Receptor Selectivity. *Bioorg. Med. Chem. Lett.;* **2019**, *29*, 11-14.

34. *Molecular Operating Environment (MOE) 2019.01; Chemical Computing Group UCL*: Montreal, QC, Canada. **2020**.

35. Wishart, D.S.; Feunang, Y.D.; Guo, A.C.; Lo, E.J.; Marcu, A.; Grant, J.R.;, Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.;, Assempour, N.; Iynkkaran, I,, Liu, Y.; Maciejewski, A.; Gale, N.; Wilson, A.; Chin, L.; Cummings, R.; Le, D.; Pon, A.; Knox, C.; Wilson, M. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. **2018**, *46*, D1074-D1082.

36. Scott, J.A.; Da Camara, C.C.; Early, J.E. Raloxifene: a selective estrogen receptor modulator. *Am. Fam. Physician.* **1999**, *60*, 1131-1139.

37. Unegbu, C.; Noje, C.; Coulson, J.D.; Segal, J.B.; Romer L. Pulmonary Hypertension Therapy and a Systematic Review of Efficacy and Safety of PDE-5 Inhibitors. *Pediatrics*. **2017**, *139*, e20161450.

38. Patel, H.B.; Lusk, K.A.; Cota, J.M. The Role of Cefepime in the Treatment of Extended-Spectrum Beta-Lactamase Infections. *J. Pharm. Pract.* **2019,** *32*, 458-463.

39. Hamazaki , H.; Hasegawa, H.; Horiuchi, A.; Teshima, H.; Hiraoka, A.; Masaoka, T.; Nasu, K.; Uchino, H.; Tatsumi, N.; Inoue, N.; Kageyama, T.; Kawagoe, H.; Tukaguchi, M.; Hukuhara, S.; Takahashi, T.; Takatsuka, H.; Kanamaru, A.; Kakishita, E.; Nagai, K.; Hara, H.; Kanayama, Y.; Sugiyama, H.; Kitani, T. Clinical evaluation of cefpirome sulfate for severe infections in patients with hematological disorders. Hanshin Study Group of Hematopoietic Disorders and Infections. *Jpn. J. Antibiot.* **1997**, *50*, 12-21.

40. Vigneri, P.; Wang, J.Y. Induction of apoptosis in chronic myelogenous leukemia cells through nuclear entrapment of BCR-ABL tyrosine kinase. *Nat. Med.* **2001**, *7*, 228-234.

41. Zhou, T.; Commodore, L.; Huang, W.S.; Wang, Y.; Thomas, M.; Keats, J.; Xu, Q.H.; Rivera, V.M.; Shakespeare, W.C.; Clackson, T.; Dalgarno, D.C.; Zhu, X.T. Structural mechanism of the Pan-BCR-ABL inhibitor ponatinib (AP24534): lessons for overcoming kinase inhibitor resistance. *Chem. Biol. Drug Des.* **2011**, *77*, 1–11.

42. Gelbert, L.M.; Cai, S.; Lin, X.; Sanchez-Martinez, C.; Del Prado, M.; Lallena, M.J.; Torres, R.; Ajamie, R.T.; Wishart, G.N.; Flack, R.S.; Neubauer, B.L.; Young, J.; Chan, E.M.; Iversen, P.; Cronier, D.; Kreklau, E.; de Dios, A. Preclinical characterization of the CDK4/6 inhibitor LY2835219: in-vivo cell cycle-dependent/independent anti-tumor activities alone/in combination with gemcitabine. *Invest. New Drugs.* **2014**, *32*, 825-37.

43. Rolfo, C.; Ruiz, R.; Giovannetti, E.; Gil-Bazo, I.; Russo, A.; Passiglia, F.; Giallombardo, M.; Peeters, M.; Raez, L. Entrectinib: a potent new TRK, ROS1, and ALK inhibitor. *Expert Opin. Investig. Drugs.* **2015**, *24*, 1493-1500.

44. Wassmann, S.; Laufs, U.; Stamenkovic, D.; Linz, W.; Stasch, J.-P.; Ahlbory, K.; Rösen, R.; Böhm, M.; Nickenig, G. Raloxifene Improves Endothelial Dysfunction in Hypertension by Reduced Oxidative Stress and Enhanced Nitric Oxide Production. *Circulation*. **2002**, *10*, 2083-2091.

45. Pretorius, M.; Brown, N. J. Endogenous Nitric Oxide Contributes to Bradykinin-Stimulated Glucose Uptake but Attenuates Vascular Tissue-Type Plasminogen Activator Release. *J. Pharm. Exp. Ther.* **2010**, *332*, 291–297.

46. Wu, B.; Mol, C. D.; Han, G. W.; Katritch, V.; Chien, E. Y. T.; Liu, W.; Cherezov, V.; Stevens, R. C. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science*. **2010**, *330*, 1066-1071.

47. Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815.

48. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J. GROMACS: Fast flexible and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718.

49. Lupala, C.S.; Rasaifar, B.; Gomez-Gutierrez, P.; Perez, J.J. Using Molecular Dynamics for the refinement of atomistic models of GPCRs by homology modeling. *J. Biomol. Struct. Dyn.* **2018**, *36*, 2436–2448.

50. Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.; Shaw, D.E.; Shelley, M.; et al. Glide: A New Approach for Rapid Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.

Conclusions

1. The results of this work support the use of molecular dynamics for the refinement of atomistic models of GPCRs constructed by homology modeling. This can be assessed by inspecting the rmsd time evolution of the transmembrane regions of the models. Moreover, the results also support that the closer in the phylogenetic tree the template used is, the higher the accuracy of the model obtained for the same sampling time.

2. The presence of a ligand in the refinement process using molecular dynamics makes refinement time to be shorter, yielding a more accurate model of the transmembrane region.

3. Comparison of atomistic models of the bradykinin receptors B1 and B2, constructed by homology modeling permitted to define five-point pharmacophores for these receptors. Comparison of these pharmacophores suggests that four of the points are common to the two receptors and can be used to design non-selective antagonists, whereas fulfillment of the fifth pharmacophore point provides selectivity to the ligands.

4. The fifth pharmacophore point of the bradykinin receptors exhibits a differential chemical nature. In B1R is a hydrogen bond accepting center, whereas in the B2R is an aromatic/lipophilic center. Ligands are selective not only because the differential nature of this pharmacophore point, but because of steric features of its location. Specifically, there is steric hindrance to ligands due to the differential nature of the side chains of the non-conserved residues Arg202 in B1 compared to its counterpart, Thr197 in B2.

5. The results of a virtual screening of a small molecule database suggests that fulfillment of the fifth pharmacophore point of the B1 and B2 receptors is a necessary condition for designing selective ligands, but not sufficient, since ligands need to be bulky enough to bind to the receptor in a limited number of poses.

6. Construction of an atomistic model of the BB1 bombesin receptor by homology modeling, followed by docking studies of the peptoid ligands PD168368 and PD176252 permitted to define a pharmacophore for receptor antagonism. This was used a query in a virtual screening of a small molecule database that yielded a few selective ligands in the low micromolar range.

7. Construction of an atomistic model of the BB2 bombesin receptor by homology modeling, followed by docking studies of the peptoid ligand PD176252 permitted to understand different possible binding modes of the ligands. This study provided useful information for further site-mutagenesis studies. Specifically, the results of the mutation of Arg^{308} or Arg^{287} for non-polar residues will be useful to define a pharmacophore.

8. Construction of an atomistic model of the BB3 bombesin receptor by homology modeling, followed by docking studies of the peptoid ligand AM-37 permitted to understand different possible binding modes of the ligands. This study provides useful information for further site-mutagenesis studies. Specifically, mutation of Ser^{205} and Ser^{124} for Pro would return to BB1 and BB2 showing an increased binding affinity to these receptors. In addition, mutation of Asp^{104}, Tyr^{108} and Arg^{316} will provide clues for the differential binding experienced by ST-36 in regard the PD-compounds.

9. A virtual screening study on the DrugBank using as query the common points of the B1 and B2 receptors pharmacophores, permitted to identify eight available drugs that prospectively could be used as non-selective bradykinin antagonists. This list includes raloxifene; sildenafil; cefepime; cefpirome; imatinib; ponatinib; abemaciclib and entrectinib. Tests of their bradykinin antagonist potency are currently underway.

Annex

Molecular features characterizing non-peptide selectivity to the human B1 and B2 bradykinin receptors

B. Rasaeifar[a], C.S. Lupala[b], P. Gomez-Gutierrez[a,c], Juan J. Perez[a,*]

[a] Dept. of Chemical Engineering, Universitat Politecnica de Catalunya, Av. Diagonal, 647, 08028 Barcelona, Spain
[b] Beijing Computational Science Research Center, Building 9, East Zone, ZPark II, No. 10 Xibeiwang East Road, Haidian District, Beijing 100193, China
[c] Allinky Biopharma, Madrid Scientific Park, Faraday, 7, 28049 Madrid, Spain

ABSTRACT

Bradykinin is produced in response to inflammation, trauma, burns, shock, allergy and some cardiovascular diseases. Actions of this peptide are mediated through two different G-protein coupled receptors, named B1 and B2 that have different pharmacological characteristics. The former is up-regulated during inflammation episodes or tissue trauma whereas, the latter is constitutively expressed in a variety of cell types. In a previous work we have characterized the molecular features that explain the observed structure-activity results for both receptors by means of molecular modeling studies, using diverse ligands for both receptors. These results were summarized in the form of two different pharmacophores that provided new insights to be used for the design of novel molecules with antagonistic profile. In the present work, we compare these pharmacophores to understand the features that characterize ligand selectivity to the two bradykinin receptors. The study shows that most of the residues involved in the binding pocket are similar in both receptors and consequently are the pharmacophores obtained. The main difference between the two pharmacophores remains on point #5 that involves a polar moiety for the B1 receptor and an aromatic ring for the B2 receptor. Accordingly, analysis of the prospective bound conformation of several non-selective small molecule ligands of the bradykinin receptors permits to conclude that fulfilment of point#5 is a requirement to produce selective ligands. However, the study also shows that this is a necessary condition only, since ligands need also to be bulky enough to avoid binding to these receptors in diverse poses. These results provide new insights for a better understanding of the molecular features that ligands are required to exhibit to be selective bradykinin ligands.

Bradykinin (BK), a nonapeptide of sequence Arg^1-Pro^2-Pro^3-Gly^4-Phe^5-Ser^6-Pro^7-Phe^8-Arg^9 is a member of the kinins, a group of peptides ubiquitously produced by the action of kallikreins on circulating kininogens. Other members of the kinin family include kallidin (KD) (Lys^0-BK) and the metabolites of KD and BK: $desArg^9$-BK and $desArg^9$-KD. Kinins are produced in response to inflammation, trauma, burns, shock, allergy and some cardiovascular diseases, provoking changes in blood pressure and vasodilation, increased vascular permeability, stimulation of sensory neurons, vascular and bronchial smooth muscle contraction, intestinal ion secretion, release of prostaglandins and cytokines, and the production of nitric oxide.[1,2] Pharmacological actions of this family of compounds are mediated by at least two G-protein coupled receptors, named B1 and B2. The former is up-regulated during inflammation episodes or tissue trauma whereas, the latter is constitutively expressed in a variety of cell types. Members of the kinin family show a diverse pharmacological profile. Thus, whereas BK and KD exhibit higher affinity for the B2 receptor, the $desArg^9$ metabolites bind only to the B1 receptor, with $desArg^9$-KD being a potent B1 receptor agonist.[2]

Despite the cardioprotective and neuroprotective action of bradykinin in various forms of ischemic conditioning,[3] due to their role in mediating pain and inflammation a great deal of interest was devoted in the past to the identification of potent kinin antagonists for therapeutical intervention. The first generation of selective antagonists described in the literature consisted of linear peptides including D-Arg-[Hyp^3,D-Phe^7]-BK (NPC-567) (Hyp = hydroxyproline); D-Arg-[Hyp^3,$Thi^{5,8}$,D-Phe^7]-BK (NPC-349) (Thi = thienylalanine); or D-Arg-[Hyp^3,D-Phe^7,Leu^8]-BK[4,5] that provided the basis for the discovery of analogs more resistant to metabolism and conformationally constrained.[6] These studies led to the discovery of Icatibant, a B2 receptor antagonist currently in the market used for the symptomatic treatment of acute attacks of hereditary angioedema in adults with C1-esterase-inhibitor deficiency.[7] Over the time a number of small molecule

---

* Corresponding author.
  E-mail address: juan.jesus.perez@upc.edu (J.J. Perez).

compounds antagonists, selective to human B1 and B2 were also disclosed.[8] All of these compounds can be described as bulky molecules with high molecular weight.

In order to understand the minimal requirements that small molecules must have to exhibit bradykinin antagonism, we recently undertook a modeling study aimed at identifying the molecular features that characterize human B1 and B2 receptor-antagonist complexes. For this purpose, since the crystallographic structure of none of the two receptors is available, we built atomistic models of the human B1 and B2 receptors by homology modeling and carried out docking studies,[9,10] using diverse selective small molecule binders described in the literature.[8] Analysis of the resulting receptor-ligand models permitted to understand the way diverse selective and non-selective small molecule ligands bind to the bradykinin receptors, suggesting that they do not necessarily bind as peptide ligands do, as shown before.[11] Furthermore, these studies permitted the development of pharmacophoric hypothesis for the human B1 and B2 receptor antagonism that were validated by an *in silico* screening study that permitted the identification of a few novel small molecules with antagonistic activity.[12,13]

Comparative analysis of the two pharmacophores can also be used further to provide information about the molecular features that confer ligand selectivity to each of the bradykinin receptors. Accordingly, we report in the present study the results of such analysis and provide a rationale for the pharmacological profile of diverse ligands, including those disclosed in our previous reports identified from an *in silico* screening process.[14]

Comparative analysis of the 3D structures of the human B1 and B2 receptors using the homology models previously described,[11,12] reveals a high identity of residues within the binding pocket of the two receptors (~50%), as anticipated from the alignment of their sequences.[15] However, comparison of the two pharmacophores reveals that although points #1–4 are common, point #5 exhibits differential features regarding to its chemical nature and location, as shown in Fig. 1. Specifically, point #5 is represented by a polar moiety in the human B1 receptor pharmacophore that is complementary to the space defined by residues Arg[202] (R5.38), Tyr[266] (Y6.51) and Asn[298] (N7.39) (Ballesteros-Weinstein notation[16] is included in parenthesis for the sake of clarity). In contrast, point #5 in B2 is defined by an aromatic/hydrophobic moiety located in a deeper region of the orthosteric site that interacts with residues of a hydrophobic cleft comprising Trp[256] (W6.48), Phe[259] (F6.51) and Tyr[295] (Y7.43).

A detailed analysis of the two receptors reveals that the reason of

the different location and nature of point #5 can be attributed in part, to the difference between Arg202 (R5.38) in B1 and its counterpart Thr197 (T5.38) in B2. Actually, the side chain of Arg202 (R5.38) points towards the center of the binding pocket and forms a hydrogen bond with Tyr266 (Y6.51) that together with the conserved Asn298 (N7.39), define pharmacophoric point #5 of B1. The differential nature of Thr197 (T5.38) and Phe259 (F6.51) (counterpart of Tyr266 (Y6.51)) makes impossible to create the same environment in B2. Moreover, the smaller size of Thr197 (T5.38) in B2 permits ligands accessing a region of aromatic residues defined by Trp256 (W6.48), Phe259 (F6.51) and Tyr295 (Y7.43). This environment, inaccessible in B1 permits to define a different pharmacophoric point #5 in B2. Another difference between the two receptors regards the differential features of residues Lys[118] (K3.33) in the B1 receptor and its counterpart Ser[111] (S3.33) in B2, occupying the same spot in transmembrane helix 3 (TM3). The former exhibits a hydrogen bond with the side chain of residue Glu[205] (E5.41), whereas its counterpart is Leu[207] (L5.41) in B2. The lack of this charge-charge interaction in B2 is responsible for TM5 to appear displaced in regard to the corresponding helix in B1, reducing the solvent-accessible surface area of the receptor. Interestingly, Lys[118] (K3.33) in B1 is key to explain the selectivity profile of the diverse kinins.[17] Recent Solid State NMR studies show evidence that Lys[118] (K3.33) interacts with the carboxyl C-terminal group of the B1 selective analogs desArg[9]-BK and desArg[9]-KD and forces the C-terminus of BK and KD to adopt a distinct orientation to avoid a repulsive interaction with the Arg[9] side chain of the peptides.[18]

According to the analysis carried out, compounds that only fulfil pharmacophoric points #1–4 are expected to be non-selective bradykinin antagonists, being point #5 responsible for conferring receptor selectivity to the ligands. Actually, all selective antagonists used for pharmacophore development fulfil at least four of the five pharmacophoric points, including point#5.[12,13] More instructive is to analyze the prospective binding mode and selectivity profile of a set of non-selective small molecule compounds disclosed in previous studies, discovered from an *in silico* screening process.[12,13] Chemical structures of the molecules studied are shown in Fig. 2 and their antagonistic activity to B1 and B2 is listed in Table 1.

We proceeded to dock these molecules onto the human B1 and B2 receptor models constructed by homology modeling as explained in the methods section, to check the number of pharmacophoric points fulfilled. Table 1 shows the number of pharmacophoric points fulfilled by the molecules to each of the two receptors. Analysis of Table 1 suggests
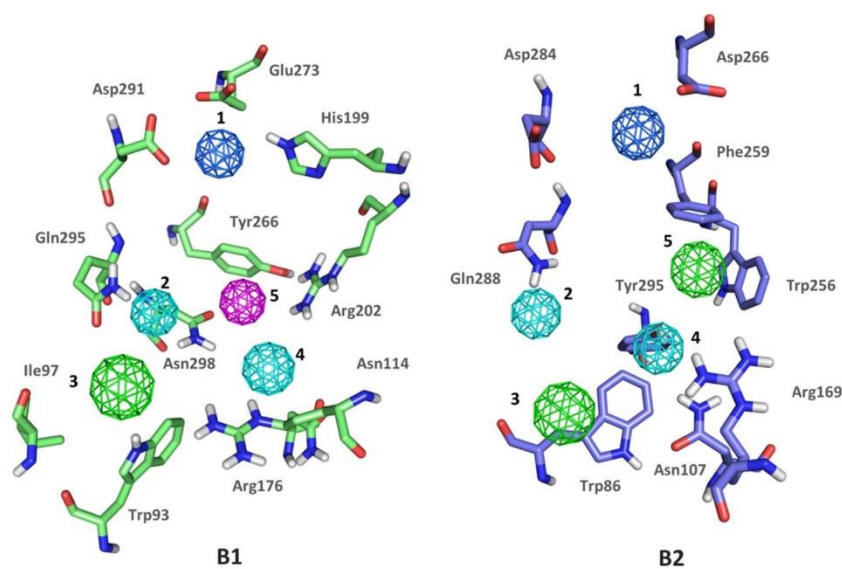

Fig. 1. Pictorial view of the binding pocket of the human bradykinin B1 (left) and B2 (right) receptors. Side chains of the main residues involved in defining the binding pocket for non-peptide ligands are explicitly depicted. Colored balls represent the pharmacophoric points that should fulfill the ligands with the following color code: dark blue represents a positive charge moiety; magenta a proton accepting center; light blue a proton donor center; light green a lipophilic center and dark green an aromatic center.
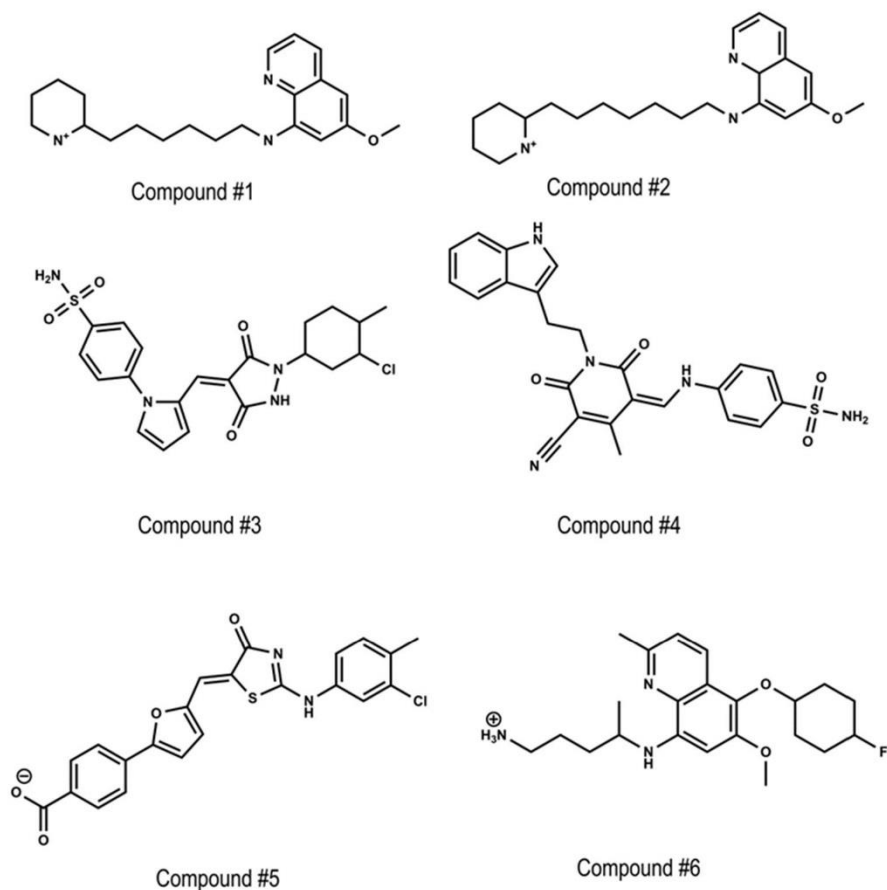
Fig. 2. Chemical structures of the small molecule bradykinin antagonists discovered in an in silico screening.

Table 1
Pharmacological profile and pharmacophoric points fulfilled for the set of compounds discovered in an in silico screening described previously.[12,13]

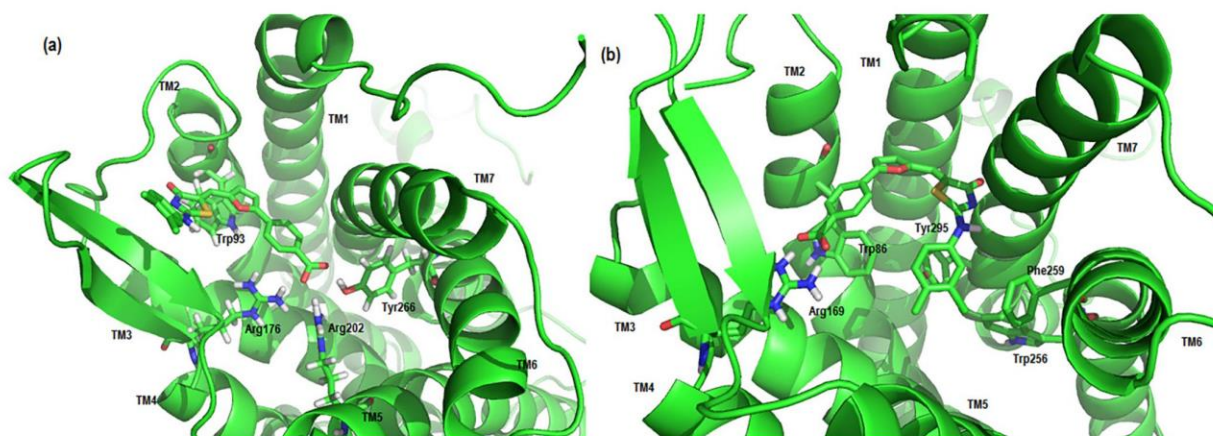| compound	B1 inhibition	B2 inhibition	Point#1	Point#2	Point#3	Point#4	Point#5 @B1	Point#5 @B2
#1	31%@50uM	36%@50uM	✓		✓			
#2	46%@50uM	65%@50uM	✓		✓			
#3	31%@10uM	41%@50uM	✓	✓	✓			
#4	32%@10uM	34%@10uM	✓	✓		✓		
#5	14%@10uM	33%@10uM			✓	✓	✓	✓
#6	12%@50uM	45%@50uM	✓		✓			

Fig. 3. Prospective bound conform of compound#5 to the bradykinin receptors. a) bound conformation into receptor B1; b) bound conformation into receptor B2.

that the antagonistic activity correlates well with the number of pharmacophoric points fulfilled. Moreover, the lack of selectivity of most of the compounds is due to their fulfillment of a subset of pharmacophoric points #1–4, as expected. Interestingly, compound #5 is non-selective, but fulfils point #5 in each of the two pharmacophores. This means that even if ligands do have moieties fulfilling pharmacophoric point#5 they can be non-selective which suggests that fulfilling point #5 is a necessary condition, but not sufficient, to get selective compounds. Inspection of its prospective bound conformation of compound#5 in the two receptors produced by docking studies, shows that the molecule is small enough to be accommodated in different poses in the two receptors, as shown in Fig. 3. In B1 the ligand carboxyl moiety interacts with Arg^{202} (R5.38) (Fig. 3a), fulfilling pharmacophoric point#5, whereas in B2 the same chemical group faces Arg^{169} (R34.57) (Fig. 3b), fulfilling pharmacophoric pont#4. In addition, in B1 the molecule fulfills pharmacophoric points#3 and #4, whereas in B2 it fulfills pharmacophoric points#3 and #5. This suggests that in order to be selective, I addition to fulfill pharmacophoric point#5, ligands need to be bulky enough to avoid binding to the receptors in diverse poses. The most obvious way for molecules to obey such a condition is the fulfillment of an additional pharmacophoric point. Actually, all the ligands that were used in pharmacophore development fulfill at least four pharmacophoric points,[12,13] whereas the most active compounds of Table 1 only fulfill three of them. This suggests that in order to get selective compounds, molecules should fulfill at least four pharmacophoric points, being one of these pharmacophoric point#5.

In the present study we report the results of a comparative analysis of the B1 and B2 bradykinin 3D models. This analysis permitted to identify some structural features that provide ligand selectivity. Specifically, the two pharmacophores have points #1–4 in common, being point #5 the one that exhibits differential features regarding to its chemical nature and location and that provides ligand selectivity. Analysis of the models reveals that the reason of the different location and nature of point #5 can be attributed in part, to the difference between Arg^{202} (R5.38) in B1 and its counterpart Thr^{197} (T5.38) in B2. Unfortunately, there are no experimental results that can corroborate this finding. However, an analysis of a set of non-selective small molecules disclosed in previous studies,[12,13] identified from an *in silico* study using the two pharmacophores, permits to give support to such an hypothesis. The reason of the non-selective profile of five of the compounds can be easily explained to be due to their fulfilment of a number of the common pharmacophoric points #1–4. However, compound #5 represents a special case. It fulfills pharmacophoric point #5 in the two receptors due to its capacity to bind in different poses in the two receptors. This permits us to conclude that fulfilling point #5 is a necessary condition to get selective compounds but it is not sufficient,

since ligands need also to be bulky to avoid binding to the receptors in diverse poses, being the most obvious way for molecules to obey such a condition, the fulfillment of an additional pharmacophoric point.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bmcl.2018.11.026.

References

1. Regoli D, Barabe J. Pharmacology of bradykinin and related kinins. *Pharmacol Rev.* 1980;32:1–46.
2. Leeb-Lundberg LMF, Marceau F, Muller-Esterl W, Pettibone DJ, Zuraw BL. International union of pharmacology. XLV. classification of the Kinin Receptor Family: from molecular mechanisms to pathophysiological consequences. *Pharmacol Rev.* 2005;57:27–77.
3. Sharma R, Randhawa PK, Singh N, Jaggi AS. Bradykinin in ischemic conditioning-induced tissue protection: evidences and possible mechanisms. *Eur J Pharmacol.* 2015;768:58–70.
4. Steranka LR, Farmer SG, Burch RM. Antagonists of B2 Bradykinin Receptors. *FASEB J.* 1989;3:2019–2025.
5. Stewart JM. Bradykinin antagonists: discovery and development. *Peptides.* 2004;25:527–532.
6. Marceau F, Regoli D. Bradykinin receptor ligands: therapeutic perspectives. *Nat Rev Drug Discovery.* 2004;3:845–852.
7. Bork K, Yasothan U, Icatibant Kirkpatrick P. *Nat Rev Drug Discovery.* 2008;7:801–802.
8. Dziadulewicz ED. Non-peptide ligands for bradykinin receptors 1995–2004. *Expert Opin Ther Patents.* 2005;15:829–859.
9. Cordomi A, Edholm O, Perez JJ. Effect of different treatments of long-range interactions and sampling conditions in molecular dynamic simulations of rhodopsin embedded in a dipalmitoyl phosphatidylcholine bilayer. *J Comput Chem.* 2007;28:1017–1030.
10. Lupala CS, Rasaifar B, Gomez-Gutierrez P, Perez JJ. Using Molecular Dynamics for the refinement of atomistic models of GPCRs by homology modeling. *J Biomol Struct Dyn.* 2018;36:2436–2448.
11. Meini S, Cucchi P, Belluccia F, et al. Site-directed mutagenesis at the human B2 receptor and molecular modelling to define the pharmacophore of non-peptide bradykinin receptor antagonists. *Biochem Pharmacol.* 2004;67:601–609.
12. Lupala CL, Gomez-Gutierrez P, Perez JJ. New insights into the stereochemical requirements of the bradykinin B2 receptor antagonist binding. *J Comput-Aided Mol Des.* 2016;30:85–101.
13. Lupala CL, Gomez-Gutierrez P, Perez JJ. New insights into the stereochemical requirements of the bradykinin B1 receptor antagonist binding. *J Mol Graphics Model.* 2016;68:184–196.
14. The computational software utilized and detail of the computational protocols can be found in Supplementary data.
15. Surgand JS, Rodrigo J, Kellenberger E, Rognan D. A chemogenomic analysis of the transmembrane binding cavity of human G-protein-coupled receptors. *Proteins.* 2006;62:509–538.
16. Ballesteros JA, Weinstein H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* 1995;25:366–428.
17. Fathy DB, Mathis SA, Leeb T, Leeb-Lundberg LM. A single position in the third transmembrane domains of the human B1 and B2 bradykinin receptors is adjacent to and discriminates between the C-terminal residues of subtype-selective ligands. *J Biol Chem.* 1998;273:12210–12218.
18. Joedicke L, Mao J, Kuenze G, et al. The molecular basis of subtype selectivity of human kinin G-protein-coupled receptors. *Nat Chem Biol.* 2018;14:284–290.

Article

New Insights into the Stereochemical Requirements of the Bombesin BB1 Receptor Antagonists Binding

Bahareh Rasaeifar, Patricia Gomez-Gutierrez and Juan J. Perez * [iD]

Department of Chemical Engineering, Universitat Politecnica de Catalunya, ETSEIB, Av. Diagonal, 647, 08028 Barcelona, Spain; bahareh.rasaeifar@upc.edu (B.R.); ondopasa@gmail.com (P.G.-G.)
* Correspondence: juan.jesus.perez@upc.edu

check for updates

Abstract: Members of the family of bombesinlike peptides exert a wide range of biological activities both at the central nervous system and in peripheral tissues through at least three G-Protein Coupled Receptors: BB1, BB2 and BB3. Despite the number of peptide ligands already described, only a few small molecule binders have been disclosed so far, hampering a deeper understanding of their pharmacology. In order to have a deeper understanding of the stereochemical features characterizing binding to the BB1 receptor, we performed the molecular modeling study consisting of the construction of a 3D model of the receptor by homology modeling followed by a docking study of the peptoids PD168368 and PD176252 onto it. Analysis of the complexes permitted us to propose prospective bound conformations of the compounds, consistent with the experimental information available. Subsequently, we defined a pharmacophore describing minimal stereochemical requirements for binding to the BB1 receptor that was used in silico screening. This exercise yielded a set of small molecules that were purchased and tested, showing affinity to the BB1 but not to the BB2 receptor. These molecules exhibit scaffolds of diverse chemical families that can be used as a starting point for the development of novel BB1 antagonists.

Keywords: bombesin receptors; neuromedin B antagonism; G-protein coupled receptors homology modeling; non-peptide neuromedin B antagonists

1. Introduction

Members of the bombesinlike family of peptides, originally isolated from the skin of diverse amphibians and later found to be widely distributed in mammals [1,2], are compounds with a wide spectrum of biological activity. Thus, in the central nervous system they are involved in satiety, control of circadian rhythm and thermoregulation, whereas in peripheral tissues, stimulation of gastrointestinal hormone release, macrophages activation and effects on development [3]. In addition, they are known to play a role in the control of cellular proliferation [4,5]. Actions of this family of peptides are mediated through at least three G-Protein Coupled Receptors: the neuromendin B receptor (known as BB1R), the gastrin-releasing peptide receptor (known as BB2R), and the orphan, since its endogenous ligand has not been disclosed yet, bombesin receptor subtype 3 (known as BB3R) [6–8]. Due to the wide spectrum of biological activities mediated by these receptors, there is considerable interest in understanding their potential use as therapeutic agents. A literature review reveals potential therapeutic use for both agonists and antagonists targeting any of the three receptors for cancer therapy [9,10]: BB3R agonists for the treatment of obesity/diabetes mellitus [11]; BB2R antagonists for the treatment of radiation-induced lung injury [12] and BB1R or BB2R antagonists for the treatment of itching in atopic dermatitis [13].

Neuromedin B (NMB) and the gastrin-releasing peptide (GRP), together with its shorter version GRP(18–27)—known as neuromedin C (NMC)—are the mammal endogenous ligands of the BB1R and

BB2R, respectively [1,2]. They are selective agonists for the respective receptors, binding with high affinity: NMB exhibits a Ki = 0.052 nM for the BB1R and about 1000 times higher for the BB2R, whereas GRP exhibits a Ki = 0.19 nM for the BB2R and about 1000 times higher for the BB1R [2]. Neither NMB nor GRP bind to the BB3R, and although its endogenous ligand has not been disclosed, several selective peptide and nonpeptide agonists have been disclosed [2]. On the other hand, a few peptide antagonists with a diverse degree of selectivity have also been disclosed for the three receptors [2,14,15]. However, the poor oral bioavailability, low absorption, rapid degradation by proteolytic enzymes and immunogenic profile of peptides make nonpeptide molecules more desirable [16,17]. Efforts in this direction resulted in the discovery of second generation peptoids PD168368 and PD176252 (Figure 1) [18], along with a set of analogs with diverse substitutions [19] that exhibit an antagonist profile for the BB1R and BB2R, with diverse degrees of selectivity. Specifically, PD168368 exhibits a Ki = 0.5 nM for the BB1R and 1700 nM for the BB2R, whereas PD176252 exhibits a Ki = 0.5 nM for the BB1R and 170 nM for the BB2R [2]. Furthermore, the same scaffold was also later used to design ML-18, a BB3R selective antagonist [20], and more recently, compounds AM-37 and ST-36 with diverse pharmacological profile to the diverse bombesin receptors [21]. However, the similarity between the chemical structures of these compounds is so big that it makes it intriguing to understand the subtle differences that provide selectivity for the different receptors. In the absence of ligand–receptor 3D complex structures, rationalization of the structure–activity needs to be carried out indirectly, although the lack of structural diversity of the known binders makes it a difficult job. Complementary to the structure–activity studies, it was pointed a few years ago in an interesting report [22] that Tyr220 in BB1R (corresponding to Phe218 in BB2R) can explain the differential behavior exhibited by PD168368 for the two receptors.
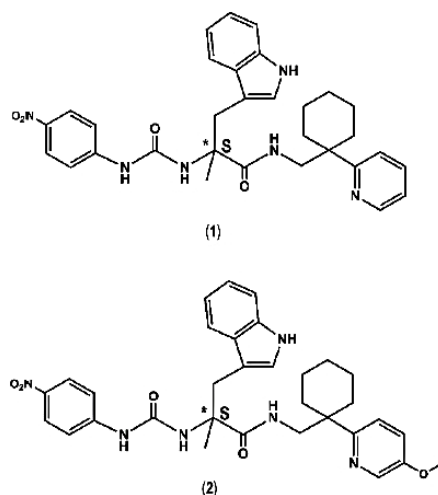

Figure 1. Chemical structures including the chirality of the asymmetric carbon (*) of the bombesin antagonists studied in the present work. PD168368 (**1**) and PD176252 (**2**).

Despite knowledge accumulated in the past to ascertain the potential use of bombesin antagonists as therapeutic agents, advancement has been hindered because of the low diversity within the ligands available. Moreover, PD168368 and PD176252 have been reported to be potent agonists of the human formyl-peptide receptors, questioning the interpretation that their reported effects can be solely attributed to their activity as BB1/BB2 antagonists [23]. Only a few compounds devoid of a peptoid scaffold have been disclosed in the literature to date: a BB1R antagonist with a dibenzodiazepine scaffold [24], the reported antagonist of the BB2 receptor NSC-77427 [25] and Bantag-1, a peptidomimetic designed by an isostere replacement that exhibits a selective antagonist profile for the BB3R [26].

In order to have a better understanding of the therapeutic potential of bombesin antagonists, there is a need to discover novel compounds based on different chemical structures. In this direction, we report, in the present work, the results of a molecular modeling study aimed at discovering novel chemical scaffolds suitable to develop novel bombesin antagonists. For this purpose, we constructed a 3D model of the BB1 receptor by homology modeling and docked PD176252 in diverse poses. Analysis of the ligand–receptor complexes together with known structure–activity studies of this family of compounds permitted to define a pharmacophore that was subsequently used for in silico screening. The results of the study permitted the identification of a limited set of small molecules that were purchased and tested at 50 µM for its capacity to antagonize NMB at the BB1R. The results of this study are disclosed in the present report.

2. Results and Discussion

Figure 2 shows the result of the multiple alignment of the diverse sequences of the class-A rhodopsin family used in the present work. As can be seen, the alignment was robust enough to show properly aligned diverse conserved motifs, specifically, a set of conserved residues including (in the Ballesteros-Weinstein numbering scheme [27]) N1.50 in TM1; L2.46, A2.47, D2.50 in TM2; D/E3.49, R3.50, Y3.51 in TM3; W4.50 in TM4; F5.47, P5.50, Y5.58 in TM5; F6.44, W6.48, P6.50 in TM6; N7.49, P7.50, Y7.53 in TM7. Focusing on the sequences of the NTS1 and BB1 receptors, the conserved residue Y5.58 in the former was an N5.58, but this difference did not induce any structural consequence. In contrast, Y7.53 in the former corresponded to Y7.54 in the latter. A closer inspection of the alignment suggests that this corresponded to a displacement and not an insertion that could induce a buldge in TM7. Other motifs such as D(E)RY in TM3, CWxP(Y/F/L) in TM6 or the NPxxY in TM7 were well aligned.
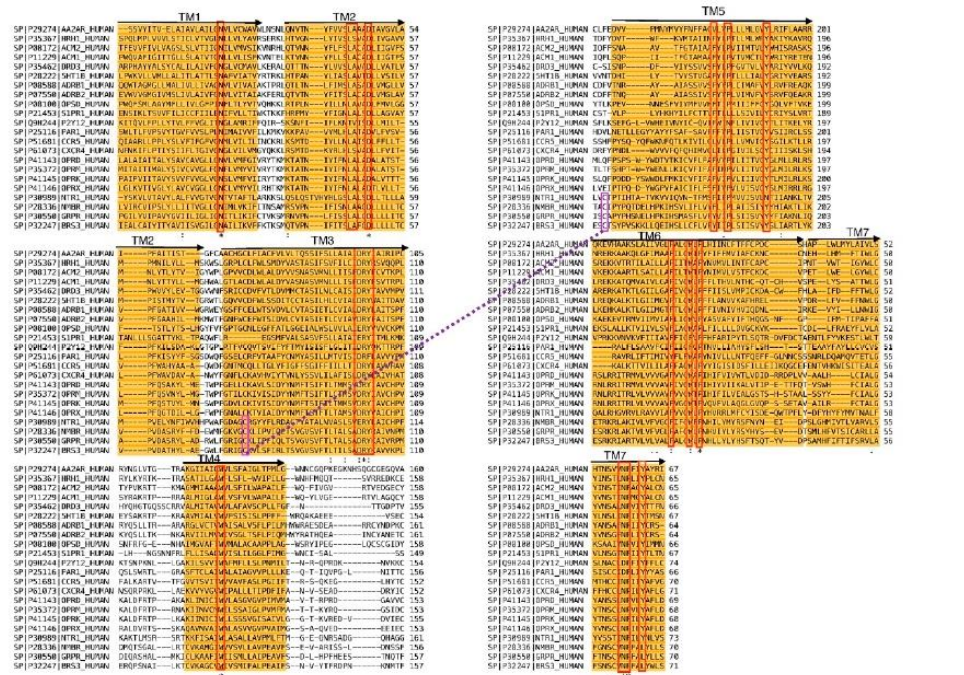

Figure 2. Multiple sequence alignment of diverse GPCRs used in the present work (see text). Transmembrane segments are colored in orange and conserved residues are in red boxes. There is also a purple line indicating a disulfide bridge.

Figure 3 shows the evolution of the root-mean-square-deviation (rmsd) computed using the Cα of the BB1 receptor structure along the MD trajectory. A can be seen, it required more than 100 ns for the structure to have the structure equilibrated as had been observed previously with other MD simulations of GPCRs [28]. The procedure permitted the construction of a final model of the BB1 receptor that was generated as the average structure obtained using the last 100 ns of the molecular dynamics trajectory. This structure was subsequently minimized in a two-step procedure, using the steepest descent method with a distance dependent dielectric constant of **2**. First, side chains were optimized with the backbone atoms constrained to be subsequently released in a second minimization step.
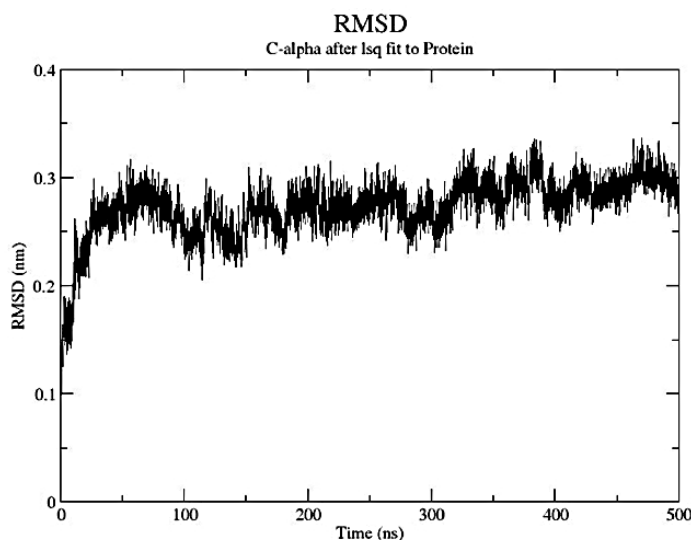
Figure 3. Time evolution of the root mean square deviation (rmsd) of the bombesin BB1 receptor during the refinement process. The rmsd was computed using the alpha carbons of the protein.

In order to understand the molecular features that ligands need to bind onto the BB1 receptor, we analyzed the complex of PD168368 and PD176252, respectively, in their prospective bound conformations. For this purpose, we exploited the available structure–activity studies [18,19], the available results from mutagenesis studies [22] as well as the small residue differences between the BB1R and the BB2R at the orthosteric site. Accordingly, the constructed model of the BB1 receptor was used for docking the antagonists in diverse conformations and orientations. For this purpose, we followed the same protocol as described above. Multiple orientations were obtained that were rank ordered using the XP scoring function of GLIDE [29]. The diverse complexes were subsequently energy minimized in vacuo with a distance dependent dielectric constant of **2**, using the steepest descent method to allow the ligand to adapt to the environment.

Taking the asymmetric carbon of the ligands as reference, the structure of both compounds can be described as composed of three branches including a nitrophenylurea moiety, an indole moiety and a 2-pryridinecyclohexane moiety. The diverse orientations and conformations found permitted the three legs to occupy diverse sites of the orthosteric site of the BB1 receptor and analyze diverse ligand–receptor interactions. As can be expected, the orientations adopted by PD168368 and PD176252 were comparable due to their size and the small structural differences between them. Actually, the difference between the two molecules consisted of an extra methoxyl group at the pryridine moiety in PD176252, as shown in Figure 1. However, this extra moiety provided PD176252 with an order of magnitude improved affinity to the BB2R that can be attributable to an additional favorable interaction with the receptor.

Therefore, in order to assess the most likely binding mode of the ligands to the receptor, we analyzed the diverse ligand–receptor interactions, aimed at the identification of key residues, and checked their conservation in the BB2R to understand if a specific interaction could be responsible for the observed differential binding affinity of the compounds for the two receptors. This assumes that differences in the free energy of binding between PD168368 and PD176252 are attributable only to the enthalpic component. Obviously, differences in the affinity of the two compounds could also be attributed to the ability of the differential chemical group to kick out a bound water molecule from the receptor, improving the affinity of a ligand by an order of magnitude [30]. As mentioned before, early site-directed mutagenesis studies demonstrated the involvement of Tyr220 of BB1 in the binding of PD168368. Actually, when Tyr220 is mutated to a phenylalanine, binding drops to that observed for the BB2R [22]. This result suggests that PD168368, and presumably PD176252, exhibit a critical interaction with Tyr220 in the BB1R bound complex that disappears when the residue is mutated to Phe220. Since the replacement of Tyr for Phe represents a conservative mutation, it can be concluded that the hydroxyl group of the tyrosine is involved in a hydrogen bond with the ligand. This represents an important structural constraint that permits us to discard a few poses from the docking study. Analysis of the diverse poses consistently showed a subset of them with the hydroxyl group of the Tyr220 side chain pointing towards the center of the aromatic ring of the nitrobenzene ring of the ligand, as shown in Figure 4. Interestingly, the interaction between both moieties was reinforced by a quadruple–quadruple interaction between the aromatic ring of the ligand and the aromatic ring of the tyrosine side chain adopting a T-shape relative orientation. This explains in part the lower affinity exhibited by PD168368 and PD176252 for the BB2R. In this case, although the quadrupole–quadrupole interaction between aromatic rings was preserved in the complex with the BB2R, the loss of a hydrogen bond can be associated with an order of magnitude drop of affinity [31]. In addition, the nitro group formed a hydrogen bond with His283 (residue conserved in the BB2R), reinforcing the role played by the nitrophenyl moiety in the affinity of the two ligands, as shown in Figure 4.
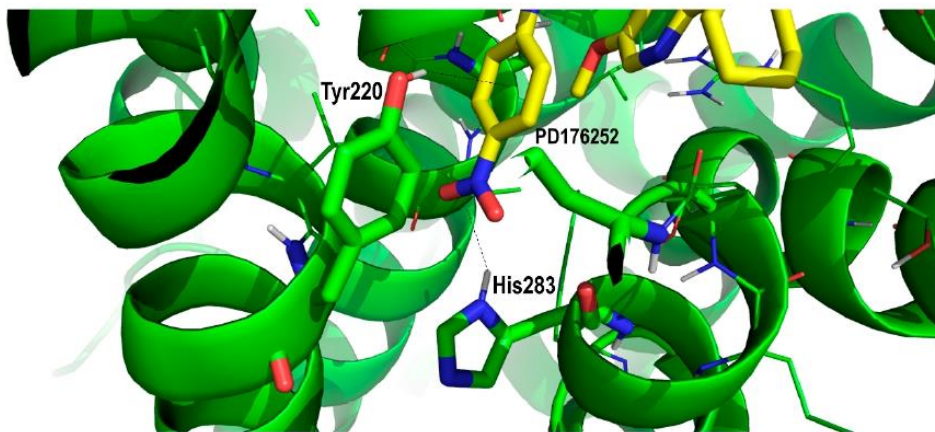

Figure 4. Close-up of PD176252 in its prospective bound conformation to the bombesin BB1 receptor, showing the interaction between its nitrophenyl moiety and diverse residues BB1R including Tyr220 and His286 (see text).

The indole moiety in both ligands sat in a hydrophobic pocket defined by residues Phe181, Pro120 and Leu215 (Figure 5). In addition, the indole moiety was oriented in such a way that it formed a hydrogen bond with Glu178. All these residues were conserved in the BB2R, so these interactions help to explain part of the affinity of the ligand, but they do not explain the higher affinity observed by PD168368 and PD176252 towards the BB1R.
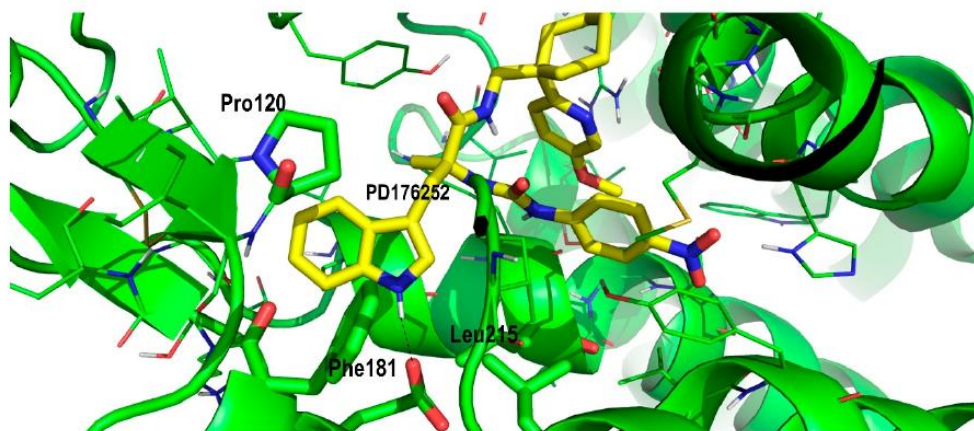
Figure 5. Close-up of PD176252 in its prospective bound conformation to the bombesin BB1 receptor, showing the interaction between its indole moiety and residues Phe181, Pro120 and Leu215 (see text).

To understand the difference in the binding affinity between PD168368 and PD176252 towards the BB2R we should focus on the role of the methoxyl group attached to the 2-pyridinecyclohexane moiety. As mentioned above, due to the similar size of the ligands, it can be hypothesized that PD168368 and PD176252 adopt a similar bound conformation to the BB1R. Analysis of the diverse poses with the nitrophenyl moiety accommodated in the proximity to Tyr220, revealed a subset of poses, where the methoxypyridinyl group nicely sat in a pocketin which the side chain of Ser126 (residue conserved in the BB2R) is located at a suitable distance to form a hydrogen bond with the oxygen of the methoxyl moiety (Figure 6). In addition, the side chain of Arg310 (also conserved in BB2R) that formed a hydrogen bond with Asp100 in the apo form, was in the position to form a hydrogen bond with the center of the aromatic ring of the PD176252 pyridine ring. The presence of a hydrogen bond between the ligand and Ser126 can explain the affinity difference between PD168368 and PD176252 towards the BB2R. However, it does not explain that the affinity of both compounds to the BB1R remained the same. This can be explained on the basis that the 2-pyridine moiety did not appear to bind in the same conformation in both ligands. In the case of PD168368, the pyridine ring was likely to be accommodated close to Arg289, in such a way that its side chain formed a hydrogen bond with the nitrogen of the heterocycle (Figure 7). Since this residue was conserved in both receptors, this effect was expected to be found in both the BB1R and the BB2R and does not contribute to explaining the affinity difference of the compound for the two receptors. However, the position of the heterocycle permitted an additional interaction with Phe105 via a quadrupole–quadrupole interaction that was not found in the BB2R, since this residue was not conserved (it corresponded to Leu102 in BB2R). In addition, the cyclohexane group accessed a hydrophobic pocket formed by residues Met287 (Leu285 in BB2R), the conserved Ile296 and the aliphatic chain of the conserved Lys210. These differences, together with the differential interaction with Tyr220, can explain the affinity difference of the compounds for the two receptors. Accordingly, the fact that PD176252 exhibited the same affinity for BB1R and BB2R could be explained as a compensation of interactions, in such a way that the gain of a hydrogen bond with Ser126 may compensate the loss of the quadrupole–quadrupole interaction with Phe105.
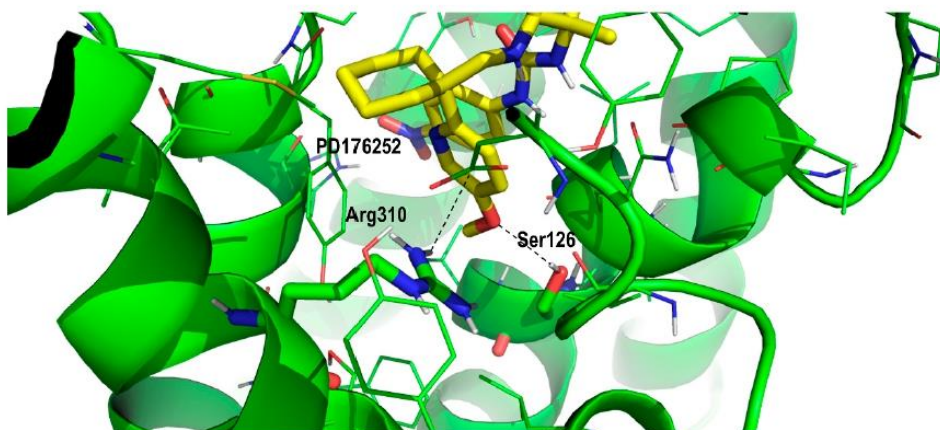
Figure 6. Close-up of PD176252 in its prospective bound conformation to the bombesin BB1 receptor, showing the interaction between its phenoxyl-2-pyridine moiety interacting with diverse residues including Ser126, Gln123 and Arg310 (see text).
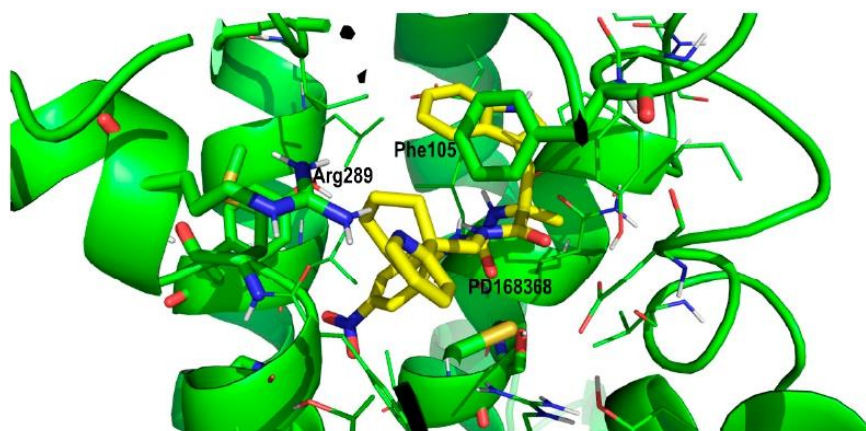

Figure 7. Close-up of PD168368 in its prospective bound conformation to the bombesin BB1 receptor, showing the interaction between its phenoxyl-2-pyridine moiety and diverse residues including Arg289 and Phe105 (see text).

According to the previous analysis, we can propose prospective conformations of PD168368 and PD176252 bound to the BB1R. Figure 8 shows pictorially the two compounds superimposed when bound to the orthosteric site. Inspection of Figure 8 shows the nitrophenyl group sitting close to Tyr220 and His283 and the indole moiety sitting in a hydrophobic pocket defined by Phe181, Pro120 and Leu215 for both ligands, as described above. These residues were all conserved in the BB2R, so it is not expected that they can provide an explanation for the affinity differences between BB1R and BB2R. On the other hand, the 2-pyridincyclohexane moiety apparently bound in a different conformation in the two ligands. In the case of PD168368, the pyridine ring adopted an extended conformation, whereas in the case of the PD176252, the corresponding dihedral angle was twisted to 60°. This differential conformation permitted the pryridine ring of PD176252 to access a hydrophobic pocket and form a hydrogen bond with the side chain of Ser126, as explained above. In contrast, PD168368 adopted a conformation that permitted the formation of a hydrogen bond with Arg289 together with a quadrupole–quadrupole interaction with Phe105.
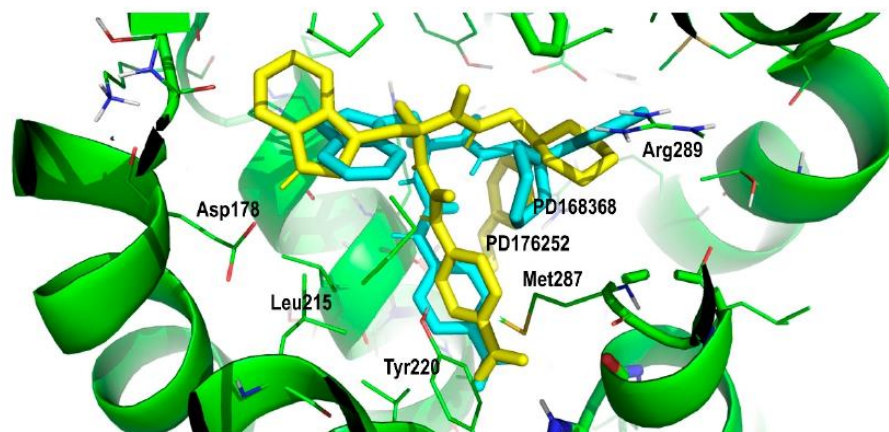
Figure 8. Superposition of PD168368 (cyan) and PD176252 (yellow) in their prospective bound conformations, respectively, to the BB1R.

In order to further assess the feasibility of both prospective bound conformations, we checked the extent that these models explained the available structure–activity results of this family of compounds [18]. Analysis of the activity of the diverse compounds showed that when the nitro group of the nitrophenyl moiety was attached in position 3 or when it was substituted by a proton accepting chemical group such as nitrile, the compound preserved its activity. However, when the nitro group was placed in position 2 or substituted by a group with lower proton accepting capability, the affinity dropped at least one order of magnitude. The same trends were observed regarding the binding affinity for the BB2R. On the other hand, substitutions on the 2-pydirydine moiety that preserved a proton accepting center in position 4 had similar behavior.

Proof of Concept

Using the prospective bound conformation of PD176252 onto the BB1R, we proceeded to identify those interactions that appeared to be key for ligand–receptor recognition, aimed at defining a pharmacophore that could be used as a query for an in silico virtual screening. Accordingly, a simple three-point pharmacophore was defined as shown in Figure 9. The pharmacophore was defined as simply as possible to identify hits with chemical scaffolds of diverse profiles. The pharmacophore involved Tyr220 in the form of a proton accepting center, aimed at discovering BB1R selective compounds, since the BB2R does not have this capability. Specifically, the 3-point pharmacophore included: (i) a proton accepting center in the direction of the OH bond of the hydroxyl moiety of Tyr220 side chain, located at 2.5 Å of the hydroxyl hydrogen; (ii) a hydrophobic center located at a point defined by the side chains of Pro200, Phe181 and Leu215; (iii) a proton donor center located in the plane defined by the atoms of the carboxyl group of Asp100 at 3.0 Å from the center of the two oxygens.

The three-point pharmacophore was used as a query for an in silico screening of diverse databases using the Molecular Operating Environment (MOE) program [32]. More precisely, the query was defined in the form of three spheres with diverse radii to introduce some tolerance in each of the pharmacophoric points. Specifically, the proton acceptor and donor spheres were defined with a 0.12 nm radius, whereas the hydrophobic was defined within a 0.18 nm radius. Two databases of 3D structures of small molecules were used for the screening process including the leadlike database included in the MOE software containing around 650,000 commercially available compounds [32] and the leads-now subset of the ZINC database containing approximately 4,200,000 unique molecules downloaded in 2015 [33]. For each compound, in addition to its 3D structure, the database includes a set of conformations generated using a build-up procedure from systematic conformational searches of molecular fragments.
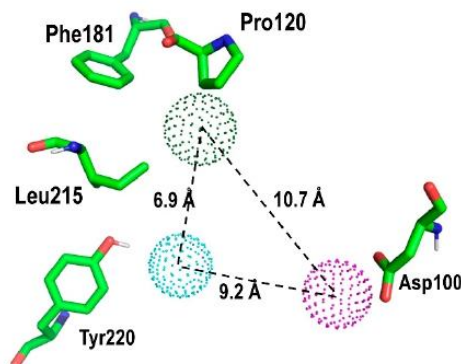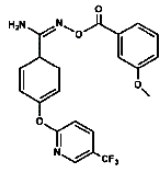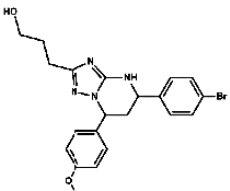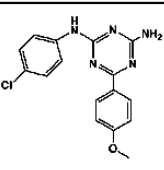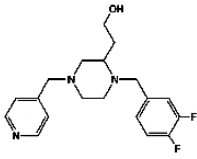
Figure 9. 3-point pharmacophore defined by the geometries of a few residues characterizing the sterochemical features involved in BB1R binding (see text).

After screening the 3D databases, a few hundred hits were identified and subjected to a diversity analysis [34]. For this purpose, molecules were encoded as bit strings using the typed atom triangle (TAT) methodology in which atoms are grouped in trios based on information about their chemical nature and mutual distance [35]. In a next step, the distance between bit strings was computed using the Tanimoto coefficient [36]. Finally, molecules were grouped in clusters using the Jarvis-Patrick algorithm [37]. This procedure permitted us to select a subset of compounds preserving the diversity of the initial set, showing the diversity of chemical scaffolds. Representative molecules of the diverse clusters were selected according to the suitability of chemical groups responsible for each of the pharmacophoric points and checking that molecules may not suffer steric clashes. A set of about fifty selected compounds were docked onto the BB1R using the GLIDE software [29] and ranked order according to the XP scoring function. Thirteen compounds among those with the best score were purchased (Appendix A) and tested at 50 μM for their capacity to displace the radioligand to the BB1R, as explained in the methods section [38]. The chemical structures of the six compounds were found to displace the radioligand used in the binding assays by more than 15%, as listed in Table 1. These results yielded a success rate of about 50%, as found in other cases [39,40]. Table 1 also lists the results of radioligand displacement experiments for the BB1R and BB2R, suggesting that the ligands are BB1R selective.

Table 1. Listing of small molecules identified in the in silico screening (see text). Column 2 shows their chemical structure and columns 3 and 4 the displacements of the corresponding BB1R and BB2R radioligands, respectively, (in percentage) at 50 μM ($n = 2$).

Compound#	Chemical Structure	BB1R(Neuromendin B Receptor) Radioligand Displacement (%)	BB2R(Gastrin-Releasing Peptide Receptor) Radioligand Displacement (%)
1		19.7	0.0
2		24.3	0.0

Table 1. *Cont.*

Compound#	Chemical Structure	BB1R(Neuromendin B Receptor) Radioligand Displacement (%)	BB2R(Gastrin-Releasing Peptide Receptor) Radioligand Displacement (%)
3		28.1	0.0
4		30.5	0.0
5		38.0	10.5
6		16.1	0.0

　　　In order to explain these results, we proceeded to dock the novel hits disclosed onto the BB1R following the same protocol as explained in the methods section. As an example, Figure 10 shows pictorially compound #5 in its prospective bound conformation to the BB1R. As can be seen, the triazine ring works as a scaffold with three branches: a phenoxyl moiety, a chlorophenyl moiety and an amine. Analysis of the prospective bound conformation showed that the oxygen of the methoxyl group fulfilled pharmacophoric point #1, exhibiting a hydrogen bond interaction with the hydroxyl moiety of Tyr220. Moreover, the chlorine atom sat at the hydrophobic site (pharmacophoric point #2) and the amine fulfilled pharmacophoric point #3 exhibiting a hydrogen bond with Asp300. In the case of the BB2R, these ligands could not attain the hydrogen bond with Tyr220, fulfilling only two pharmacophoric points, explaining that their affinity for the BB2R is much lower than for the BB1R.
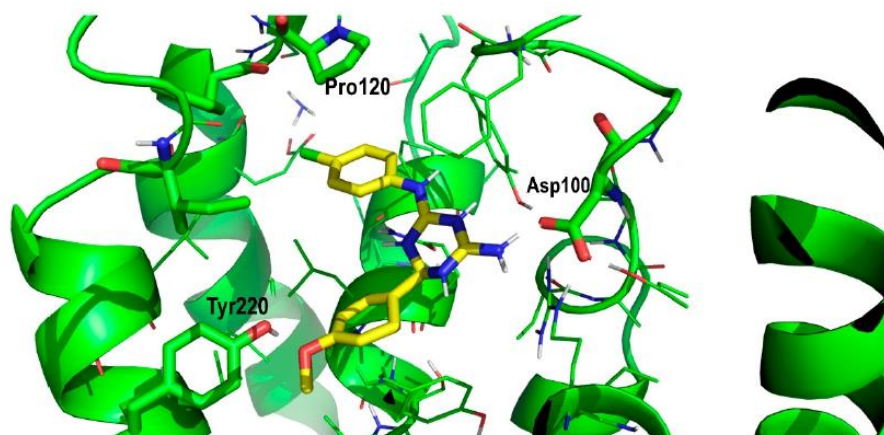
Figure 10. Pictorial view of the proposed binding mode of Compound **#5** to the bombesin BB1 receptor.

Finally, although the affinities exhibited by the hits presently disclosed is not very high, they represent a set of small molecules with high diversity that can be exploited further to discover novel selective antagonists for the BB1R.

3. Materials and Methods

3.1. Molecular Modeling

A crude model of the human bombesin BB1 receptor was constructed by homology modeling using the rat neurotensin receptor NTS1 (PDB entry: 4GRV) as a template [41]. The template was selected for being one of the few receptors with a known crystallography structure located in the same branch of bombesin in the GPCRs phylogenetic tree [28,42]. Since the 4GRV structure corresponds to a fusion protein of NTS1 and the T4 Lysozyme, the template structure was edited by removing the coordinates of the latter and joining the segments of the ECL3 left at both sides. In a subsequent step, the sequences of the BB1 and the NST1 receptors were aligned. This process is critical for ensuring the accuracy of the models constructed by homology modeling [43]. Since in this case, sequence identity is low (~23%; 43% sequence similarity), we undertook a multiple sequence alignment to improve its quality [44]. Accordingly, we selected a set of 20 sequences of diverse GPCRs of the class-A with a known crystallographic structure plus the three bombesin receptors that were aligned using the CLUSTALW software (version 2.0.12) [45]. To finalize the alignment process, the two sequences of interest were checked to ensure the alignment of diverse known conserved residues in the class A family of GPCRs and motifs such as D(E)RY in TM3, CWxP(Y/F/L) in TM6 or the NPxxY in TM7, as well as the location of the conserved disulfide bridge between ECL2 and TM3. In a further step, the alignment was used to thread the sequence of the BB1 receptor onto the backbone of the template structure, using the Molecular Operating Environment (MOE) program (version 2019.01) [32]. Thirty models were produced after the incorporation of alternative sidechain conformations, using an extensive rotamer library embedded in MOE, generated from a high-resolution structural database. Once hydrogens were added to the models using the protonate3D method [46], each of them was energy minimized using a contact energy function to relieve any serious steric strains. Models were checked for inter-residue contacts as well as for backbone conformations through the Ramachandran map and scored. Finally, the model with the lower root-mean square deviation (rmsd) in regard to the average structure with the highest score was selected as a crude model and considered for refinement.

Refinement of the receptor was carried out using molecular dynamics with a ligand bound in the orthosteric site. The presence of a ligand provides a more efficient refinement of the constructed model [28]. Accordingly, the antagonist PD176252 (Figure 1) [17] was docked in different conformations

into the orthosteric site of the crude model using the GLIDE software (version 6.7) [29]. Multiple orientations were obtained that were rank ordered using the XP scoring function. The pose with the highest score was energy minimized in vacuo with a distance dependent dielectric constant of **2** using the steepest descent method to permit a reorientation of the ligand. Subsequently, the ligand–receptor complex was embedded in a lipid bilayer. Specifically, the protein was embedded in a box consisting of a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipid and water molecules generated and equilibrated according to the procedure described previously [47]. The box had an initial size of $8.9 \times 8.3 \times 10.5$ nm^3 (XYZ), organized in such a way that the bilayer plane was oriented on the XY plane. The protein was placed in the center of the box, and the overlapping molecules were removed. Specifically, all water molecules with oxygen atoms closer than 0.40 nm to a nonhydrogen atom of the protein, as well as all lipid molecules with at least one atom closer than 0.25 nm to a nonhydrogen atom of the protein, were removed. This resulted in a final system containing 193 lipids and ca. 12,000 water molecules. Removal of these atoms introduced small voids between the protein and water or lipid molecules that disappeared during the first part of the molecular dynamics (MD) simulation, in which a progressive adjustment of the lipid bilayer and water molecules to the protein takes place. Next, 101 randomly selected water molecules were replaced by 47 sodium and 54 chloride ions, providing a neutral system with a concentration approximately 0.2 M in sodium chloride. This concentration is similar to that found in biological organisms, although they exhibit different intra- and extracellular ion concentrations. Then, the system was energy minimized to avoid steric classes using the steepest descent method and subjected to a 500 ns MD simulation at a constant pressure using the GROMACS package 4.6 [48]. Molecules were described using the all-atom OPLS force field [49] currently implemented in GROMACS, except for water molecules that were modeled using the TIP3P model [50]. The system was subjected to periodic boundary conditions in the three coordinate directions. The temperature was kept constant at 300 K using separate thermostats for the protein, water, ions and lipid molecules. The time constant for the thermostats was set to 0.1 ps, except for water, for which a smaller value of 0.01 ps was used. The pressure in the three coordinate directions was kept at 0.1 MPa by independent Berendsen barostats using a time constant of 1.0 ps. The equations of motion were integrated using the leapfrog algorithm with an integration step of **2** fs. All bonds involving hydrogen atoms within the protein and lipid molecules were kept frozen using the LINCS algorithm [51]. The bonds and the angle of water molecules were fixed using the analytical SETTLE method. Lennard–Jones interactions were computed using a cutoff of 1.0 nm. Electrostatic interactions were treated using the particle-mesh Ewald procedure [52].

3.2. Binding Assays

BB1 antagonism assays were carried out following a protocol described elsewhere [42]. Specifically, human recombinant bombesin BB1 receptors expressed in CHO-K1 cells were used in modified HEPES-KOH buffer pH 7.4 (Thermo Scientific, Waltham, MA, USA). A 0.2 µg protein aliquot was incubated with 0.1 nM [^{125}I][Tyr^4]-bombesin for 60 min at 25 °C. Nonspecific binding was estimated in the presence of 1 µM neuromedin B. Membranes were filtered and washed; the filters were then counted to determine [^{125}I][Tyr^4]-bombesin (K_d = 0.13 nM) specifically bound. Hits were screened at 50 µM. Compound binding was calculated as the percentage of the inhibition of the binding of a radioactively labeled ligand.

4. Conclusions

This paper reports the results of a modeling study aimed at shedding some light on the stereochemical requirements for small molecule binders to the BB1 bombesin receptor. For this purpose, we first constructed a 3D model of the bombesin BB1 receptor by homology modeling using the rat neurotensin receptor as a template. Then, the model was refined using molecular dynamics in a system composed by the receptor embedded in a bilayer of POCP lipids, water and sodium chloride. The MD simulation was carried out with the ligand PD176252 bound to the receptor in

its orthosteric site to accelerate the refinement process. After 500 ns sampling, the refined structure of the receptor was computed as the average of the diverse configurations sampled during the last 100 ns of the trajectory. Subsequently, the modeled 3D structure of the receptor was used to dock the antagonists PD168368 and PD176252 in its orthosteric site. Analysis of the complexes, guided by structure–activity and mutagenesis studies available, permitted us to propose prospective complexes of the bound conformation of each of the ligands to the BB1 receptor. The results of this study directly connect diverse pieces of information that were available in the literature and can be used as the basis for designing new experiments and small molecule ligands.

As a proof of principle, we also carried out an in silico screening using a simple pharmacophore defined from the complex of PD176252 bound to the BB1R. Specifically, a three-point pharmacophore that involves a point exclusive for the BB1 receptor was used for this purpose. The in silico study permitted us to identify a set of small molecules with affinity for the BB1 receptor that were also disclosed. Interestingly, none of the molecules exhibited affinity for the BB2 receptor. The set of molecules have scaffolds of a diverse chemical nature that can be used as a starting point for the development of novel BB1 antagonists.

Appendix A

Catalog Numbers and Suppliers of the Compounds Listed in Table 1.

Compound #1: AKOS000796000 (AKos GmbH, Stuttgart, Germany); Compound #2: Amb19684292 (Ambinter c/o Greenpharma SAS, Orleans, France); Compound #3: CB358 (Menai Organics Ltd., Bangor, UK); Compound #4: Amb1220897 (Ambinter c/o Greenpharma SAS, Orleans, France); Compound #5: Amb3992353 (Ambinter c/o Greenpharma SAS, Orleans, France); Compound #6: Amb11089933 (Ambinter c/o Greenpharma SAS, Orleans, France).

References

1. Jensen, R.T.; Moody, T.W. Bombesin Peptides (Cancer). In *Hand-Book of Biologically Active Peptides*; Kastin, A.J., Ed.; Elsevier: Amsterdam, The Netherlands, 2013; pp. 506–511.
2. Ramos-Alvarez, I.; Moreno, P.; Mantey, S.A.; Nakamura, T.; Nuche-Berenguer, B.; Moody, T.W.; Coy, D.H.; Jensen, R.T. Insights into bombesin receptors and ligands: Highlighting recent advances. *Peptides* **2015**, *72*, 128–144. [CrossRef]
3. Weber, H.C. Regulation and signaling of human bombesin receptors and their biological effects. *Curr. Opin. Endocrinol. Diabetes Obes.* **2009**, *16*, 66–71. [CrossRef] [PubMed]
4. Flores, D.G.; De Farias, C.B.; Leites, J.; De Oliveira, M.S.; Lima, R.C.; Tamajusuku, A.S.; Leone, L.P.D.; Meurer, L.; Brunetto, A.L.; Schwartsmann, G.; et al. Gastrin releasing peptide receptors regulate proliferation of C6 glioma cells through a phosphatidylinositol 3-kinase-dependent mechanism. *Curr. Neurovasc. Res.* **2008**, *5*, 99–105. [CrossRef] [PubMed]
5. Moody, T.W.; Moreno, P.; Jensen, R.T. Neuropeptides as lung cancer growth factors. *Peptides* **2015**, *72*, 106–111. [CrossRef] [PubMed]
6. Benya, R.V.; Kusui, T.; Pradhan, T.K.; Battey, J.F.; Jensen, R.T. Expression and characterization of cloned human bombesin receptors. *Mol. Pharmacol.* **1995**, *47*, 10–20. [PubMed]
7. Fathi, Z.; Corjay, M.H.; Shapira, H.; Wada, E.; Benya, R.; Jensen, R.; Viallet, J.; Sausville, E.A.; Battey, J.F. BRS-3: A novel bombesin receptor subtype selectively expressed in testis and lungcarcinoma cells. *J. Biol. Chem.* **1993**, *268*, 5979–8594. [PubMed]

8. Jensen, R.T.; Battey, J.F.; Spindel, E.R.; Beny, R.V. International Union of Pharmacology. LXVIII. Mammalian Bombesin Receptors: Nomenclature, Distribution, Pharmacology, Signaling, and Functions in Normal and Disease States. *Pharm. Rev.* **2008**, *60*, 1–42. [CrossRef]

9. Moreno, P.; Ramos-Alvarez, I.; Moody, T.W.; Jensen, R.T. Bombesin related peptides/receptors and their promising therapeutic roles in cancer imaging; targeting and treatment. *Expert Opin. Ther. Targets* **2016**, *20*, 1055–1073. [CrossRef]

10. Park, H.J.; Kim, S.R.; Kim, M.K.; Choi, K.S.; Jang, H.O.; Yun, I.; Bae, M.-K. Neuromedin B receptor antagonist suppresses tumor angiogenesis and tumor growth in vitro and in vivo. *Cancer Lett.* **2011**, *312*, 117–127. [CrossRef]

11. Gonzalez, N.; Moreno, P.; Jensen, R.T. Bombesin receptor subtype 3 as a potential target for obesity and diabetes. *Expert Opin. Ther. Targets* **2015**, *19*, 1153–1170. [CrossRef]

12. Zhou, S.; Nissao, E.; Jackson, I.L.; Leong, W.; Dancy, L.; Cuttitta, F.; Vujaskovic, Z.; Sunday, M.E. Radiation-Induced Lung Injury Is Mitigated by Blockade of Gastrin-Releasing Peptide. *Am. J. Pathol.* **2013**, *182*, 1248–1254. [CrossRef] [PubMed]

13. Ehling, S.; Fukuyama, T.; Ko, M.C.; Olivry, T.; Bäumer, W. Neuromedin B Induces Acute Itch in Mice via the Activation of Peripheral Sensory Neurons. *Acta Derm. Venereol.* **2019**, *99*, 587–593. [CrossRef] [PubMed]

14. Cristau, M.; Devin, C.; Oiry, C.; Chaloin, O.; Amblard, M.; Bernad, N.; Heitz, A.; Fehrentz, J.A.; Martinez, J. Synthesis and Biological Evaluation of Bombesin Constrained Analogues. *J. Med. Chem.* **2000**, *43*, 2356–2361. [CrossRef] [PubMed]

15. Gonzalez, N.; Mantey, S.A.; Pradhan, T.K.; Sancho, V.; Moody, T.W.; Coy, D.H.; Jensen, R.T. Characterization of putative GRP- and NMB-receptor antagonist's interaction with human receptors. *Peptides* **2009**, *30*, 1473–1486. [CrossRef]

16. Perez, J.J.; Corcho, F.; Llorens, O. Molecular modeling in the design of peptidomimetics and peptide surrogates. *Curr. Med. Chem.* **2002**, *9*, 2209–2229. [CrossRef]

17. Perez, J.J. Designing Peptidomimetics. *Curr. Top. Med. Chem.* **2018**, *18*, 566–590. [CrossRef]

18. Ashwood, V.; Brownhill, V.; Higginbottom, M.; Horwell, D.C.; Hughes, J.; Lewthwaite, R.A.; McKnight, A.T.; Pinnock, R.D.; Pritchard, M.C.; Suman-Chauhan, N.; et al. PD 176252-The First High Affinity Non-peptide Gastrin-Releasing Peptide (BB2) Receptor Antagonist. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2589–2594. [CrossRef]

19. Carrieri, A.; Lacivita, E.; Belviso, B.D.; Caliandro, R.; Mastrorilli, P.; Gallo, V.; Niso, M.; Leopoldo, M. Structural determinants in the binding of BB2 receptor ligands: In silico, x-ray and NMR studies in PD176252 analogues. *Curr. Top. Med. Chem.* **2017**, *17*, 1599–1610. [CrossRef]

20. Moody, T.W.; Mantey, S.A.; Moreno, P.; Nakamura, T.; Lacivita, E.; Leopoldo, M.; Jensen, R.T. ML-18 is a non-peptide bombesin receptor subtype-3 antagonist which inhibits lung cancer growth. *Peptides* **2015**, *64*, 55–61. [CrossRef]

21. Moody, T.W.; Tashakkori, N.; Mantey, S.A.; Moreno, P.; Ramos-Alvarez, I.; Leopoldo, M.; Jensen, R.T. AM-37 and ST-36 are Small Molecule Bombesin Receptor Antagonists. *Front. Endocrinol.* **2017**, *8*, 176. [CrossRef]

22. Tokita, K.; Hocart, S.J.; Katsuno, T.; Mantey, S.A.; Coy, D.H.; Jensen, R.T. Tyrosine 220 in the 5th Transmembrane Domain of the Neuromedin B Receptor Is Critical for the High Selectivity of the Peptoid Antagonist PD168368. *J. Biol. Chem.* **2001**, *276*, 495–504. [CrossRef] [PubMed]

23. Schepetkin, I.A.; Kirpotina, L.N.; Khlebnikov, A.I.; Jutila, M.A.; Quinn, M.T. Gastrin-Releasing Peptide/Neuromedin B Receptor Antagonists PD176252, PD168368, and Related Analogs Are Potent Agonists of Human Formyl-Peptide Receptors. *Mol. Pharmacol.* **2011**, *79*, 77–90. [CrossRef] [PubMed]

24. Fu, J.; Shuttleworth, S.J.; Connors, R.V.; Chai, A.; Coward, P. Discovery and optimization of a novel Neuromedin B receptor antagonist. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4264–4267. [CrossRef]

25. Martinez, A.; Zudaire, E.; Julian, M.; Moody, T.W.; Cuttitta, F. Gastrin-releasing peptide (GRP) induces angiogenesis and the specific GRP blocker 77427 inhibits tumor growth in vitro and in vivo. *Oncogene* **2005**, *24*, 4106–4113. [CrossRef] [PubMed]

26. Moreno, P.; Mantey, S.A.; Nuche-Berenguer, B.; Reitman, M.L.; Gonzalez, N.; Coy, D.H.; Jensen, R.T. Comparative pharmacology of bombesin receptor subtype-3 nonpeptide agonist MK-5046, a universal peptide agonist; and peptide antagonist Bantag-1 for human bombesin receptors. *J. Pharmacol. Exp. Ther.* **2013**, *347*, 110–116. [CrossRef] [PubMed]

27. Ballesteros, J.A.; Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* **1995**, *25*, 366–428.

28. Lupala, C.S.; Rasaifar, B.; Gomez-Gutierrez, P.; Perez, J.J. Using Molecular Dynamics for the refinement of atomistic models of GPCRs by homology modeling. *J. Biomol. Struct. Dyn.* **2018**, *36*, 2436–2448. [CrossRef]

29. Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.; Shaw, D.E.; Shelley, M.; et al. Glide: A New Approach for Rapid Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749. [CrossRef]

30. Venkatakrishnana, A.J.; Maa, A.K.; Fonseca, R.; Latorraca, N.R.; Kelly, B.; Betz, R.M.; Asawa, C.; Kobilka, B.K.; Dror, R.O. Diverse GPCRs exhibit conserved water networks for stabilization and activation. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 3288–3293. [CrossRef]

31. Bissantz, C.; Kuhn, B.; Stahl, M. A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem.* **2010**, *53*, 5061–5084.

32. *Molecular Operating Environment (MOE), Version 2019.01*; Chemical Computing Group UCL: Montreal, QC, Canada, 2020.

33. Sterling, T.; Irwin, J.J. ZINC 15—Ligand Discovery for Everyone. *J. Chem. Inf. Model.* **2015**, *55*, 2324–2337. [CrossRef] [PubMed]

34. Perez, J.J. Managing molecular diversity. *Chem. Soc. Rev.* **2005**, *34*, 143–152. [CrossRef] [PubMed]

35. Bender, A.; Jenkins, J.L.; Scheiber, J.; Chetan, S.; Sukuru, K.; Glick, M.; Davies, J.W. How Similar Are Similarity Searching Methods? A Principal Component Analysis of Molecular Descriptor Space. *J. Chem. Inf. Model.* **2009**, *49*, 108–119. [CrossRef] [PubMed]

36. Lipkus, A.H. A proof of the triangular inequality for the Tanimoto distance. *J. Math. Chem.* **1999**, *26*, 263–265. [CrossRef]

37. Jarvis, R.A.; Patrick, E.A. Clustering Using a Similarity Measure Based on Shared Near Neighbors. *IEEE Trans. Comp.* **1973**, *22*, 1025–1034. [CrossRef]

38. Ohki-Hamazaki, H.; Iwabuchi, M.; Maekawa, F. Development and function of bombesin-like peptides and their receptors. *Int. J. Dev. Biol.* **2005**, *49*, 293–300. [CrossRef]

39. Krishna, S.; Singh, D.K.; Meena, S.; Datta, D.; Siddiqi, M.I.; Banerjee, D. Pharmacophore-Based Screening and Identification of Novel Human Ligase I Inhibitors with Potential Anticancer Activity. *J. Chem. Inf. Model.* **2014**, *54*, 781–792. [CrossRef]

40. Gomez-Gutierrez, P.; Campos, P.M.; Perez, J.J. Identification of a Novel Inhibitory Allosteric Site of MAP Kinases. *PLoS ONE* **2016**, *11*, e0167379. [CrossRef]

41. Noinaj, N.; White, J.F.; Shibata, Y.; Love, J.; Kloss, B.; Xu, F.; Gvozdenovic-Jeremic, J.; Shah, P.; Shiloach, J.; Tate, C.G.; et al. The crystal structure of the neurotensin receptor NTS1 in complex with neurotensin (8–13). *Nature* **2012**, *490*, 508–513.

42. Costanzi, S.; Skorski, M.; Deplano, A.; Habermehl, B.; Mendoza, M.; Wang, K.; Biederman, M.; Dawson, J.; Gao, J. Homology modeling of a Class A GPCR in the inactive conformation: A quantitative analysis of the correlation between model/template sequence identity and model accuracy. *J. Mol. Graph. Model.* **2016**, *70*, 140–152. [CrossRef]

43. Nayeem, A.; Sitkoff, D.; Krystek, S., Jr. A comparative study of available software for high-accuracy homology modeling: From sequence alignments to structural models. *Protein Sci.* **2006**, *15*, 808–824. [CrossRef] [PubMed]

44. Cavasotto, C.N.; Palomba, D. Expanding the horizons of G protein-coupled receptor structure-based ligand discovery and optimization using homology models. *Chem. Commun.* **2015**, *51*, 13576–13594. [CrossRef] [PubMed]

45. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948. [CrossRef] [PubMed]

46. Labute, P. Protonate3D: Assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins* **2009**, *75*, 187–205. [CrossRef] [PubMed]

47. Cordomi, A.; Edholm, O.; Perez, J.J. Effect of different treatments of long-range interactions and sampling conditions in molecular dynamic simulations of rhodopsin embedded in a dipalmitoyl phosphatidylcholine bilayer. *J. Comput. Chem.* **2007**, *28*, 1017–1030. [CrossRef]

48. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J. GROMACS: Fast flexible and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718. [CrossRef]

49. Kaminski, G.; Friesner, R.A.; Tirado-Rives, J.; Jorgensen, W.L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474–6487. [CrossRef]

50. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926. [CrossRef]

51. Miyamoto, S.; Kollman, P.A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992**, *13*, 952–962. [CrossRef]

52. Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089. [CrossRef]

1 *Communication*

2 Molecular features of non-selective small molecule
3 antagonists of the Bradykinin Receptors

4 **Bahareh Rasaeifar, Patricia Gomez-Gutierrez and Juan J. Perez***

5 Department of Chemical Engineering. Universitat Politecnica de Catalunya. ETSEIB. Av. Diagonal, 647; 08028
6 Barcelona, Spain; bahareh.rasaeifar@upc.edu (B.R.); ondopasa@gmail.com (P.G.-G.)
7 * Correspondence:juan.jesus.perez@upc.edu;

9 **Abstract:** Angiotensin converting enzyme 2 (ACE2) downregulation is a key negative factor for the
10 severity of lung edema and acute lung failure observed in patients infected with SARS-CoV-2. ACE2
11 downregulation affects the levels of diverse peptide mediators of the renin-agiotensin-
12 aldestosterone and kallikrein-kinin systems, compromising vascular hemostasis. Increasing
13 evidence suggests that the inflammatory response observed in covid-19 patients is initiated by the
14 action of kinins on the bradykinin receptors. Accordingly, the use of bradykinin antagonists should
15 be considered as a strategy for therapeutic intervention against covid-19 illness progression.
16 Presently, icatibant is the only bradykinin antagonist drug approved. In the present report, we
17 investigated the molecular features characterizing non-selective antagonists targeting the
18 bradykinin receptors and carried out a *in silico* screening of approved drugs, aimed at the
19 identification of compounds with a non-selective bradykinin antagonist profile that can be
20 evaluated for drug repurposing. The study permitted to identify eight compounds as prospective
21 non-selective antagonists of the bradykinin receptors, including raloxifene; sildenafil; cefepime;
22 cefpirome; imatinib; ponatinib; abemaciclib and entrectinib.

23 **Keywords:** bradykinin receptors; bradykinin pharmacophore; non-peptide bradykinin antagonists;
24 bradykinin repurposing; Covid-19 therapy

25

26 1. Introduction

27   Beginning in December 2019, a novel coronavirus designated SARS-CoV-2 was identified as the
28 pathogen causing an international outbreak of respiratory illness termed Covid-19, originated in
29 Wuhan, Hubei Province, China. Despite the virus has a fatality rate of only ~2-3% it exhibits a high
30 transmission rates, resulting in a high overall death toll that forced the World Health Organization
31 to declare SARS-CoV-2 as a pandemic infectious disease of international concern on March 11, 2020
32 [1]. Until July 18, 2020, there have been 14,108.240 confirmed cases of Covid-19 with 602.695
33 confirmed deaths [2]. Unfortunately, ~20% of the people infected are likely to develop pneumonia of
34 varying severity that may evolve to acute respiratory distress syndrome (ARDS), sepsis, and death
35 [3]. Presently, clinical treatment of Covid-19 is mainly symptomatic by using anti-inflammatories like
36 dexamethasone [4] or cytokine inhibitors, combined with antibiotics to treat secondary infections.
37 Knowledge of the mechanism behind SARS-CoV-2 infection will help to identify other therapeutic
38 agents to be used for the treatment of patients with Covid-19. This report focuses in the mechanism
39 of infection and the implication of the renin-angiotensin-aldosterone and the kallikrein/kinin systems
40 in illness progression [5].
41   The renin-angiotensin-aldosterone system (RAAS) [6] and the kallikrein/kinin system (KKS) [7]
42 are involved in the regulation of intravascular volume, blood pressure and tissue repair via
43 inflammatory and proliferative mechanisms. The angiotensin converting enzyme (ACE) and the

44 angiotensin converting enzyme 2 (ACE2) are key players in both systems. Cross-talk between the two
45 systems is summarized in Figure 1. ACE is a carboxydipeptidase that produces the octapeptide
46 angiotensin II from its inactive precursor angiotensin I, orchestrating a plethora of actions including
47 sodium reabsorption and increase of blood pressure mediated through the AT1 receptor and
48 vasodilation and natriuresis mediated through the AT2 receptor [8]. On the other hand, ACE2 is an
49 integral membrane carboxypeptidase with its catalytic domain at the extracellular side that
50 counterbalances the actions of ACE. Specifically, ACE2 degrades angiotensin II to produce
51 angiotensin (1-7), a peptide that elicits vasodilation and natriuresis via activation of the Mas receptor
52 [9]. Furthermore, ACE2 also converts angiotensin I into angiotensin (1-9), a peptide that elicits
53 vasodilation and anti-inflammatory effects through activation of the AT2 receptor [10]. Angiotensin
54 (1-9) is further converted into angiotensin (1-7) by the action of ACE [8]. On the other hand, kallikreins
55 are serine proteases that produce bradykinin (BK) and kallidin (Lys-BK) -two members of the kinin
56 family- from kininogens in response to inflammation, trauma, burns, shock, allergy and some
57 cardiovascular diseases [11]. Other members of the kinin family include the corresponding des-Arg
58 analogs: des-Arg^9-BK and des-Arg^{10}-Lys-BK.
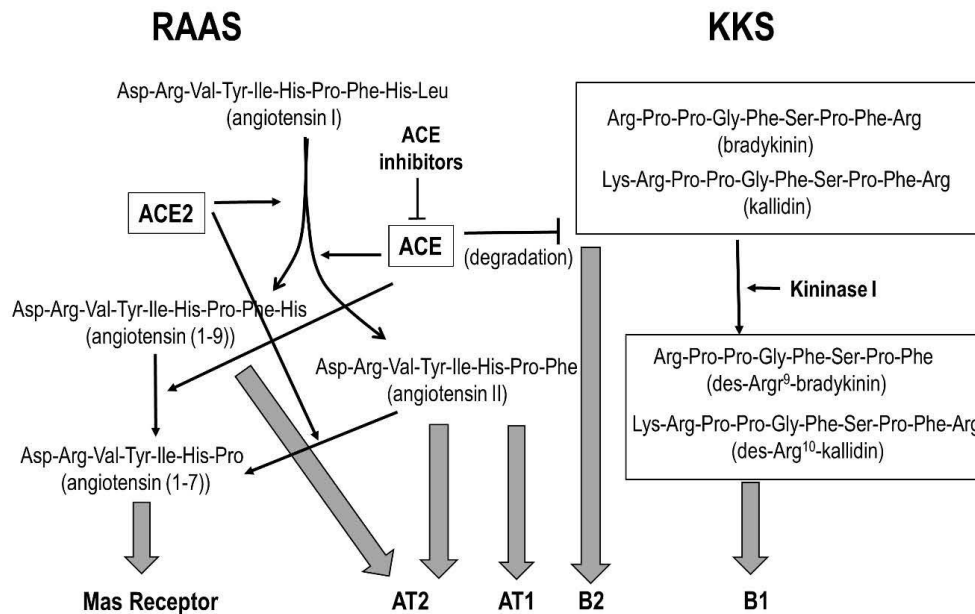
Figure 1. Cross-talk between RAAS and KKS. The angiotensin converting enzymes ACE and ACE2 are key
players of RAAS, regulating the production of diverse mediators (see text), producing a plethora of physiological
actions through the activation of different receptors (solid arrows). Thus, activation of the angiotensin AT1
receptor produces vasoconstriction, hypertrophy and fibrosis; whereas activation of the AT2 and Mas receptors
produce vasodilation, antihypertrophy and antifibrosis. On the other hand, ACE regulates the levels of kinins
that produce vasodilatation and increased vascular permeability through the B1 and B2 receptors.

67 Kinins produce a plethora of physiological actions including vasodilatation and increased
68 vascular permeability via activation of the B1 and B2 receptors [12, 13]. The former is upregulated
69 during inflammation episodes or tissue trauma, whereas the latter is constitutively expressed in a
70 variety of cell types. BK and Lys-BK are agonists of the B2 receptor, whereas the des-Arg analogs are
71 agonists of the B1 receptor [14]. ACE and the ACE2 enzymes are actors involved in the cross-talk

72　between RAAS and KKS. The former upregulates angiotensin II and downregulates BK, whereas the
73　latter upregulates angiotensin (1-9) and downregulates Arg^9-detached kinins, respectively [15].
74　　　　SARS-CoV-2 binds with high affinity to ACE2 facilitating cell fusion and entry [16, 17].
75　Endocytosis of the SARS-CoV-2 spike protein-ACE2 complex into endosomes reduces surface ACE2
76　expression, being detrimental for its role in tissue protection; producing a clear interference on RAAS
77　mediated homeostasis functions. Taking into account that ACE2 is more abundantly present in the
78　epithelia of the lungs and on lymphocytes [18], its downregulation is a key negative factor for severity
79　of lung edema and acute lung failure observed in patients infected by SARS-CoV. Actually,
80　downregulation of ACE2 translates into altered levels of diverse mediators of the SAARS and KKS.
81　Specifically, levels of angiotensin II are increased and in turn, those of the des-Arg kinins due to a
82　lower availability of ACE to degrade BK and Lys-BK, whereas levels of angiotensin (1-9) and
83　angiotensin (1-7) are decreased. About fifteen years ago, it was hypothesized that the observed
84　physiological effects produced in patients infected by SARS-CoV were due to the actions of
85　angiotensin II on the AT1 and AT2 receptors [19, 20]. Presently, there is growing evidence that
86　inflammation may be triggered through the des-Arg peptides/B1 axis-mediated signaling pathway
87　[21- 24]. This new perspective suggests that inhibition of BK signaling may be a suitable therapy to
88　avoid the cytokine storm associated with the Covid-19 illness [25].
89　　　　Based on this novel mechanistic hypothesis, selective and non-selective BK antagonists should
90　be considered as therapeutic agents for the treatment of covid-19 [26, 27]. Despite the enormous effort
91　devoted in the past to design peptide and non-peptide selective ligands targeting the BK receptors
92　[13, 28], icatibant (Figure 2) is the only BK antagonist presently approved as therapeutic agent for the
93　symptomatic treatment of acute attacks of hereditary angioedema in adults with C1-esterase-
94　inhibitor deficiency [29]. Despite being a B2 selective antagonist, the compound is presently involved
95　in a clinical trial to assess its benefits for the treatment of the covid-19 illness [30]. Considering the
96　lack of BK antagonists in the market and the urgency to have new treatments for the covid-19
97　available, drug repurposing is a valuable strategy for quickly discover novel therapeutic uses of
98　already approved drugs. Specifically in this case, the discovery of approved therapeutic agents with
99　a BK antagonist profile.
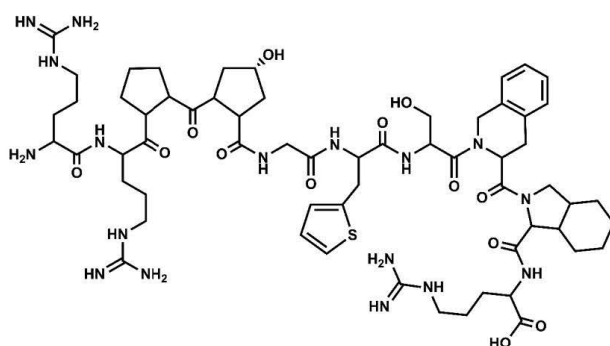
D-Arg-Arg-Pro-Hyp-Gly-Thi-Ser-D-Tic-Oic-Arg
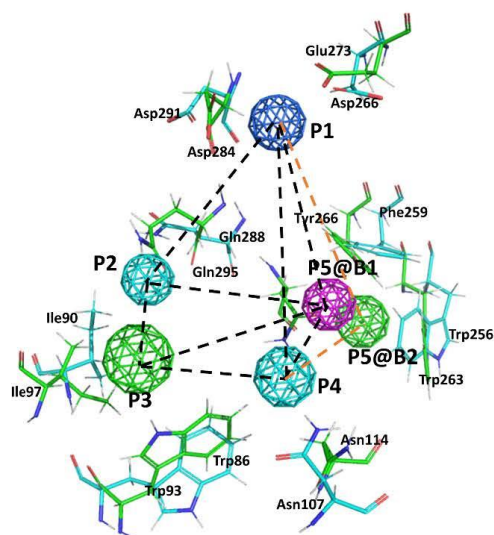
100
101　**Figure 2.** Chemical structure of icatibant, together with its residue sequence. Hyp=hydroxylproline; Thi=
102　thiophenyl-alanine; Tic= 1,2,3,4-tetrahydroisoquinolin-2-ylcarbonyl; Oic= (3aS,7aS)-octahydroindol-2-
103　ylcarbonyl.
104

105   Virtual screening methods can be very valuable in drug repurposing, provided we count on
106 specific structural knowledge of the therapeutic target of interest [31]. Specifically, for BK we recently
107 reported the results of a modeling study addressed to analyze the stereochemical features required
108 for non-peptide selective ligands to bind to the B1 and B2 receptors [32, 33]. Furthermore, the results
109 of the study also permitted to identify the stereochemical features associated with selective binding
110 to each of them [34]. As a complement to that work, we discuss in the present contribution the
111 characterization of the molecular features that confer a non-selective binding profile to small
112 molecule ligands targeting the B1 and B2 receptors and show preliminary results from a virtual
113 screening aimed at the identification of approved drugs with a non-selective profile to the BK
114 receptors.

115 **2. Results and Discussion**

116 *2.1. Stereochemical features of non-selective small molecule ligands targeting the B1 and B2 receptors*

117   Due to the lack of crystallographic structures of the BK receptors, the construction of 3D models
118 at atomic resolution by homology modeling of the B1 and B2 receptors was recently performed and
119 reported [32, 33]. Models were subsequently used to undertake a docking study that permitted the
120 analysis of diverse ligand-receptor complexes of known selective small molecule compounds. From
121 this study, the corresponding pharmacophores describing the stereochemical features that ligands
122 must fulfill for binding to each of the two receptors were defined. The two receptors share a high
123 sequence identity of 28% (sequence homology is 43%) that reaches ~50% when the orthosteric sites
124 are compared. Accordingly, it is expected to find common residues in their respective orthosteric
125 sites. Comparison of the pharmacophores shows that they exhibit four points in common (Figure 3):
126 a positive charge (P1); a hydrogen bond donor/acceptor (P2); an aromatic ring (P3); and a hydrogen
127 bond donor/acceptor (P4). Point P5 in the two pharmacophores discriminates binding to the two
128 receptors: in B1 is hydrogen bond acceptor, whereas in B2 is hydrophobic/aromatic moiety.
129 Accordingly, ligands fulfilling pharmacophore points P1-P4, common to the two receptors are
130 expected to be non-selective bradykinin antagonists. As a proof of concept, we disclosed a short list
131 of non-selective hits identified through the virtual screening of a large database of small molecules
132 [34]. In contrast, the design of selective ligands is trickier. In addition of the differential chemical
133 nature of P5 in the two receptors, analysis of the 3D models of the ligand-receptor complexes suggests
134 that there is a steric hindrance that prevents selective B1 ligands to bind B2 and, vice versa. The steric
135 hindrance is produced by the differential nature of the side chains of the non-conserved residues
136 Arg^{202} in B1 compared to its counterpart, Thr^{197} in B2 [34].

137

138 **Figure 3.** Superposition of the B1 and B2 receptor pharmacophores. Colored spheres represent
139 pharmacophore points (P1-P5) according to the following color code: dark blue represents a positive
140 charge moiety; magenta a hydrogen bond accepting center; light blue a hydrogen bond
141 donor/acceptor center; green an aromatic/lipophilic center. @B1 and @B2 is used to differentiate P5 for
142 the B1 and B2 receptors, respectively. Consensus distances between common pharmacophore points
143 are (black dotted lines): d(1,2)=9 Å; d(1,3)=14 Å; d(1,4)=10.5 Å; d(2,3)=6 Å; d(2,4)=7 Å; d(3,4)=7.5 Å.
144 Specific distances for the B1 pharmacophore: d(1,5)=9.5 Å; d(2,5)=9.3 Å; d(3,5)=9.5 Å; d(4,5)=5.7 Å;
145 whereas for the B2 pharmacophore are (orange dotted lines): d(1,5)=11 Å; d(2,5)=9 Å; d(3,5)=8.8 Å;
146 d(4,5)=8.4 Å. Side chains of the main residues involved in defining the binding pocket for non-peptide
147 ligands are explicitly depicted: green for the B1 and blue for the B2, respectively.
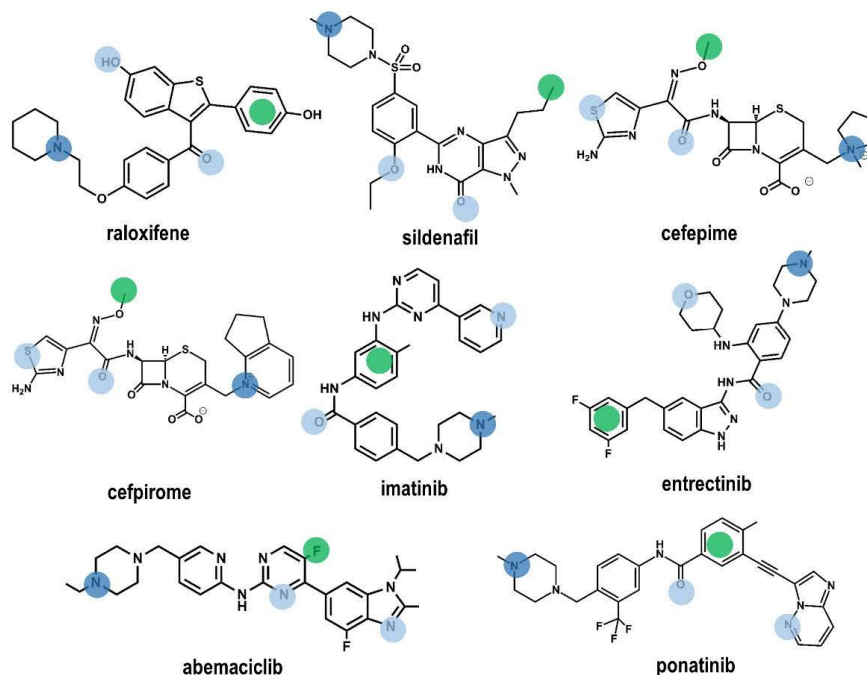
148 *2.2. Drug repurposing*

149 Using the four common pharmacophore points P1-P4 as a query, we carried a *in silico* screening
150 on the DrugBank using the Molecular Operating Environment (MOE) program [35]. The DrugBank
151 is a database containing comprehensive information of all FDA approved drugs [36]. In order to carry
152 out the virtual screening study, we first generated the corresponding 3D DrugBank database using
153 the Database Viewer module of MOE [35]. The database contains for each molecule its 3D structure
154 together with a set of conformations, generated using a build-up procedure from systematic
155 conformational searches of molecular fragments. Virtual screening was carried out on a subset of
156 1703 molecules selected according to their molecular weight between 200 and 600. Hits obtained were
157 subsequently docked onto the 3D models of the B1 and B2 receptors, respectively to check for possible
158 steric hindrance. Preliminary results of the virtual screening yielded eight drugs (Figure 4): raloxifene
159 [37], a selective estrogen receptor modulator; sildenafil [38], a phosphodiesterase type 5 inhibitor;
160 cefepime [39] and cefpirome [40], two β-lactam antibiotics; imatinib [41] and ponatinib [42], two bcr-
161 Abl tyrosine kinase inhibitors; abemaciclib [43], a dual inhibitor of cyclin-dependent kinases 4 and 6;
162 and entrectinib [44] a non-selective tyrosine kinase inhibitor. According to the results of this study,
163 these compounds exhibit the characteristics to be non-selective BK ligands. Figure 4 shows the
164 location of the common pharmacophore points on their 2D chemical structures. Evaluation of the
165 ability of these compounds to act as B1 and B2 receptor antagonists is currently underway.

166
167 **Figure 4.** Chemical structures of the eight compounds identified by virtual screening of the DrugBank database.
168 Color dots are drawn on top of the moieties responsible for fulfilling pharmacophore points P1-P4 according to
169 the color code described in Figure 1.
170

171 Among the drugs identified, there are no studies reporting a direct BK antagonistic profile of
172 any of them. However, in the case of raloxifene there are studies that show a synergistic action with
173 bradykinin. Actually, rats treated with raloxifene show an increased reduction of systolic blood
174 pressure on administration of bradykinin, suggesting an enhanced bioavailability of NO in these
175 animals [45]. Also, sildenafil does not exhibit a synergetic action in the reduction of BK induced
176 glucose uptake in humans when administered together with N^{ω}-monomethyl-L-arginine a nitric-
177 oxide synthase inhibitor [46], indirectly suggesting that sildenafil may not interact with the BK
178 receptors.

179 **3. Materials and Methods**

180 The 3D models of the B1 and B2 were constructed by homology modeling using the chemokine
181 CXCR4 receptor as template (pdb entry code 3ODU) [47], following the procedure explained
182 elsewhere [32, 33]. Initial crude models of the receptors were constructed by threading the sequences
183 of the B1 and B2 receptors onto the backbone of the template following the sequence alignment and
184 subsequently validated using the Modeller 9 version 8 (9v8) software [48]. Next, models were refined
185 using molecular dynamics simulations using a system consisting of each of the respective receptors
186 embedded in a lipid bilayer of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipids and
187 water molecules, using GROMACS 4.6 package [49] as described elsewere [50]. B1 and B2 small
188 molecule pharmacophores were defined after docking studies of diverse non-peptide selective
189 ligands to each of the two receptors [32, 33]. Docking studies were carried out using a set of unique
190 conformations resulted from thorough conformational searches for the diverse ligands studied and
191 rank ordered using the XP score function of GLIDE [51].

192 **4. Conclusions**

Around 20% of the people infected with SARS-CoV-2 are likely to develop pneumonia of varying severity that may evolve to acute respiratory distress syndrome (ARDS), sepsis, and death. SARS-CoV-2 binds with high affinity to ACE2, mediating cell fusion and entry. ACE2 downregulation was pointed as the origin of the observed inflammatory response of sever cases of Covid-19, mediated by angiotensin II through the AT1 receptor. However, there is an increasing evidence that inflammation is mediated through the bradykinin B1 receptor due to the increased levels of des-Arg⁹BK. Accordingly, antagonists of the bradykinin receptors could be useful therapeutic agents to block the inflammatory signaling process. Presently, icatibant is the only bradykinin antagonist approved drug in the market and there are clinical studies in progress to assess its efficacy for the treatment of Covid-19. However, icatibant is a B2 selective antagonist and it is desirable to have also B1 selective or non-selective antagonists available.

In order to have novel therapeutic treatments available in a short time, drug repurposing is a valuable procedure. Repurposing of already approved drugs has several advantages like their known safety/tolerability profiles, availability and low cost. In order to speed up the identification of approved drugs for novel therapeutical uses, virtual screening can be a valuable tool, provided that structural information on the target is available.

As a continuation of a previous study devoted to identify the molecular features required by compounds to exhibit an antagonist profile to the bradykinin receptors, we discussed in the present manuscript those molecular features that provide a non-selective profile to them. These features were used as a query to carry out a virtual screening on the DrugBank, a database containing all approved drugs. The study yielded eight molecules that were subsequently docked onto the 3D models of the B1 and B2 receptors respectively, to check for possible steric hindrance. Evaluation of their profile as bradykinin antagonists is currently under investigation.

References

1. World Health Organization. General's Opening Remarks at the Media Briefing on COVID-19-18 March 2020; World Health Organization: Geneva, Switzerland, **2020**.

2. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). Available on-line at: https://coronavirus.jhu.edu/map.html. (accessed on July, 18th **2020**).

3. Wu, Z.; McGoogan, J.M. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. *J. Am. Med. Assoc.* **2020**, *323*, 1239-1242.

4. Horby, P.; Lim, W. S.; Emberson, J. R.; Mafham, M.; Bell, J. L.; Linsell, L.; Staplin, N.; Brightling, C.; Ustianowski, A.; Elmahi, E.; Prudon, B.; Green, C.; Felton, T.; Chadwick, D.; Rege, K.; Fegan, C.; Chappell, L. C.; Faust, S. N.; Jaki, T.; Jeffery, K.; Montgomery, A.; Rowan, K.; Juszczak, E.; Baillie, J. K.; Haynes, R.; Landray. M. J. Dexamethasone in Hospitalized Patients with Covid-19 - Preliminary Report. *N. Engl. J. Med.* **2020**; 10.1056/NEJMoa2021436. doi:10.1056/NEJMoa2021436. Online ahead of print.

5. van de Veerdonk, F.; Netea, M.G.; van Deuren, M.; van der Meer, J.W.; de Mast, Q.; Bruggemann, R.J.; van der Hoeven, H. Kinins and Cytokines in COVID-19: A Comprehensive Pathophysiological Approach. *Preprints* **2020**, 2020040023 (doi: 10.20944/preprints202004.0023.v1).

6. Crowley, S. D.; Coffman, T. M. Recent advances involving the renin–angiotensin system. *Exp. Cell Res.* **2012**, *318*, 1049-1056.

7. Schmaier, A. H. The plasma kallikrein-kinin system counterbalances the renin-angiotensin system. *J. Clin. Invest.* **2002**, *109*, 1007–1009.

241   8.   Bernstein, K. E.; Ong, F. S.; Blackwell, W.-L.; Kandarp, B.; Shah, H.; Giani, J. F.; Gonzalez-Villalobos, R. A.;
242   Shen, X. Z.; Fuchs, S. A Modern Understanding of the Traditional and Nontraditional Biological Functions
243   of Angiotensin-Converting Enzyme. *Pharmacol Rev.* **2013**, *65*, 1-46.

244   9.   Santos, R. A.; Simoes e Silva, A.C.; Maric, C.; Silva, D.M.; Machado, R.P.; de Buhr, I.; Heringer-Walther, S.;
245   Pinheiro, S.V.; Lopes, M.T.; Bader, M.; Mendes, E.P.; Lemos, V.S.; Campagnole-Santos, M.J.; Schultheiss,
246   H.P.; Speth, R.; Walther, T. Angiotensin-(1-7) is an endogenous ligand for the G protein-coupled receptor
247   Mas. *Proc. Natl. Acad. Sci. USA.* **2003**, *100*, 8258–8263.

248   10. Flores-Munoz, M.; Work, L.M.; Douglas, K.; Denby, L.; Dominiczak, A.F.; Graham, D.; Nicklin, S.
249   Angiotensin-(1–9) attenuates cardiac fibrosis in the stroke-prone spontaneously hypertensive rat via the
250   angiotensin type 2 receptor. *Hypertension.* **2012**, *59*, 300–307.

251   11. Golias, C.; Charalabopoulos, A.; Stagikas, D.; Charalabopoulos, K.; Batistatou, A. The kinin system -
252   bradykinin: biological effects and clinical implications. Multiple role of the kinin system-bradykinin.
253   *Hippokratia.* **2007**, *11*, 124-128.

254   12. Regoli, D.; Barabe, J. Pharmacology of bradykinin and related kinins. *Pharmacol. Rev.* **1980**, *32*, 1-46.

255   13. Marceau, F.; Bachelard, H.; Bouthillier, J.; Fortin, Morissette, J.-P. G.; Bawolak, M.-T.; Charest-Morin, X.;
256   Gera, L. Bradykinin receptors: Agonists, antagonists, expression, signaling, and adaptation to sustained
257   stimulation. Int. Immunopharmacol. **2020**, *82*, 106305.

258   14. Leeb-Lundberg, L. M. F.; Marceau, F.; Muller-Ester, W.; Pettibone, D. J.; Zuraw, B. L. International Union
259   of Pharmacology. XLV. Classification of the Kinin Receptor Family: from Molecular Mechanisms to
260   Pathophysiological Consequences. *Pharmacol. Rev.* **2005**, *57*, 27–77.

261   15. Su, B. J. Different cross-talk sites between the renin–angiotensin and the kallikrein–kinin systems *J. Renin-*
262   *Angiotensin-Aldosterone Syst.* **2014**, *15*, 319–328.

263   16. Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.;
264   Herrler, G.; Wu, N.-H.; Nitsche, A.; Müller, M. A.; Drosten, C.; Pöhlmann, S. SARS-CoV-2 Cell Entry
265   Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell.* **2020**, *181*,
266   1–10.

267   17. Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K.-Y.; Wang, Q.g;
268   Zhou, H.; Yan, J.; Qi, J. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell.*
269   **2020**, *181*, 894-904.

270   18. Hamming, I.; Timens, W.; Bulthuis, M.; Lely, A.; Navis, G.; van Goor, H. Tissue distribution of ACE2
271   protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J.*
272   *Pathol.* **2004**, *203*, 631–637.

273   19. Imai, Y.; Kuba, K.; Rao, S.; Huan, Y.; Guo, F.; Guan, B.; Yang, P.; Sarao, R.; Wada, T.; Leong-Poi, H.;
274   Crackower, M. A.; Fukamizu, A.; Hui, C.-C.; Hein, L.; Uhlig, S.; Slutsky, A. S.; Jiang, C.; Penninger, J. M.
275   Angiotensin-converting enzyme 2 protects from severe acute lung failure. *Nature.* **2005**, *436*, 112-116.

276   20. Ni, W.; Yang, X.; Yang, D.; Bao, J.; Li, R.; Xiao, Y.; Hou, C.; Wang, H.; Liu, J.; Yang, D.; Xu, Y.; Cao, Z.; Gao,
277   Z. Role of angiotensin-converting enzyme 2 (ACE2) in COVID-19. *Crit. Care.* **2020**, *24*, 422.

278   21. Sodhi, C. P.; Wohlford-Lenane, C.; Yamaguchi, Y.; Prindle, T.; Fulton, W. B.; Wang, S.; McCray, Jr.; P. B.;
279   Chappell, M.; Hackam, D. J.; Jia, H. Attenuation of pulmonary ACE2 activity impairs inactivation of des-
280   Arg9 bradykinin/BKB1R axis and facilitates LPS-induced neutrophil infiltration. *Am. J. Physiol. Lung Cell*
281   *Mol. Physiol.* **2018**, *314*, L17–L31.

282   22. van de Veerdonk, F. L.; Netea, M. G.; van Deuren, M.; van der Meer, J. WM.; de Mast, Q.; Brüggemann, R.
283   J.; van der Hoeven, H. Kallikrein-kinin blockade in patients with COVID-19 to prevent acute respiratory
284   distress syndrome. *eLife.* **2020**, *9*, e57555.

285   23. de Maat, S.; de Mast, Q.; Danser, A.H. J.; van de Veerdonk, F. L.; Maas, C. Impaired Breakdown of
286   Bradykinin and Its Metabolites as a Possible Cause for Pulmonary Edema in COVID-19 Infection. *Semin.*
287   *Thromb. Hemost.* **2020**, doi: 10.1055/s-0040-1712960. Online ahead of print.

288   24. Dagnino, A. . A.; Campos, M. M.; Silva, R. B. M. Kinins and Their Receptors in Infectious Diseases.
289   *Pharmaceuticals.* **2020**, *13*, 215.

290 25. Mahmudpour, M.; Roozbeh, J.; Keshavarz, M.; Farrokhi, S.; Nabipour, I. COVID-19 cytokine storm: The
291   anger of inflammation. *Cytokine*. **2020**, *133*, 155151.

292 26. Rameshrad, M.; Ghafoori, M.; Mohammadpour, A. H.; Nayeri, M. J. D.; Hosseinzadeh, H. A comprehensive
293   review on drug repositioning against coronavirus disease 2019 (COVID19). *Naunyn-Schmiedeberg's Archiv.*
294   *Pharmacol.* **2020**, *393*, 1137–1152.

295 27. Tolouian, R.; Vahed, S. Z.; Ghiyasvand, S.; Tolouian, A.; Ardalan, M. COVID-19 interactions with
296   angiotensin-converting enzyme 2 (ACE2) and the kinin system; looking at a potential treatment. *J. Renal*
297   *Inj. Prev.* **2020**, *9*, e19.

298 28. Marceau, F.; Regoli, D. Bradykinin Receptor Ligands: Therapeutic Perspectives. *Nat. Drug Discov.* **2004**, *3*,
299   845-852.

300 29. Bork, K.; Yasothan, U.; Kirkpatrick, P. Icatibant. *Nat. Rev. Drug Discov.* **2008**, *7*, 801-802.

301 30. https://clinicaltrials.gov/ct2/show/NCT04488081?term=icatibant&draw=3&rank=19 (accessed on July, 18th
302   **2020**).

303 31. Sree Gns, H.; GR, S.; Muraharia, M.; Krishnamurthy, M. An update on Drug Repurposing: Re-written saga
304   of the drug's fate. *Biomed. Pharmacother.* **2019**, *110*, 700-716.

305 32. Lupala, C.L.; Gomez-Gutierrez, P.; Perez, J.J. New insights into the stereochemical requirements of the
306   bradykinin B1 receptor antagonist binding. *J. Mol. Graphics Model.* **2016**, *68*,184–196.

307 33. Lupala, C.L.; Gomez-Gutierrez, P.; Perez, J.J. New insights into the stereochemical requirements of the
308   bradykinin B2 receptor antagonist binding. *J. Comp.-Aided Mol. Design*, **2016**, *30*, 85-101.

309 34. Rasaeifar, B.; Lupala, C. S.; Gomez-Gutierrez, P.; Perez, J. J. Molecular Features Characterizing Non-peptide
310   B1 and B2 Bradykinin Receptor Selectivity. *Bioorg. Med. Chem. Lett.;* **2019**, *29*, 11-14.

311 35. *Molecular Operating Environment (MOE) 2019.01; Chemical Computing Group UCL*: Montreal, QC, Canada.
312   **2020**.

313 36. Wishart, D.S.; Feunang, Y.D.; Guo, A.C.; Lo, E.J.; Marcu, A.; Grant, J.R.; Sajed, T.; Johnson, D.; Li, C.;
314   Sayeeda, Z.;, Assempour, N.; Iynkkaran, I,, Liu, Y.; Maciejewski, A.; Gale, N.; Wilson, A.; Chin, L.;
315   Cummings, R.; Le, D.; Pon, A.; Knox, C.; Wilson, M. DrugBank 5.0: a major update to the DrugBank
316   database for 2018. *Nucleic Acids Res.* **2018**, *46*, D1074-D1082.

317 37. Scott, J.A.; Da Camara, C.C.; Early, J.E. Raloxifene: a selective estrogen receptor modulator. *Am. Fam.*
318   *Physician.* **1999**, *60*, 1131-1139.

319 38. Unegbu, C.; Noje, C.; Coulson, J.D.; Segal, J.B.; Romer L. Pulmonary Hypertension Therapy and a
320   Systematic Review of Efficacy and Safety of PDE-5 Inhibitors. *Pediatrics.* **2017**, *139*, e20161450.

321 39. Patel, H.B.; Lusk, K.A.; Cota, J.M. The Role of Cefepime in the Treatment of Extended-Spectrum Beta-
322   Lactamase Infections. *J. Pharm. Pract.* **2019**, *32*, 458-463.

323 40. Hamazaki , H.; Hasegawa, H.; Horiuchi, A.; Teshima, H.; Hiraoka, A.; Masaoka, T.; Nasu, K.; Uchino, H.;
324   Tatsumi, N.; Inoue, N.; Kageyama, T.; Kawagoe, H.; Tukaguchi, M.; Hukuhara, S.; Takahashi, T.;
325   Takatsuka, H.; Kanamaru, A.; Kakishita, E.; Nagai, K.; Hara, H.; Kanayama, Y.; Sugiyama, H.; Kitani, T.
326   Clinical evaluation of cefpirome sulfate for severe infections in patients with hematological disorders.
327   Hanshin Study Group of Hematopoietic Disorders and Infections. *Jpn. J. Antibiot.* **1997**, *50*, 12-21.

328 41. Vigneri, P.; Wang, J.Y. Induction of apoptosis in chronic myelogenous leukemia cells through nuclear
329   entrapment of BCR-ABL tyrosine kinase. *Nat. Med.* **2001**, *7*, 228-234.

330 42. Zhou, T.; Commodore, L.; Huang, W.S.; Wang, Y.; Thomas, M.; Keats, J.; Xu, Q.H.; Rivera, V.M.;
331   Shakespeare, W.C.; Clackson, T.; Dalgarno, D.C.; Zhu, X.T. Structural mechanism of the Pan-BCR-ABL
332   inhibitor ponatinib (AP24534): lessons for overcoming kinase inhibitor resistance. *Chem. Biol. Drug Des.*
333   **2011**, *77*, 1–11.

334 43. Gelbert, L.M.; Cai, S.; Lin, X.; Sanchez-Martinez, C.; Del Prado, M.; Lallena, M.J.; Torres, R.; Ajamie, R.T.;
335   Wishart, G.N.; Flack, R.S.; Neubauer, B.L.; Young, J.; Chan, E.M.; Iversen, P.; Cronier, D.; Kreklau, E.; de
336   Dios, A. Preclinical characterization of the CDK4/6 inhibitor LY2835219: in-vivo cell cycle-
337   dependent/independent anti-tumor activities alone/in combination with gemcitabine. *Invest. New Drugs.*
338   **2014**, *32*, 825-37.

339 44. Rolfo, C.; Ruiz, R.; Giovannetti, E.; Gil-Bazo, I.; Russo, A.; Passiglia, F.; Giallombardo, M.; Peeters, M.; Raez,
340   L. Entrectinib: a potent new TRK, ROS1, and ALK inhibitor. *Expert Opin. Investig. Drugs.* **2015**, *24*, 1493-
341   1500.

342   45.   Wassmann, S.; Laufs, U.; Stamenkovic, D.; Linz, W.; Stasch, J.-P.; Ahlbory, K.; Rösen, R.; Böhm, M.;
343   Nickenig, G. Raloxifene Improves Endothelial Dysfunction in Hypertension by Reduced Oxidative Stress
344   and Enhanced Nitric Oxide Production. *Circulation*. **2002**, *10*, 2083-2091.
345   46.   Pretorius, M.; Brown, N. J. Endogenous Nitric Oxide Contributes to Bradykinin-Stimulated Glucose Uptake
346   but Attenuates Vascular Tissue-Type Plasminogen Activator Release. *J. Pharm. Exp. Ther.* **2010**, *332*, 291–
347   297.
348   47.   Wu, B.; Mol, C. D.; Han, G. W.; Katritch, V.; Chien, E. Y. T.; Liu, W.; Cherezov, V.; Stevens, R. C. Structures
349   of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science*. **2010**, *330*,
350   1066-1071.
351   48.   Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**,
352   *234*, 779–815.
353   49.   Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J. GROMACS: Fast flexible
354   and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718.
355   50.   Lupala, C.S.; Rasaifar, B.; Gomez-Gutierrez, P.; Perez, J.J. Using Molecular Dynamics for the refinement of
356   atomistic models of GPCRs by homology modeling. *J. Biomol. Struct. Dyn.* **2018**, *36*, 2436–2448.
357   51.   Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.;
358   Shaw, D.E.; Shelley, M.; et al. Glide: A New Approach for Rapid Accurate Docking and Scoring. 1. Method
359   and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
360

361