UNIVERSITÀ DEGLI STUDI DI PADOVA

SCUOLA DI DOTTORATO IN SCIENZE MOLECOLARI
INDIRIZZO SCIENZE FARMACEUTICHE
CICLO XXI

TESI DI DOTTORATO

G Protein-Coupled Receptors as Potential Drug Target: From Receptor Topology to Rational Drug Design, an in-silico Approach

DIRETTORE DELLA SCUOLA: Prof. MAURIZIO CASARIN

SUPERVISORE: Prof. STEFANO MORO

DOTTORANDA: ERIKA MORIZZO

31 GENNAIO 2009

Contents

List of Figures

List of Abbreviations

ACO	Ant Colony Optimization
AR	Adenosine Receptor
CGS	2-[4-(2-carboxyethyl)phenethyl]amino-5'-(N-ethylcarbamoyl)adenosine
DPCPX	8-cyclopenyl-1,3-dipropylxanthine
EL	Extracellular Loop
GA	Genetic Algorithm
GPCR	G Protein-Coupled Receptors
hβ_2-AR	human β_2-Adrenergic Receptor
hA$_1$AR	human A$_1$ Adenosine Receptor
hA$_3$AR	human A$_3$ Adenosine Receptor
hA$_{2A}$AR	human A$_{2A}$ Adenosine Receptor
hA$_{2B}$AR	human A$_{2B}$ Adenosine Receptor
I-AB-MECA	N^6-(4-amino-3-iodobenzyl)-5'-(N-methylcarbamoyl)adenosine
IL	Intracellular Loop
LBHM	Ligand-Based Homology Modeling
LGA	Lamarckian Genetic Algorithm
MD	Molecular Dynamics
MOE	Molecular Operating Environment
NECA	5'-(N-ethylcaboxamido)adenosine
PTP	Pyrido-Triazolo-Pyrazine
PYRM	Pyrimidine
QN	Quinoline
QZ	Quinazoline
RBHM	Rhodopsin-Based Homology Modeling
RMSD	Root Mean Square Deviation
SA	Simulated Annealing
SAR	Structure-Activity Relationship
tβ_1-AR	turkey β_1-Adrenergic Receptor
TM	Transmembrane
TQX	Triazoloquinozalinone
TS	Tabu Search

Abstract

G Protein-Coupled Receptors as Potential Drug Target: From Receptor Topology to Rational Drug Design, an in-silico Approach

Abstract: G protein-coupled receptors (GPCRs) constitute a very large family of heptahelical, integral membrane proteins that mediate a wide variety of physiological processes, ranging from the transmission of the light and odorant signals to the mediation of neurotransmission and hormonal actions. GPCRs are dysfunctional or deregulated in several human diseases and are estimated to be the target of more than 40% of drugs used in clinical medicine today.

The crystal structures of rhodopsin and the recent published crystal structures of human β_2-adrenergic receptor and human A_{2A} Adrenergic Receptor provide the information of the three-dimensional structure of GPCRs, which supports homology modeling studies and structure-based drug-design approaches. Rhodopsin-based homology modeling has represented for many years a widely used approach to built GPCR three-dimensional models. Structural models can be used to describe the interatomic interactions between ligand and receptor and how the binding information is transmitted through the receptor. Both agonist and antagonist like states can be described by several different conformational receptor states depending on the nature of both ligand and receptor. Considering different complementarities, we might explore different conformations of the same pharmacological state.

We investigated the molecular pharmacology of adenosine receptors and, in particular, the human A_3 adenosine receptor (hA$_3$AR) by using an interdisciplinary approach to speed up the discovery and structural refinement of new potent and selective hA$_3$AR antagonists. Human A$_3$AR belongs to adenosine receptors family of GPCRs, which consists of four distinct subtypes: A_1, A_{2A}, A_{2B}, A_3 that are ubiquitously expressed in the human body.
The hA$_3$AR, which is the most recently identified adenosine receptor, is implicated in a variety of important physiological processes. Activation of A$_3$ARs increases the release of inflammatory mediators, such as histamine from rodent mast cells, and it inhibits the production of tumor necrosis factor-α. The activation of the hA$_3$AR seems to be involved in immunosuppression and in the response to ischemia of the brain and heart. Agonists or antagonists of A$_3$ARs are potential therapeutic agents for the treatment of ischemic and inflammatory diseases.

The first model of human A_3AR has been built using a conventional rhodopsin-based homology modeling approach. The model has been used to probe atomic level specific interactions, detected using site-directed mutagenesis analysis. The rhodopsin-based model of the hA_3AR in its resting state (antagonist-like state) has been revisited, taking into account a novel strategy to simulate the possible receptor reorganization induce by the antagonist-binding. We called this new strategy ligand-based homology modeling (LBHM). It is an evolution of a conventional homology modeling algorithm: any selected atoms will be included in energy tests and in minimization stages of the modeling procedure. Ligand-based option is very useful when one wishes to build a homology model in the presence of a ligand docked to the primary template. Starting from the conventional rhodopsin-based homology model and applying our ligand-based homology modeling implementation we can generate other antagonist-like conformational states of hA_3AR in which the ligand recognition cavity is expanded. Using different antagonist-like conformational states, we are able to rationalize the observed activities for all the compounds analyzed. Many severe analysis concerning false-positives and false-negatives situations are usually conducted.

To strictly validate this methodology as novel tool to address the multi-conformational space of GPCRs, we have analyzed different classes of known human A_3 antagonists in the corresponding putative ligand binding site: for example triazoloquinoxalin-1-one derivatives, arylpyrazolo-quinoline derivatives and pyrazolo-triazolo-pyrimidines derivatives. These studies led to the identification of groups for every class of antagonists that, introduced one by one in a suitable position, afford high hA_3AR affinity and good selectivity.

Starting from these binding requirements, we decided to perform an in silico molecular simplification approach to identify a suitable fragmentation route of the 4-amino-triazoloquinoxalin-1-one scaffold and explore which of the structural features were essential to guarantee efficient ligand-receptor recognition.

With the availability of new three dimensional templates different from rhodopsin, we built new models of hA_3AR. All the models were used for a molecular dynamic simulation in a POPC bilayer to investigate the topological fluctuation of the binding pocket.

Keywords: GPCR, A_3 Adenosine Receptor, Adenosine Receptor Antagonists, Molecular Docking, Homology Modeling, Ligand Based Homology Modeling, Molecular Dynamics.

Riassunto

I recettori accoppiati alle proteine G come potenziali bersagli terapeutici: dalla topologia recettoriale alla progettazione di nuovi ligandi, un approccio in-silico.

Riassunto: I recettori accoppiati alle proteine G (GPCR) costituiscono una grande famiglia di proteine integrali di membrana caratterizzate da sette eliche transmenmbrana, che mediano un'ampia gamma di processi fisiologici che vanno dalla trasmissione della luce e dei segnali olfattivi alla mediazione della neurotrasmissione e dell'azione degli ormoni. I GPCR mancano di una corretta regolazione in molte patologie umane ed è stato stimato che costituiscano il target del 40% dei medicinali utilizzati attualmente in clinica.

La struttura cristallografica della rodopsina e le strutture più recenti del recettore β adrenergico e del recettore adenosinico A_{2A} forniscono l'informazione strutturale che sta alla base della costruzione di modelli per omologia e degli approcci di structure-based drug design dei GPCR. La costruzione di modelli di GPCR per omologia basati sulla struttura della rodopsina ha rappresentato per molti anni un approccio ampiamente utilizzato. Questi modelli possono essere usati per descrivere le interazioni interatomiche tra ligando e recettore e come le informazioni sono trasmesse attraverso il recettore. Diversi stati conformazionali del recettore possono essere in grado di descrivere la conformazione del recettore che lega l'agonista e quella che lega l'antagonista, a seconda della natura di ligando e recettore. Se si considerano diverse complementarietà, si possono esplorare diversi stati conformazionali di uno stesso stato farmacologico.

Noi abbiamo studiato la farmacologia molecolare dei recettori adenosinici e, in particolare, del recettore adenosinico A_3 umano (hA_3AR), utilizzando un approccio interdisciplinare al fine di massimizzare la scoperta e l'ottimizzazione strutturale di nuovi antagonisti potenti e selettivi per il hA_3AR. Il hA_3AR fa parte della famiglia dei recettori adenosinici che consiste in quattro diversi sottotipi (A_1, A_{2A}, A_{2B}, A_3) che sono espressi in tutto il corpo umano. Il recettore adenosinico A_3 è stato identificato più recentemente ed è implicato in importanti processi fisologici. L'attivazione del hA_3AR aumenta il rilascio di mediatori dell'infiammazione, come l'istamina dalle mast-cellule, e inibisce la produzione del TNF-α. L'attivazione del hA_3AR sembra essere coinvolta nell'immunosoppressione e nella risposta ischemica di cuore e cervello. Agonisti o antagonisti del hA_3AR sono potenziali agenti terapeutici

nel trattamento di patologie ischemiche e infiammatorie.

Il primo modello di hA$_3$AR è stato costruito usando un approccio convenzionale di homology modeling basato sulla rodopsina ed è nel suo stato che lega l'antagonista. Dopo essere stato utilizzato per verificare le interazioni a livello molecolare che erano state evidenziate da studi di mutagenesi, il modello è stato rivisto prendendo in considerazione una nuova strategia che simula la possibile riorganizzazione del recettore indotta dal legame con l'antagonista. Abbiamo chiamato questa strategia ligand-based homology modeling. È un'evoluzione dell'algoritmo convenzionale di homology modeling: ogni atomo selezionato viente preso in considerazione nei test energetici e nelle fasi di minimizzazione della procedura di modeling. L'opzione ligand-based è molto utile quando si vuole costruire un modello per omologia in presenza di un ligando nella sua ipotetica conformazione di legame nel templato iniziale. A partire dal modello ottenuto dalla rodopsina e applicando la tecnica del LBHM, possiamo generare altri stati conformazionali del recettore hA$_3$AR che legano l'antagonista, nei quali la cavità di riconoscimento del ligando è espansa. Usando diversi stati conformazionali che legano l'antagonista, possiamo razionalizzare l'attività misurata sperimentalmente di tutti i composti analizzati. Sono condotte severe analisi relative a falsi positivi e falsi negativi.

Per validare la metodologia come nuovo strumento per indirizzare lo spazio multiconformazionale dei GPCR, abbiamo analizzato diverse classi di antagonisti con attività nota sul hA$_3$AR: ad esempio derivati triazolochinossalinonici, derivati arilpirazolo-chinolinici e derivati pirazolo-triazolo-pirimidinici. Questi studi hanno portato all'identificazione di gruppi per ogni classe di antagonisti che, se introdotti in una precisa posizione, portano ad un'alta affinità e ad una buona selettività per il hA$_3$AR.

A partire dalle caratteristiche risultate importanti per il legame, abbiamo applicato una tecnica di semplificazione molecolare in silico per identificare una possibile via di frammentazione della struttura 4-amino-triazolochinoassalin-1-onica ed esplorare quali sono le caratteristiche strutturali essenziali per garantire un'efficiente riconoscimento ligando-recettore.

Con la disponibilità di nuove strutture tridimensionali da utilizzare come templati diversi dalla rodopsina, abbiamo costruito nuovi modelli del recettore hA$_3$AR. Tutti i modelli sono stati usati per una simulazione di dinamica molecolare in un doppio strato fosfolipidico, per analizzare le fluttuazioni topologiche della tasca di legame.

Parole Chiave: GPCR, Recettore Adenosinico A$_3$, Docking Molecolare, Homology Modeling, Ligand Based Homology Modeling, Dinamica Molecolare

Introduction

1.1   G Protein-Coupled Receptors

G Protein-Coupled Receptors (GPCRs) are among the largest and most important family of signal transduction membrane proteins. GPCRs represent an efficient signaling system used by cells to transmit molecular information from the extracellular side to the intracellular side. [1,2]

They play a crucial role in many essential physiological processes, ranging from the transmission of the light and odorant signals to the mediation of neurotransmission, hormonal actions, cell growth and immune defense. GPCRs mediate responses interacting with a variety of bioactive molecules including ions, lipids, aminoacids, peptides, proteins and small organic molecules. [3,4] Signal transduction is controlled by GPCRs: the agonist binding promotes allosteric interactions between the receptor and the G protein, that catalyses the GDP-GTP exchange and transfer the signal to intracellular effectors, such as enzymes and ions channels. (Figure 1.1) [5,6]
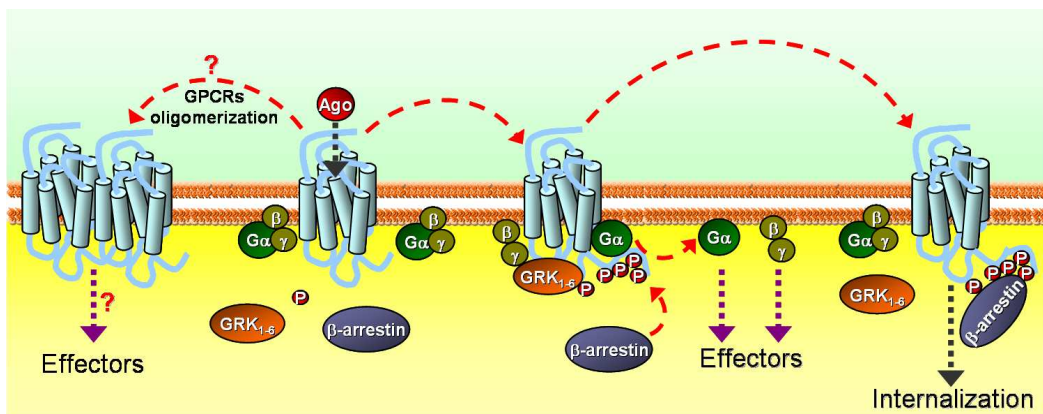
Figure 1.1: GPCR signaling.

However, GPCRs interact also with several other important proteins involved in the control of cellular homeostasis such as arrestins, [7,8] or PDZ domain-containing proteins. [9] In particular, cytosolic proteins of the arrestin family bind specifically to GPCRs phosphorilated by G protein-coupled receptor kinases (GRKs). [10] This complex (phosphorilated receptor/arrestin)

prevents the further coupling of that receptor to its G protein, reducing over time the capacity of second messenger synthesis. However, arrestins serve equally important roles in regulation internalization and alternative signaling events. [10]

The signaling pattern of GPCRs can be generated bypassing G protein intervention. It is generally accepted that GPCRs can lead to a dimeric or multimeric quaternary structure that plays a role in G protein independent signaling, although the exact mechanism are not entirely elucidated. Increasing evidence suggests that many GPCRs exist as homodimers and heterodimers and their oligomeric assembly could have important functional roles. [11,12] Key questions that remain to be answered include the prevalence and relevance of these in native tissue and the implications of heterodimerization for pharmacology and, potentially, for drug design. [13]

The total number of GPCRs with and without introns in the human genome was estimated to be approximately 950, of which 500 are odorant or taste receptors and 450 are receptors for endogenous ligands (approximately 2% of the coding genes). [14]
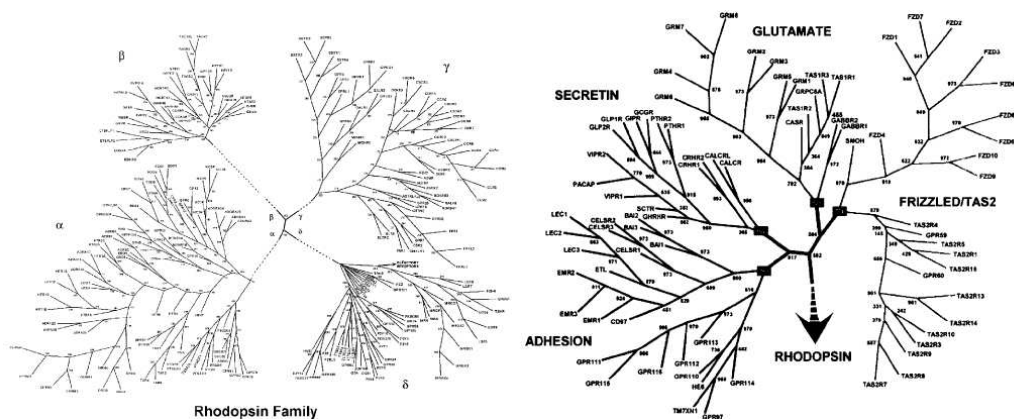
Figure 1.2: On the left: phylogenetic relationship between the GPCRs in the human genome. On the right: the phylogenetic relationship between GPCRs in the human rhodopsin family.

Several classification systems have been used to sort out this superfamily (Figure 1.2). According to sequence analyses, GPCRs have been clustered in a number of family or classes. The different classification systems include the A to F system, the 1 to 5 system and the GRAFS system. Thus the A (named 1 or rhodopsin in the 1 to 5 or the GRAFS system, respectively) is the rhodopsin-like class/family; B (or 2 or secretin) is the secretin class/family; C (3 or glutamate) is the metabotropic glutamate and pheromone class/family; D (or 4) is the fungal pheromone class/family; [15] E is the cAMP receptor class/family; and F (or 5 or frizzled) is the frizzled/smoothened family. [4,16, 17] Family A is by far the largest and the most studied. The overall homology
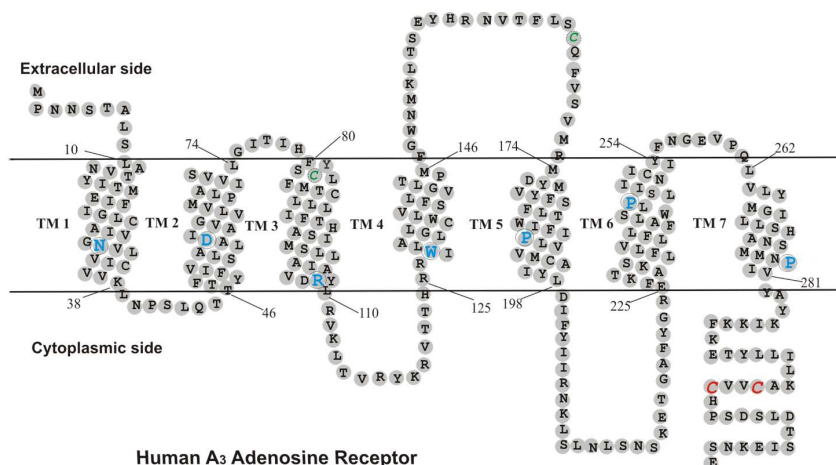
Figure 1.3: Schematic representation of the membrane topology of the human A_3 adenosine receptor. Each of the 7 TMs have at least one characteristic residue (blue colour), which is found among the majority of family A receptors (Asn30(1.50); Asp58(2.50); Arg108(3.50); Trp135(4.50); Pro189(5.50); Pro245(6.50); and Pro279(7.50)). Disulfide bridge formation between Cys83 (3.25) and Cys166 (EL2) (green colour), palmitoylation sites (Cys300 and/or 303, red colour) in the C terminus.

among all family A receptors is low and restricted to a small number of highly conserved key residues distributed in each of the seven helices. [4,16,17]

Usually with native GPCRs, activation is initiated by agonist binding. However, GPCRs can achieve the active states independently of agonists, that is, they can become constitutively active. Constitutively active GPCRs can be involved in the pathogenesis of human diseases and they are also invaluable tools to discover the signal transduction pathways of hundreds of orphan GPCRs, which are potential targets of novel drugs. [18] On the other hand, a number of constitutively active GPCR mutants have been found, which are involved in the pathogenesis of human disease. [19,20]

Disregulation of GPCRs has been found in a growing number of human diseases, [21,22] and GPCRs have been estimated to be the target of about half of the drugs used in clinical medicine today. Thus understanding how GPCRs function at the molecular level is an important goal of biological research. [23,24]

Some fundamental structural features are common to members of family A GPCRs. Sequence comparison among GPCRs revealed the presence of different receptor families that does not share sequence similarity even if specific fingerprints exist in all GPCR classes.

All GPCRs have in common a central core domain consisting of seven transmembrane helices (TM1 to TM7) that are connected by three intracellular (IL1, IL2 and IL3) and three extracellular (EL1, EL2 and EL3) loops. Two cysteine residues (one in TM3 and one in EL2), which are conserved in most

GPCRs, form a disulfide link.

Each TM region contains at least one highly conserved residue. This residue is used as reference for the Ballesteros and Weinstein nomenclature system: every amino acid of TM regions is identified by a number that refers to the transmembrane segment of the GPCR, followed by a number that refers to the position relative to reference residue that has arbitrarily the number 50 (Asn1.50, Asp2.50, Arg3.50, Trp4.50, Pro5.50, Pro6.50 and Pro7.50 in TM1-7, respectively). [25]

Aside from sequence variation, GPCRs differ in the length and function of their N-terminal extracellular domain, their C-terminal intracellular domain and their intra- and extracellular loops. Each of these domains provides very specific properties to these receptor proteins (Figure 1.3).

1.2   Structural features of crystal structures of GPCRs

The evolution of the field of computer-aided design of GPCR ligands (both agonists and antagonists) has depended on the availability of a suitable molecular receptor template. Despite the enormous biomedical relevance of GPCRs, high resolution structural information on their active and inactive states is still lacking.

An elucidation of structural features of available class A GPCRs structures has been recently published by Mustafi and Palczewski. [26] The GPCRs structures available in the Protein Data Bank [27] are listed in table 1.1.

1.2.1   Rhodopsin - Crystal Structures

Rhodopsin had represented for many years the only structural information available for GPCRs and it had been widely used as template for the resting state of members of family A. [46]

The first highly resolved structure of rhodopsin was published by Palczewski and collaborators in 2000. [28] The 2.8 resolution structure, deposited in the Protein Data Bank under the identifier 1F88, showed all major structural features as predicted from years of biochemical, biophysical and bioinformatics studies and presented the same overall topology of bacteriorhodopsin. The arrangements of seven helices of bovine rhodopsin and the one of bacterial rhodopsin were found to be different. The structure of rhodopsin presents more organized extramembrane region than that of bacteriorhodopsins, demonstrating the functional differences between these two retinal binding proteins. Rhodopsin is composed of the protein opsin covalently linked to 11-*cis*-retinal through Lys296. The molecule size of bovine rhodopsin is intermediate among the members of the GPCR family.

Table 1.1: GPCRs crystal structures available in the Protein Data Bank.

PDB ID	Release Date	Resolution	GPCR
1F88	8/4/2000	2.80	Bovine Rhodopsin [28]
1HZX	7/4/2001	2.80	Bovine Rhodopsin [29]
1L9H	5/15/2002	2.60	Bovine Rhodopsin [30]
1GZM	11/20/2003	2.65	Bovine Rhodopsin [31]
1U19	10/12/2004	2.20	Bovine Rhodopsin [32]
2HPY	8/22/2006	2.80	Bovine Rhodopsin [33]
2G87	9/2/2006	2.60	Bovine Rhodopsin [34]
2I35	10/17/2006	3.80	Bovine Rhodopsin [35]
2I36	10/17/2006	4.10	Bovine Rhodopsin [35]
2I37	10/17/2006	4.15	Bovine Rhodopsin [35]
2J4Y	9/25/2007	3.40	Bovine Rhodopsin [36]
2PED	10/30/2007	2.95	Bovine 9-cis-Rhodopsin [37]
2RH1	10/30/2007	2.40	Human β2-Adrenergic Receptor [38]
2R4R	11/6/2007	3.40	Human β2-Adrenergic Receptor [39]
2ZIY	5/6/2008	3.70	Squid rhodopsin [40]
2Z73	5/13/2008	2.50	Squid rhodopsin [40]
3D4S	6/17/2008	2.80	Human β2-Adrenergic Receptor [41]
3CAP	6/24/2008	2.90	Bovine Opsin [42]
2VT4	6/24/2008	2.70	Turkey β1-Adrenergic Receptor [43]
3DQB	9/23/2008	3.20	Bovine Opsin [44]
3EML	14/10/2008	2.60	Human A_{2A} Adenosine Receptor [45]

The protein contains 348 amino acids and it folds into seven TM helices: the structure include 194 residues that make up seven TM helices (35 to 64 for TM1, 71 to 110 for TM2, 107 to 139 for TM3, 151 to 173 for TM4, 200 to 225 for TM5, 247 to 277 for TM6 and 286 to 306 for TM7). In addition to these helices, a short helix is located at the cytosolic end of TM7, perpendicular to the membrane, and it is called helix 8 (HX8). Helices 1, 4, 6 and 7 are bent at proline residues.

The extracellular and intracellular regions of rhodopsin consist of three inter-helical loops as well as two tails, N-term and C-term respectively.

Intra- and extracellular domains present a clear contrast concerning the packing: whereas ELs associate significantly with each other and with the N-term, only few interactions are observed among the ILs. In particular, while EL1 and EL2 run along the periphery of the molecule, a part of EL2 folds deeply into the center of rhodopsin. Residues Arg177 to Glu181 form an antiparallel β-sheet with residues Ser186 to Asp190, which is deeper inside the molecule and is just below the 11-cis-retinal and is a part of the chromophore-binding pocket. Cys187 (EL2) forms a disulfide bond with Cys110 (3.25) at the ex-

tracellular end of TM3. The cytoplasmic loops were poorly determined in the structures. This is the region with the highest B-factor and these loops are probably mobile in solution. In the structure 1F88 residues are missing in IL3 from 236 to 239 and in the C-term from 328 to 333. [28]

It should be noted that the IL3 is known to vary considerably among related GPCRs, so the flexibility and variability of this region may be critical for functionality and specificity in G-protein activation.
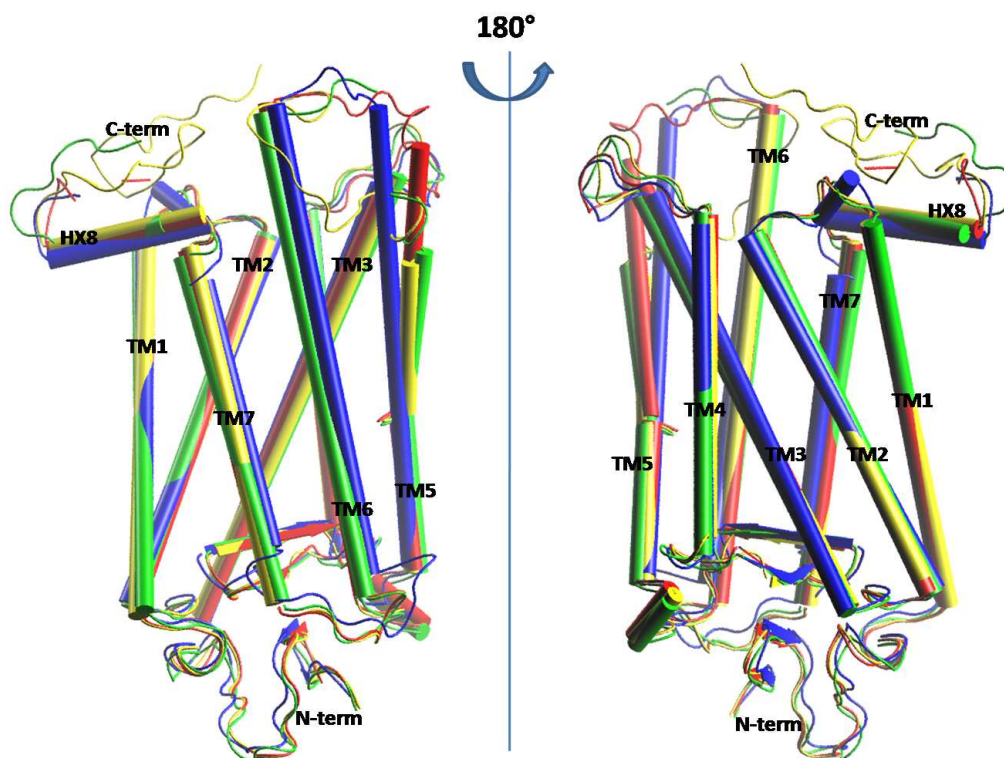
Figure 1.4: Side view, parallel to the membrane surface, of the superimposed structures of bovine rhodopsin: 1GZM in red, 1U19 in yellow, 2I37 in green (bovine meta II-like rhodopsin, photoactivated), 3DQB in blue (bovine opsine). The intracellular side is at the top. The main differences are in the intracellular side and, in particular, in the IL2 between TM3 and TM4, in the IL3 between TM5 and TM6 and in the C-term.

Further refinement of rhodopsin and 11-*cis*-retinal generated crystallographic structure deposited in the PDB under the identifier 1HZX. [29] Differences between 1F88 and 1HZX structures are located mainly in the IL2 and C-term.

Improved resolution was obtained with the following crystal structures that were published from 2002 to 2004: 1L9H (2.60 Å resolution), [30] 1GZM (2.65 Å resolution) [31] and 1U19 (2.20 Å resolution). [32] The crystal structure

IL9H provided a more detailed view of the TM region where several water molecules are found to play critical roles. [30]

Improvement of the resolution limit to 2.2 Å has been achieved by new crystallization conditions of 1U19 that completed the description of the protein backbone and is in general agreement with earlier diffraction studies. In this structure, structural information of IL3 and C-term are complete and the structure of the 11-*cis*-retinal chromophore and its binding site have been defined with greater precision, including the configuration about C6-C7 single bond of the 11-*cis*-retinal Schiff base and revealing significant negative pre-twist of the C11-C12 double bond, which is suggested to be critical for the function of rhodopsin. [32]

Li and coworkers determined the structure 1GZM of bovine rhodopsin at 2.65 Å resolution using untwinned native crystals in the space group P31. The new structure revealed mechanistically important details unresolved previously. New water molecules were identified and they extended H-bonding networks. The main difference with previously reported structures is in the intracellular side: the IL2 (residues 141-149) is L-shaped in both crystal forms, but lies more parallel with the membrane surface in 1GZM, the cytoplasmic ends of TM5 and TM6 have been extended by one turn, therefore the IL3 loop is elevated above the membrane surface like a spiral extension of helix 5. [31]

In the phototransduction cascade, rhodopsin plays a key role. Upon absorption of a photon, isomerization of the cromophore, 11-*cis*-retinal, to an all-trans conformation induces changes in the opsin structure, converting it from an inactive to an activated signaling state that interacts with the G protein. Rhodopsin progresses through a series of photointemediates that present different shape and dissimilar retinal ligands. Three dimensional structures of bathorhodopsin and lumirhodopsin were obtained by Nakamichi and Okada in 2006 and they are deposited in the PDB under the identifiers 2HPY [33] and 2G87. [34]
Equilibrium is formed between the later photointermediates MI and MII. MII correspond to the fully activated receptor. Advances in purification protocol and crystallization conditions permitted to Salom at al. the growth of ground state crystals that upon exposure to light transformed rhodopsin into a photoactivated deprotonated intermediate resembling the MII biological state. This structure (PDB ID 2I37) presents a resolution of 4.1 Å that results in lack of resolved residues. The photoactivated structure did not have residues Val230 to Gln238, Lys311 to Phe313 and Asp330 to Ala248 resolved. The x-ray crystallographic data reveal that the dimer is stabilized by a series of intermolecular contacts previously observed in other three dimensional structures but rotated by 180°around a hydrophobic center. [35]

In 2007 was resolved the first structure of a recombinantly produced G protein-coupled receptor (PDB ID 2J4Y). [36] The mutant N2C/D282C was
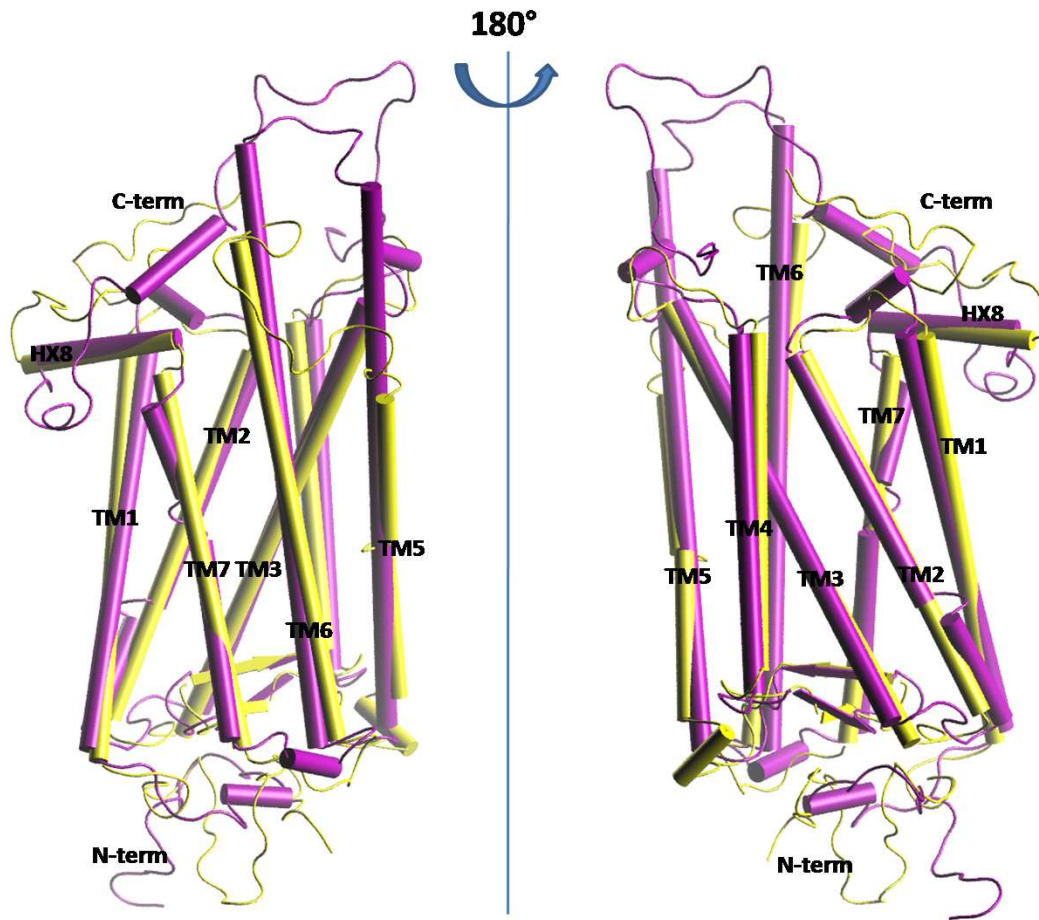
Figure 1.5: Side view, parallel to the membrane surface, of the superimposed structures of bovine rhodopsin (PDB ID 1U19) in yellow and squid rhodopsin (PDB ID 2ZIY) in magenta. The intracellular side is at the top.

designed to form a disulfide bond between the N-terminus and EL3. The disulfide introduces only minor changes but fixes the N-terminal cap over the β-sheet lid covering the ligand binding site. Moreover the structure of isorhodopsin was solved in which the native 11-*cis*-retinal of rhodopsin is replaced with the analog 9-*cis*-retinal (PDB ID 2PED). No significant structural differences were noted between rhodopsin and isorhodopsin. [37]

In 2008 the discovery of x-ray crystallographic structure of squid rhodopsin elucidated the differences between invertebrate and vertebrate structures. Two structures are available: 2ZIY (3.70 Å resolution) [40] and 2Z73 (2.50 Å resolution). [47] Squid rhodopsin contains a well structured cytoplasmic region involved in the interaction with G-proteins. TM5 and TM6 are longer and extrude into the cytoplasm. The distal C-terminal tail contains a short hydrophilic α-helix after the palmitoylated cysteine residues. The residues in
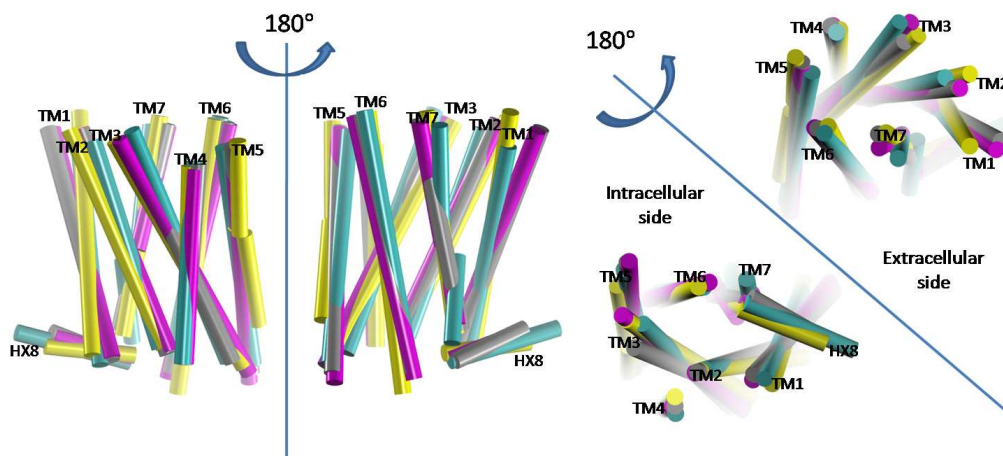
Figure 1.6: Superposition of the TM regions of the crystallographic structures of rhodopsin (PDB ID 1U19) in yellow, β_2-Adrenergic receptor (PDB ID 2RH1) in magenta, β_1-Adrenergic receptor (PDB ID 2VT4) in grey and A$_{2A}$ adenosine receptor (PCB ID 3EML) in cyan.

the distal C-term tail interact with the neighboring residues in the IL2, the extruded TM5 and TM6, and the short helix HX8 (Figure 1.5).

Two crystal structures of ligand-free native opsin from bovine retinal rod cells were solved in 2008: the 2.90 Å resolution structure published by Park et al. (PDB ID 3CAP) [42] and the 3.20 Å resolution structure published by Scheerer et al. (PDB ID 3DQB). [44] The structural analysis show only slight changes relative to rhodopsin for TM1 to TM4. The main differences are found in the intracellular ends of TM5, TM6 and TM7 and in the IL2 and IL3. These structural changes, some of which were attributed to an active GPCR state, reorganize the empty retinal-binding pocket to disclose two openings that may serve the entry and exit of retinal.

1.2.2 Beta Adrenergic Receptors - Crystal Structures

Adrenergic receptors belong to class A of GPCRs as well as rhodopsin. The crystal structure of a human β_2-adrenergic receptor-T4 lysozyme fusion protein bound to the partial inverse agonist carazolol at 2.4 Å resolution was firstly reported in 2007 by Cherezov, Rosenbaum and coworkers (PDB ID 2RH1). [38,48]

A 3.4Å/3.7Å resolution structure of human beta2 adrenergic receptor in a lipid environment, bound to the inverse agonist carazolol and in complex with a Fab that binds to the IL3 was also reported by Rasmussen, Choi and collaborators (PDB ID 2R4R). [39] The receptor was highly engineered, the protein was mutated and N-term and C-term were not resolved in the structures. Anyway the structurally conserved TM region provides a common
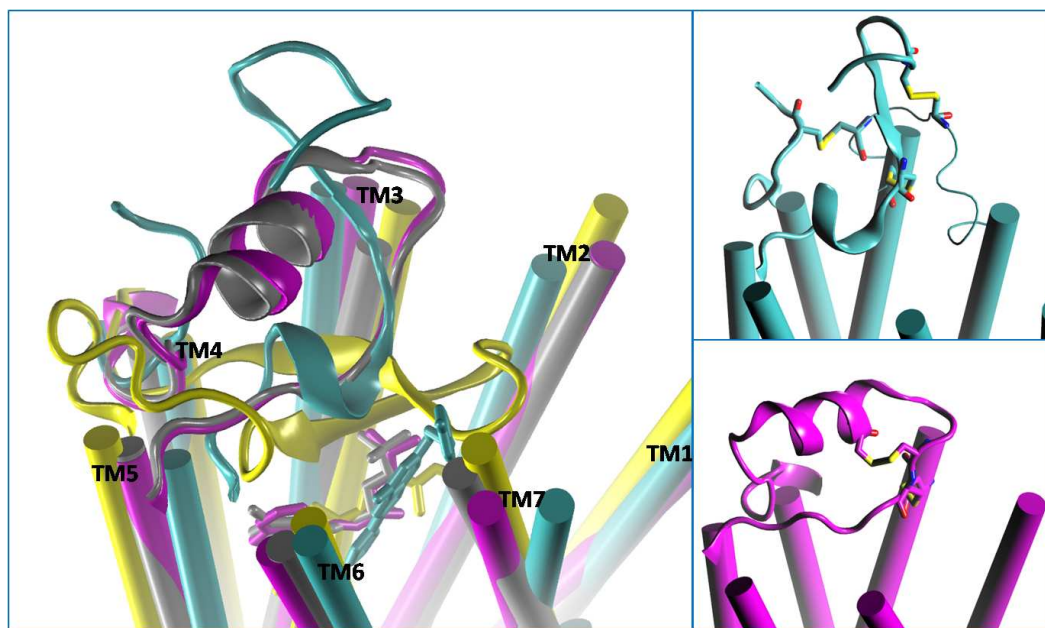
Figure 1.7: Representation of EL2. (left) TM regions of the superimposed structures of rhodopsin with retinal (PDB ID 1U19) in yellow, β_2-Adrenergic receptor with carazolol (PDB ID 2RH1) in magenta, β_1-Adrenergic receptor with cyanopindolol (PDB ID 2VT4) in grey and A_{2A} adenosine receptor with ZM241385 (PCB ID 3EML) in cyan. (right) On the top, representation of the TM regions and EL2 of A_{2A} adenosine receptor. Three disulfide bridges, one with TM3 and two with EL1 are highlighted. On the bottom, representation of the TM regions and EL2 of β_2-Adrenergic receptor. Two disulfide bridges are highlighted, one with TM3 and one internal link between two cysteine residues of EL2.

core with the one of rhodopsin (Figure 1.6). The structures provide a high-resolution view of a human G protein-coupled receptor bound to a diffusible ligand. Ligand-binding site accessibility is enabled by the EL2, which is held out of the binding cavity by a pair of closely spaced disulfide bridges and a short helical segment within the loop: in contrast to rhodopsin, β_2 adrenergic receptor presents a more open structure (Figure 1.7). The largest difference is in helix1, which is relatively straight and lacks the proline kink found in rhodopsin. Differences were shown also in the IL2 between rhodopsin and β_2-adrenergic receptor. No information are available for IL3 because the receptor was adapted to bind the T4 lysozyme in 2RH1 [38,48] and the Fab antibody in 2R4R. [39]

No significant structural differences were highlighted in the 2.8 Å resolution crystal structure of a thermally stabilized human β-adrenergic receptor bound to cholesterol and the partial inverse agonist timolol (PDB ID 3D4S). [41]

A crystallized mutant form of turkey β_1-adrenergic receptor in complex with high-affinity antagonist cyanopindolol is deposited in the Protein Data Bank under the identifier 2VT4. [43] In the protein six residues were mutated

and large portions of the structure were not resolved. In the crystal structure of turkey β_1-adrenergic receptor the IL2 forms a short α-helix parallel to the membrane surface. The conformation of the EL2 is similar to the one of β_2-adrenergic receptor and the binding pocket is open to the extracellular side.
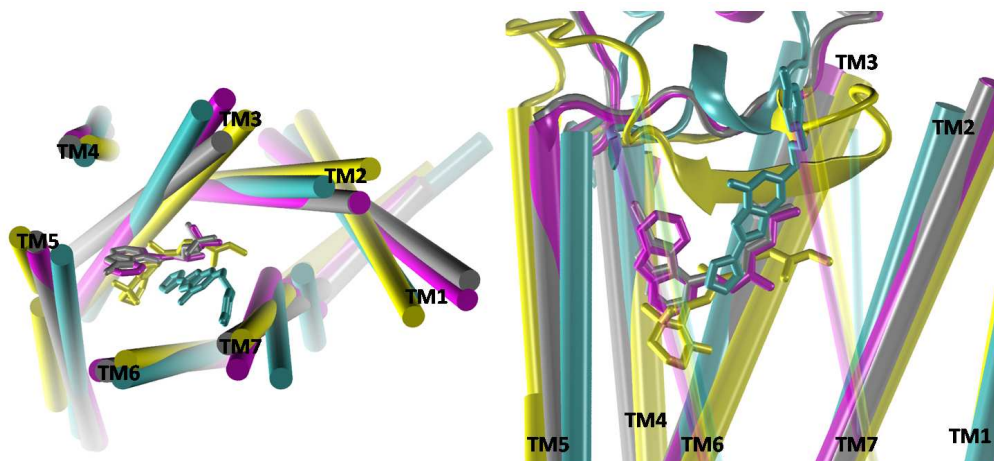
Figure 1.8: Position of ligands in the crystallographic structures of GPCRs. (left) Extracellular side view of the TM regions of the superimposed structures of rhodopsin with retinal (PDB ID 1U19) in yellow, β_2-Adrenergic receptor with carazolol (PDB ID 2RH1) in magenta, β_1-Adrenergic receptor with cyanopindolol (PDB ID 2VT4) in grey and A2A adenosine receptor with ZM241385 (PCB ID 3EML) in cyan. (right) Side view of the superimposed structures facing TM6 and TM7 (transparent). TM regions and EL2 are shown. The position of ZM241385 is significantly different from the position of retinal and amine ligands of β-adrenergic receptors, which are deeper in the binding pockets.

1.2.3   Adenosine Receptor - Crystal Structure

In 2008 the crystal structure of the human A_{2A} adenosine receptor in complex with a high-affinity subtype-selective antagonist, ZM241385, has been determined (PDB ID 3EML). [45] To crystallize the 2.60 Å resolution structure was applied the T4L fusion strategy, where most of the third cytoplasmic loop was replaced with lysozyme and the C-term tail was truncated from Ala317 to Ser412. This crystal structure presents three features different from previously reported GPCR structures. First, the EL2 is considerably different from β_1-AR, β_2-AR and bovine/squid rhodopsins and it lacks any clearly secondary structural element and possesses three disulfide linkages, one with TM3 (Cys77-Cys166) and two with EL1 (Cys71-Cys159 and Cys74-Cys146) (Figure 1.7). This contributes to the formation of a disulfide bond network that forms a rigid, open structure that allows the solvent to access the binding cavity. Secondly, ZM241385 is perpendicular to the membrane plane, co-linear with TM7 and it interacts with both EL2 and EL3. The ligand posi-
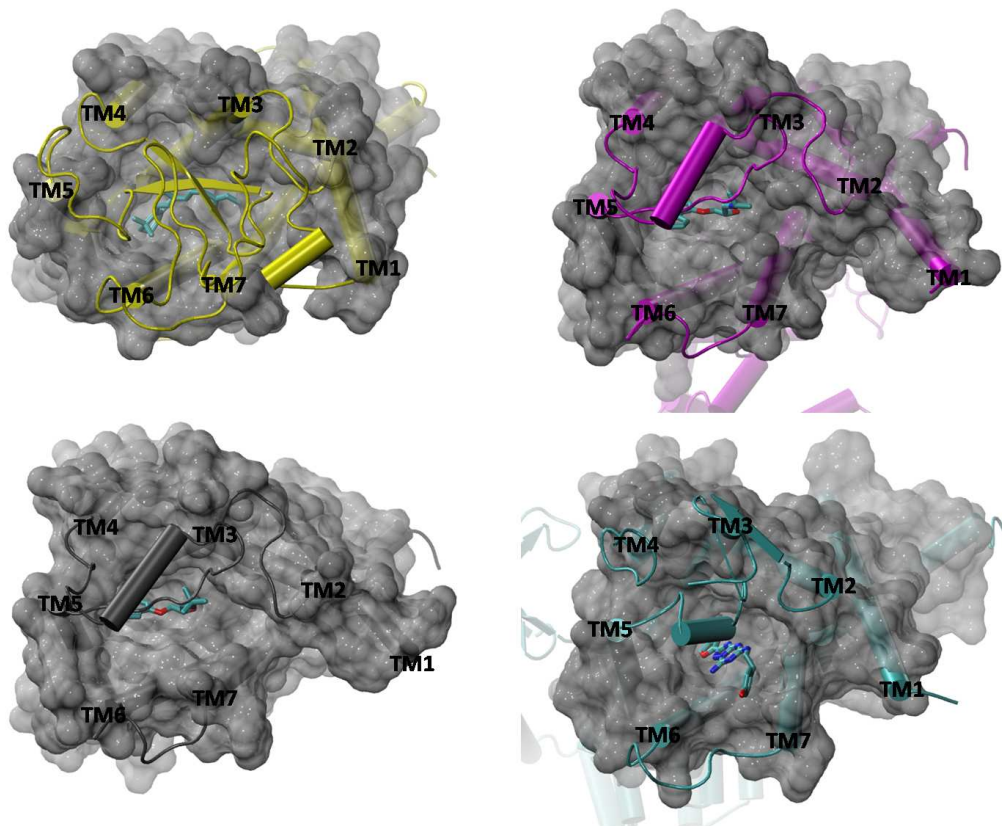
Figure 1.9: Extracellular side view of the crystal structures. On the top: bovine rhodopsin 1F88 (left), β_2-adrenergic receptor 2RH1 (right); on the bottom: β_1-adrenergic receptor 2VT4 (left), A_{2A} adenosine receptor 3EML (right). Backbones of the proteins are represented as cartoon, the TM regions are represented with a molecular surface and ligands are in stick.

tion is significantly different from the position of retinal and amine ligands of β adrenergic receptors (Figure 1.8). Finally, the helical arrangement is similar among GPCRs, however the binding pocket of the A_{2A} adenosine receptor is shifted closer to TM6 and TM7 and less interactions are allowed with TM3 and TM5 (Figure 1.9). [45]

1.3 Adenosine Receptors

A_3 adenosine receptors (ARs) belong to a small family of GPCRs, which consists of four distinct subtypes, A_1, A_{2A}, A_{2B}, and A_3 ARs are ubiquitously expressed in the human body. [49] Many cells express several ARs subtypes, although in different densities. All subtypes, including the A_3 receptor, have been cloned from a variety of species including rat and human. [49] Species differences for A_3 receptors are larger than for other ARs subtypes, particularly between rodent and human (h) receptors (only 74% sequence identity

between rat and hA$_3$ amino acid sequence). This results in different affinities of ligands, particularly antagonists, for rat versus hA$_3$ receptors.

A$_3$ ARs are negatively coupled to adenylate cyclase via G$_{i2,3}$. [49,50] Coupling of the A$_3$AR to G$_{q/11}$ leading to a stimulation of phospholipase C and its coupling to phospholipase D have also been demonstrated. [51] A$_3$AR stimulation can lead to activation of ERK1/2. In fact, A$_3$AR agonists stimulate PI3K-dependent phosphorylation of Akt leading to the reduction of basal levels of ERK1/2 phosphorylation, which in turn inhibits cell proliferation. [52] After exposure to agonist, A$_3$ARs undergo rapid desensitization via phosphorylation by G-protein receptor kinase 2 (GRK2) at the intracellular terminal chain (particularly at threonine 318 on the rat receptor). [53]
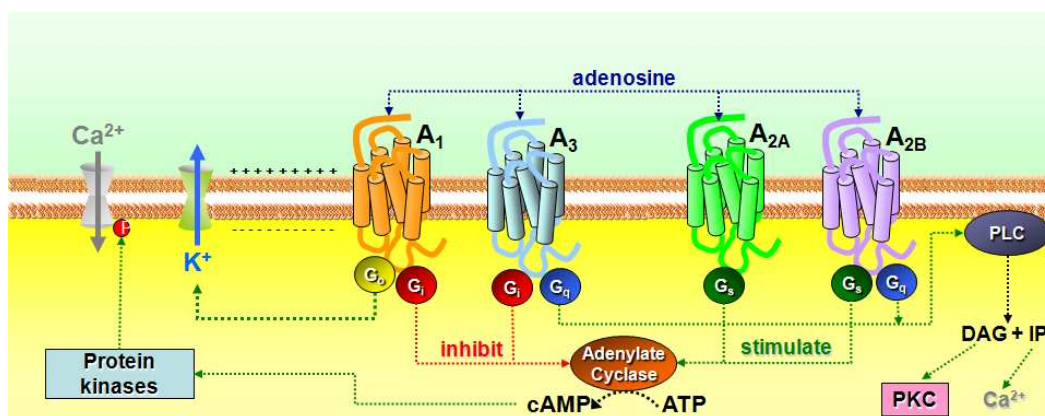
Figure 1.10: Signal transduction pathways associated with the activation of the human adenosine receptors.

The A$_3$AR, which is the most recently identified AR, is implicated in a variety of important physiological process. [50] Activation of the A$_3$AR increases the release of inflammatory mediators, such as histamine, from rodent mast cells, [54] and inhibits the production of tumor necrosis factor-α(TNF-α). [55] The activation of the A$_3$AR is also suggested to be involved in immunosuppression and in the response to ischemia of the brain and heart. [56] It is becoming increasingly apparent that agonists or antagonists of the A$_3$AR have potential as therapeutic agents for the treatment of ischemic and inflammatory diseases. [57]

1.4   Methodology Survey

The development of computers with increased calculation power gave to the scientific community new resources to develop data analysis and complex mathematical model building. In science, computers can be used to apply complex models to study different aspects of nature.

In this thesis, several computational tools were applied to study protein and other molecules, their interaction, their dynamics and to predict some of their behaviors. In this section the methods, which have been used in this project, are described as well as their strenghts and weakness.

1.4.1   Homology Modeling

Extensive information on primary and secondary structure are stored in various databases. Protein sequence determination is now routine work in molecular biology laboratories. Sequences of more than three million proteins are now available in the UniProt database [58]. The translation of sequences into 3D structure on the basis of X-ray crystallography or NMR investigations, however, takes much more time. The 3D structures of more than 55000 proteins available in the PDB [27,59] (as at the end of January 2009). In certain circumstances it can take, depending on the kind of proteins, more than a year to perform a complete structure determination. This is the reason why the number of known protein sequence is much larger than the number of complete 3D structures that have been determined.

Since a general rule for the folding of a protein has not yet been developed, it is necessary to base structural predictions on the conformations of available homologous reference proteins.

When a sequence is found homologous to another one, for which the 3D structure is available, the comparative modeling approach (which is also called *homology modeling approach*) is the method of choice for predicting the structure of the unknown protein. This computational approach is based on the notion that the primary structure of proteins is conserved, through evolution, to a lesser extent than the higher level structures, namely secondry, tertiary and quaternary.

An amino acid sequence (target) can be modeled on the structure of a second protein (template) which are predicted to have the same folding. Based on the sequence alignment of the two proteins, the pairs of residues are spatially matched with the generation of the new coordinates for the target structure. Thus, the quality of the sequence alignment which determines the residues pairs is of primary importance. Usually, conserved regions, like secondary structure elements or patterns of residues implicated in the protein function, are identified in the structure of the template. Later, the alignment is optimized to match these conserved regions. The out-coming structure can be structurally refined with different protocols like energy minimization or simulated annealing. The resulting structure has to be checked for sterochemical quality, like φ and ψ angles distributions and bond lengths, angles etc., and for its feasibility of explaining already available biochemical data.

In addition, when the alignment reveals one or more long gaps, underlining structural variations between the two proteins, care must be taken on

the structure generation. When new loops have to be built, meaning that the target sequence have non-correspondent stretches in the template, coordinates can be either assigned randomly and energy minimized or taken from experimentally known ones of other structures. The reliability of these additional loops depends on the length of these parts and the distance between the template extremities. The longer is the insertion, compared to the three-dimensional gap, the less reliable is the result [60,61].

1.4.2 Molecular Docking

Molecular Docking is a method that predicts the structure of the intermolecular complex formed between two or more molecules. Docking is frequently used to predict the binding orientation of small molecule drug candidates to their protein targets in order to predict the affinity and activity of the small molecule. Hence docking plays an important role in the rational design of drugs.

Reproducig the conformational space accessible to a macromolecule is a very difficult task and involves unavoidable approximation. Docking procedures can thus be classified into three categories depending on the approximation level:

- *rigid body docking*: both protein and ligand are treated as rigid bodies,

- *semiflexible docking*: only the ligand is condisered flexible,

- *fully flexible docking*: both ligand and protein are treated as flexible molecules.

Since ligands are much smaller than macromolecules, ligand flexibility is computationally easier to handle and thus today it is standard in docking routines.

The ideal docking methos would allow both ligand and receptor to explore their conformational degrees of freedom. However, such calculations are computationally very demanding and most of the methods only consider the conformational space of the ligand and the receptor is invariably assumed to be rigid.

The success of a docking program depends on two components: the *search algorithm* and the *scoring function*.

1.4.2.1 Search Algorithms

In molecular docking the search algorithm is used to generate ligand structures. The algorithms can be grouped into deterministic and stochastic approaches. Deterministic algorithms are reproducible, whereas stochastic algorithms include a random factor and are thus not fully reproducible.

Incremental Construction Methods In an incremental construction algorithm the ligand is not docked as a complete molecule at once, but is instead divided into single fragments and incrementally reconstructed inside the active site. FlexX treats the ligand as flexibe and the protein as rigid. It divedes the ligands along its rotational bonds into rigid fragments, first docks a base fragment into the active site and then reattaches the remaining fragments. FlexX degines interaction sites for each possible interacting group of the active site and the ligand. The interaction sites are assigned an interaction type (hydrogen bond acceptor, hydrogen bond donor, etc.) and are modeled by an interaction geometry consisting of an interaction center and a spherical surface. The base fragment is oriented by searching for placements where three interaction between the protein and the ligand can occur. The remaining ligand componetns are then incrementally attached to the core.

Genetic Algorithms A Genetic Algorithm is a computer program that mimics the process of evolution by manipulating a collection of data structures called *chromosomes*. Each of these chromosomes encodes a possible solution to the problem to be solved. *Gold* [62] and *MoeDock* [63] use GA for docking a ligand to a protein. Each chromosome encodes a possible protein-ligand complex conformation. Each chromosome is assigned a fitness score on the basis of the relative quality of that solution in terms of protein-ligand interactions. Starting from an initial, randomly generated parent population of chromosomes, the GA repeatedly applies two major genetic operators, crossover and mutation, resulting in children chromosomes that replace the least-fit member of the population. The crossover operator requires two parents and produces two children, whereas the mutation operator requires one parent and produces one child. Crossover thus combines features from two different chromosomes in one, whereas mutation introduces random perturbations. The parent chromosomes are randomly selected from the existing population with a bias toward the best, thus introducing an evolutionary pressure into the algorithm. This enphasis on the survival of the best individuals ensures that, over time, the population should move toward an optimal solution, that is to the correct binding mode. *AutoDock* 4.0 [64] uses a Lamarckian genetic algorithm (LGA). The characteristic of an LGA is that the environmental adaptation of an individual's phenotype are described into its genotype. In AutoDock 4.0 each generation is thus followed by a local search, enery minimization, on a user-defined proportion of the population and resulting ligand coordinates are stored in the chromosome, replacing the parent.

Tabu Search A Tabu search algorithms is characterized by imposing restrictions to enable a search process to negotiate otherwise difficult regions. These restrictions take the form of a tabu list that stores a number of previously

visited solutions. By preventing the search from revisiting these regions, the exploration of new search space is encouraged.

While GA usually converges quickly at the close proximity of a global minimum, it can be trapped in local minima. Using a tabu list helps in avoiding this drawback. TS is available as search algorithm in MoeDock [63].

Simulated Annealing Simulated Annealing is a special molecular dynamics simulation, in which the system is cooled down at regular time intervals by decreasing the simulation temperature. The system thus gets trapped in the nearest local minumum conformation. Disadvantage of simulated annealing are that the result depends on the initial placement of the ligand and that the algorithm doesn not explore the solution space exhaustively. SA is available as search algorithm in MoeDock [63].

Glide Algorithm The Glide (Grid-Based Ligand Docking With Energetics) [65] algorithm approximates a systematic search of positions, orientations, and conformations of the ligand in the receptor binding site using a series of hierarchical filters. The shape and properties of the receptor are represented on a grid by several different sets of fields that provide progressively more accurate scoring of the ligand pose. The fields are computed prior to docking. The binding site is defined by a rectangular box confining the translations of the mass center of the ligand. A set of initial ligand conformations is generated through exhaustive search of the torsional minima, and the conformers are clustered in a combinatorial fashion. Each cluster, characterized by a common conformation of the core and an exhaustive set of rotamer group conformations, is docked as a single object in the first stage. The search begins with a rough positioning and scoring phase that significantly narrows the search space and reduces the number of poses to be further considered to a few hundred. In the following stage, the selected poses are minimized on precomputed OPLS-AA van der Waals and electrostatic grids for the receptor. In the final stage, the 5-10 lowest-energy poses obtained in this fashion are subjected to a Monte Carlo procedure in which nearby torsional minima are examined, and the orientation of peripheral groups of the ligand is refined. The minimized poses are then rescored.

Plants The docking algorithm PLANTS is based on a class of stochastic optimization algorithms called ant colony optimization (ACO). ACO is inspired by the behavior of real ants finding a shortest path between their nest and a food source. The ants use indirect communication in the form of pheromone trails which mark paths between the nest and a food source. In the case of protein-ligand docking, an artificial ant colony is employed to find a minimum energy conformation of the ligand in the binding site. These ants are used

to mimic the behavior of real ants and mark low energy ligand conformations with pheromone trails. The artificial pheromone trail information is modified in subsequent iterations to generate low energy conformations with a higher probability. [66]

1.4.2.2   Scoring Function

The free energy of binding is given by the Gibbs-Helmoltz equation:

$$\Delta G = \Delta H - T\Delta S \tag{1.1}$$

with ΔG giving the free energy of binding, ΔH the enthalpy, T the temperature in Kelvin and ΔS the entropy. ΔG is related to the binding constant K_i by the equation

$$\Delta G = -RTlnK_i \tag{1.2}$$

with R being the gas constant. There is a wide variety of different techniques available for predicting the binding free energy of a small molecule ligand on the basis of the given 3D structure of a protein-ligand complex.

Empirical Scoring Function   Empirical scoring functions use several terms describing properties known to be important in drug binding to cunstruct a master equation for predicting binding affinity. Multilinear regression is used to optimize the coefficients to weight the computed terms using a training set of protein-ligand complexes for which both the binding and an experimentally detemined high resolution 3D structure are known. Chemscore and Glidescore are some examples.

Force-field-based Scoring Function   These scoring functions are based on the nonbonded terms of a classical molecular mechanics force field. A Lennard-Jones potential descibes van der Waals interactions, whereas the Coulomb energy describes the electrostatic components of the interactions. A major disadvantage of empirical scoring functions lies in the fact that it is unclear to what extent they can be applied to protein-ligand complexes that were not represented in the training set used for deriving the master equation. Goldscore and MOE Energy score are some examples.

Knowledge-based Scoring Function   A more recently developed approach avoiding these disadvantages uses knowledge-based scoring funtions with potential of mean force. The score is defined as the sum over all interatomic interactions of the protein-ligand complex. Advantages of this approach are that no fitting to experimentally measured binding free energies of the complexes in the training set is needed, and that solvation and entropic terms are treated implicitly.

1.4.3 Molecular Dynamics

Molecular systems, where non-bonded interactions between atoms are present, possess intrinsic movements due to the changing distribution of their internal energy. Theoretical and empirical studies of proteins should take into account their dynamical behaviors. Movements of proteins are understood as a variety of different atomic dispositions which are specific for each protein system and are ruled by physical-chemical properties such as steric hindrance of side chains or attractive and repulsive charges. In general, this molecular conformational changes can be either little, with simple structure fluctuations due to the energy present at a given temperature within the system, or large as consequence of major modifications, such as phosphorylation of residue and binding of ligands.

Molecules can be described by mathematical models where the atomic positions, radii, masses and charges as well as the covalent bonds (length, angles) of their topologies are considered.

In molecular dynamics, succesive configurations of the system are generated by integrating Newton's laws of motion. The result is a trajectory that specifies how the positions and velocities of the particles in the system vary with time. The trajectory is obtained by solving the differential equations embodied in Newton's second law ($F=ma$):

$$\frac{d^2 x_i}{dt^2} = \frac{F_{x_i}}{m_i} \tag{1.3}$$

This equation describes the motion of a particle of mass m_i along one coordinate (x_i) with F_{x_i} being the force in the particle in that direction. Initial atomic velocities are used to start the compute of the kinetic component. Forces are then used to calculate the new atomic positions and velocities by integration of the equation of motion after a defined period of time (time step). The iteration of this cycle yields to the deterministic evolution (dependent from the previous steps) of the system respect to the time.

The well known limitation of this method is how atoms are described. While using molecular mechanics (MM) model, the atoms of a simulated protein are described as balls with partial charges and the bonds are depicted as harmonic springs. The omission of all electrons speed up the calculation permitting longer time scale simulation but decrease the accuracy of the system evolution. Another issue of MD simulation is the lenght of the computed time life of a macromolecule. Certain biological phenomena concerning motions of proteins occur in a time scale which is not achievable by normal MD simulations.

The production of a trajectory usually involves three steps: the initialization of the system, its equilibration and production phase. During initialization velocities are given to the atoms to calculate the first round of forces.

When no velocities are available from a previous MD simulation, they are assigned randomly according to the Maxwell-Boltzmann distribution at given temperature. During equilibration the system is let evolve shortly to adjust velocities and to bring the system at the nearest thermic equilibrium an then the production phase.

Working with proteins some steps have to be added, this is due to the fact that these macromolecules are half way between liquid and solid state. In other words, the covalent bonds oscillations have to be restrained to reduce the number of degrees of freedom for the system. In the case that the solvent is wanted to be described explicitly in the trajectory, a certain number of water molecules have to added around the protein. The whole system needs to be energetically minimized to avoid bad steric contacts. Then a first round of MD is used to relax the solvent while the protein atoms are restrained in their initial positions. The next step consists in warming up the system, to the targeted temperature, i.e. 300 K, and to adjust the velocities. This is an important step for diminish the influence of the randomly assigned initial velocities in the final trajectory. The system is thus equilibrated for pressure and temperature using algorithms which every tot steps scale the velocities to match the set pressure and temperature within a given period of time. Eventually, the production phase is run and the system properties are collected for further analysis.

The reproducibility of this technique is an important issue because of the chaotic nature of multi-body dynamics. The several thousands particles affect the velocity of the single one by multiple interactions resulting in random trajectories. The word reproducibility is thus intended for averages of properties of the system calculated for relatively long simulations. Computational simulations of proteins should investigate a thermodynamic equilibrium of the system. The farther from the equilibrium the less reliable is the final trajectory.

Homology Modeling of Human A$_3$ Adenosine Receptor

2.1 Introduction

Rhodopsin was the first GPCR to be studied in detail. In 2000, the first three dimensional crystals of bovine rhodopsin were obtained. [67] These quickly led to a three dimensional high resolution structure for this GPCR, which for the first time provided a sufficiently detailed view that the disposition of the retinal in the structure could be determined. [28] Despite extensive efforts, rhodopsin had been for many years the only GPCR with structural information available. Rhodopsin is highly abundant from natural sources and structurally stabilized by the covalently bound ligand 11-*cis*-retinal, which maintains the receptor in a dark-adapted, non-signaling confromation. In contrast, all other GPCRs are activated by diffusible ligands and are expressed at relatively low levels in native tissues. These receptors are structurally more flexible and equilibrate among multiple conformational states, some of which are prone to instability. [68]

In the past few years several crystallographic structures of GPCRs, different from rhodopsin, were published. In 2007, Kobilka and coworkers resolved two crystallographic structures of human β_2-Adrenergic Receptor at 2.40 and 3.40 Å resolution. [38,39,48] In 2008 on PDB has been published another crystallographic structures: the one of human β_2 Adrenergic Receptor at 2.8 Å resolution [41], the structure of β_1-Adrenergic Receptor of turkey at 2.70 Å resolution [43] and recently the crystal structure of a human A$_{2A}$ Adenosine Receptor at 2.6 Å resolution. [45]

Some structures provide also information about interaction with a ligand. Human A$_{2A}$AR is the most different. The ligand ZM241385 is perpendicular to the membrane plane, co-linear with TM7 and it interacts with both EL2 and EL3. The ligand position is significantly different from the position of retinal and amine ligands of β-AR (Figure 1.8). Finally, the helical arrangement is similar among GPCRs, however the binding pocket of the A$_{2A}$ adenosine receptor is shifted closer to TM6 and TM7 and less interactions are allowed with TM3 and TM5. [45]

These structural information are the basis of homology modeling of hA$_3$AR. Structural models have been used for molecular docking (see Chapters

3 and 4) and molecular dynamics studies (see Chapter 5).

2.2   Materials and Methods

2.2.1   Sequence Allignement

Based on the assumption that GPCRs share similar TM boundaries and over-
all topology, a homology model of the hA_3 receptor was constructed. The
sequence of hA_3 receptor was retrieved from SwissProt Database [58] (ID:
P33765 [69,70]). First, the amino acid sequences of TM helices of the A_3 recep-
tor were aligned with those of the crystal structures selected [28,38,42,43,45],
guided by the highly conserved amino acid residues, including the DRY motif
(Asp3.49, Arg3.50, and Tyr3.51) and three proline residues (Pro4.60, Pro6.50,
and Pro7.50) in the TM segments of GPCRs.

2.2.2   Homology Modeling with MOE

The same boundaries were applied for the TM helices of the A_3 receptor as
they were identified from the X-ray crystal structure for the corresponding
sequences of the crystal structre used as template, the backbone coordinates
of which were used to construct the seven TM helices for the hA_3 receptor.
The loop domains of the hA_3 receptor were constructed by the loop search
method implemented in MOE.
In particular, loops are modeled first in random order. For each loop, a
contact energy function analyzes the list of candidates collected in the segment
searching stage, taking into account all atoms already modeled and any atoms
specified by the user as belonging to the model environment. These energies
are then used to make a Boltzmann-weighted choice from the candidates,
the coordinates of which are then copied to the model. Any missing side
chain atoms are modeled using the same procedure. Side chains belonging to
residues whose backbone coordinates were copied from a template are modeled
first, followed by side chains of modeled loops. Outgaps and their side chains
are modeled last.
Special caution has to be given to the second extracellular loop (EL2), which
can limit the size of the active site. Hence, amino acids of this loop could
be involved in direct interactions with the ligands. A driving force to this
peculiar fold of the EL2 loop might be the presence of a disulfide bridge
between cysteines in TM3 and EL2. Since this covalent link is conserved in
all receptors modeled in the current study, the EL2 loop was modeled using
a constrained geometry around the EL2-TM3 disulfide bridge.
After the heavy atoms were modeled, all hydrogen atoms were added, and
the protein coordinates were then minimized with MOE using the AMBER94
force field [71]. The minimizations were carried out by the 1000 steps of

steepest descent followed by conjugate gradient minimization until the rms gradient of the potential energy was less than 0.1 kcal mol^{-1} Å^{-1}. Protein stereochemistry evaluation was performed by several tools (Ramachandran and Chi plots measure phi/psi and chi1/chi2 angles, clash contacts reports) implemented in MOE suite [63].

2.3 Results and Discussion

The availability and the selection of a suitable template structure is a critical step in the homology modeling process. The structural information available for the GPCR family are limited, even if the number of GPCR crystal structure published on the PDB increased in past few years.

GPCRs are formed by a single polypeptide chain that crosses the cell membrane seven times with seven α-helical transmembrane domains (7TMs) bundled together in a very similar manner. Supporting the idea of a common folding of the seven TMs, sequence comparison revealed specific amino acid patterns characteristic of each TM and highly conserved in the great majority of Class A GPCRs. These conserved residues constitute the basis for the identification of the seven TMs within GPCR amino acid sequences. They are also the foundation of the GPCR residue indexing system introduced by Ballesteros and Weinstein. [25]

Bovine rhodopsin provided the first high resolution structural information, and for many years, rhodopsin-based homology modeling had been the most widely used approach to obtain three dimensional models of GPCRs. The results of AR modeling based on rhodopsin has been extensively reviewed. [72] With the availability of new crystallographic structures it is still questionable which one should be the more appropriate template for GPCRs modeling and, in particular, for ARs.

2.3.1 Sequence Alignment Analysis

The percentages of identity of the aligned sequences of the ARs in comparison to GPCRs having an available X-ray crystallographic structure are listed in table 2.1, and the alignment of the sequences is shown in figure 2.1. The percent identity increases from a comparison with bovine rhodopsin to a comparison with hGPCRs. The percent identity is higher if the N-terminus and the C-terminus are not taken into consideration, and the increase is even greater when comparing only TM regions.

Naturally, the $A_{2A}AR$ can be considered the best template for homology modeling of the other ARs according to the percent identity of the aligned sequences, but there are some important differences among the ARs that have to be considered in choosing the template for homology modeling. The

primary structures of A_1AR, $A_{2B}AR$, and A_3AR have a similar number of amino acid and, in general, these AR subtypes are among the smaller members of the GPCR family. For example, the human homologs of the A_1AR, $A_{2B}AR$, and A_3AR consist of 326, 328, and 318 amino acid residues, respectively. [70,73,74] In contrast, the $hA_{2A}AR$ consists of 409 amino acids, [75] and all cloned species homologs of the $A_{2A}AR$ are of similar mass. This relatively large size is manifested in the carboxyl-terminal tail of the receptor, which is much longer than any of the other AR subtypes.

Table 2.1: Percentages of identity of the aligned sequences of ARs and the crystallographic structures available for GRCRs.

		b-rhodopsin	$h\beta_2AR$	Turkey β_1AR	$hA_{2A}AR$
All	hA_1AR	13.8	19.1	17.2	39.1
	$hA_{2A}AR$	17.8	23.5	22.6	100
	$hA_{2B}AR$	17.8	22.5	20.1	46.6
	hA_3AR	14.1	19.9	17.4	31.3
All except	hA_1AR	15.6	25.6	24.9	50.8
	$hA_{2A}AR$	20.5	27.9	28.3	100
N-term and C-term	$hA_{2B}AR$	22.2	27.9	28.7	61.5
	hA_3AR	15.6	25.6	24.6	41.9
TM regions	hA_1AR	17.7	29.5	31.4	57.7
	$hA_{2A}AR$	22.3	31.8	33.2	100
	$hA_{2B}AR$	22.7	30.5	33.6	69.5
	hA_3AR	17.3	29.5	30.5	49.5
EL2	hA_1AR	14.3	14.8	11.1	32.4
	$hA_{2A}AR$	14.3	11.1	22.2	100
	$hA_{2B}AR$	14.3	18.5	22.2	41.2
	hA_3AR	14.3	11.1	11.1	23.5

The TM regions of the GPCRs possess the same overall topology, and the sequence alignment is guided by the most conserved residues in every helix. The size of each helix differs between the crystallographic structures, but the loops constitute the most variable region. The second extracellular loop (EL2) is of particular interest for building homology models of GPCRs used for drug design because of its role in the ligand recognition (Figure 1.7). The crystallographic structure of $hA_{2A}AR$ shows a disulfide bond between Cys259 and Cys262 in the intracellular side of the receptor and, in particular, three disulfide linkages that involve the EL2: one between Cys77 and Cys166, that is conserved among the members of family A of GPCRs and connects EL2 and TM3, and two between EL2 and EL1, that are unique to the $A_{2A}AR$ (Cys71-Cys159 and Cys74-Cys146). [45] The EL2 of the $A_{2A}AR$ defines the extracellular surface properties of the structure and is considerably different

Figure 2.1: Sequence alignment of hARs (A$_1$, A$_{2A}$, A$_{2B}$, A$_3$), bovine rhodopsin, hβ_2 adrenergic receptor and turkey β_1 adrenergic receptor. In grey are highlighted the transmembrane regions, in red the highly conserved residues and in yellow cysteines that form disulfide linkages that involve the second extracellular loop. For A$_1$, A$_{2B}$, A$_3$ARs only the cysteine residues that form the conserved disulfide bridge between TM3 and EL2 are highlighted in yellow, because information about other disulfide bonds are not available.

from that of rhodopsin. The extensive disulfide bond network forms a rigid, open structure exposing the ligand binding cavity to solvent, possibly allowing free access for small molecule ligands. [45]

The turkey β_1 adrenergic receptor and hβ_2 adrenergic receptor structures have the conserved disulfide bridge between EL2 and TM3 (Cys114-Cys189 for β_1AR and Cys106-Cys191 for β_2AR). In addition to this conserved structural constraint, they have a second disulfide bond that involves the EL2 (Cys192-Cys198 for β_1AR and Cys184-Cys190 for β_2AR). [38,43,48] However, rhodopsin has only one cysteine residue in the EL2, which forms a disulfide bond between EL2 and TM3. [28]

The sequences of the hA$_1$AR and the hA$_3$AR contain only one cysteine residue in the EL2 (Cys169 for A$_1$AR and Cys166 for A$_3$AR). These residues form the disulfide bridge, common to GPCRs, with the respective cysteine residues of TM3 (Cys80 for A$_1$AR and Cys83 for A$_3$AR). The hA$_{2B}$AR has three cysteine residues in the EL2. The cysteine in EL2 that forms the disulfide bridge with TM3 is conserved, as well as the cysteine residue within TM3, and the linkage between these residues is also conserved. No mutagenesis data are available for the other cysteines.

On A$_{2A}$AR there are other four cysteines that are connected by two disulfide bridges: Cys71-Cys159 and Cys74-Cys146. These residues correspond to Cys72, Thr162, Phe75 and Cys154 respectively on A$_{2B}$AR, if we consider the alignment that allows the higher percentage of identity. In this case no other disulfide bonds are formed, and only one cysteine of EL2 is involved in a disulfide linkage, i.e. the one with TM3 that is conserved among GPCRs. In addition, there are two more cysteine residues in EL2 (Cys166 and Cys167); depending on the alignment, one of these residues can be aligned with Cys159 of A$_{2A}$AR and form a second disulfide bond that connects EL2 with Cys72 of the A$_{2B}$AR. It remains to be clarified how many disulfide bonds are actually present in the structure of hA$_{2B}$AR. Nevertheless, the presence of three disulfide links on EL2 is a peculiarity of the hA$_{2A}$AR. This is an important point that has to be considered when the A$_{2A}$AR serves as the template for homology modeling of ARs to be used in drug design. The conformation of the A$_{2A}$AR binding pocket is influenced by EL2, which is strictly dependent on the presence of three disulfide linkages.

2.3.2   Homology Models of A$_3$ Adenosine Receptor

Different A$_3$AR models have been published describing the hypothetical interactions with known A$_3$AR ligands having different chemical scaffolds, and almost all of these models were constructed using bovine rhodopsin as a template. As we have widely discussed before, the new structures of GPCRs solved in the past two years provide a new starting point for homology modeling. In particular, the recent publication of A$_{2A}$AR provides important structural in-

formation for the AR family. Next to the structural information provided by the crystallographic data, mutagenesis studies can help identify the residues that are involved in ligand recognition. Site-directed mutagenesis of the A_3AR shows an important role for specific residues in TM3, TM6 and TM7. [76–81]

The three different models of hA_3AR can be constructed using as templates:

- the bovine rhodopsin (PDB ID 1F88);

- the $h\beta_2$-adrenergic receptor (PDB ID 2RH1);

- the $hA_{2A}AR$ (PDB ID 3EML).

The main differences between the templates are found within EL2, IL3 and the extracellular end of TM1. The structure-based drug design approach is mainly affected by differences in EL2, because residues of this loop can directly interact with ligands in the binding pocket. The EL2 of both squid and bovine rhodopsin assumes a β-sheet secondary structure, either in the structure with bound retinal or in the ligand-free structure. In the $h\beta_2AR$ there is an α-helix in EL2 that is structurally similar to the β_1AR of turkey, while the $A_{2A}AR$ does not have a defined secondary structure in the EL2.
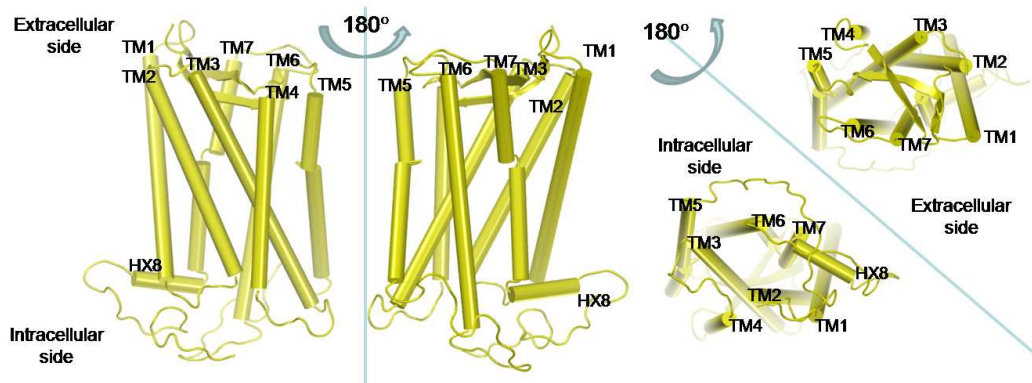
Figure 2.2: Topology of the hA_3AR built using bovine rhodopsin as template.

The first model of hA_3AR that we built was based on rhodopsin (Figure 2.2). As for the high-resolution structure of rhodopsin, the hA_3AR model reveals a seven-helical bundle with a central cavity surrounded by helices 3, 5, 6 and 7. Helix 4 is not part of the cavity wall and makes contacts only with helix 3. The access to the central cavity is not allowed because the EL2 closes the binding pocket and determines a volume of the cavity of 660 $Å^3$. EL2 is characterized by a β-sheet secondary structure and it is connected to TM3 with the conserved disulfide linkage between Cys83 and Cys166. This model has been widely used to identify putative ligand-receptor interactions and to

Figure 2.3: Topology of the hA$_3$AR built using β_2-Adrenergic Receptor as template.

understand and quantify the structure activity relationship (SAR) of known hA$_3$AR antagonists through a high-throughput docking strategy. [82–87]

Two other models of the hA$_3$AR were built using as a template the hβ_2-adrenergic receptor and the turkey β_1-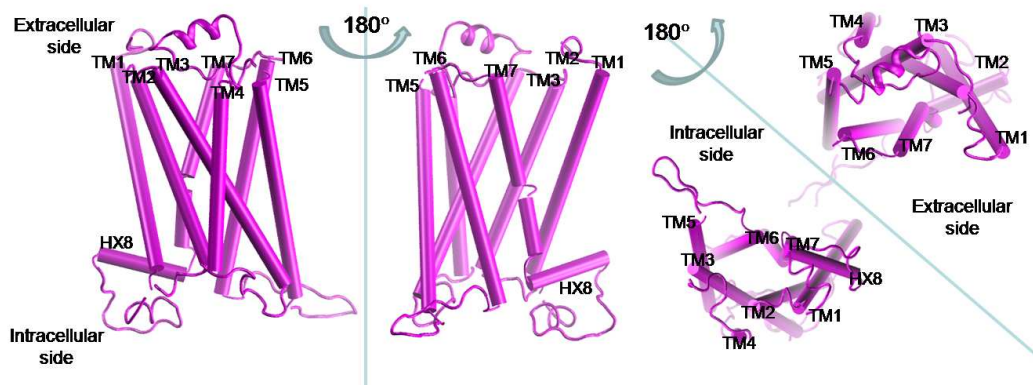adrenergic receptor. The RMSD of the entire structures superposed is around 4 Å, it is 2.8 Å without considering the N-terminus (from residue 1 to 8), C-terminus (from residue 302 to 318), and IL3 (from residue 208 to 224), which are the most variable regions. The RMSD is only 1.8 Å considering only the helical backbone. These models do not present relevant differences at the active-site level, and therefore we are considering only the one built using β_2-adrenergic receptor as template (Figure 2.3).

Even though one of the two disulfide bridges in the EL2 is missing, the conformation of the EL2 of the hA$_3$AR model is similar to the EL2 of the adrenergic receptor template: an α-helix secondary structure enables the accessibility to the ligand-binding site. In the template, this conformation may be stabilized by an intra-loop disulfide bond, which is missing in the model of hA$_3$AR. The putative location of ligands in the two templates is very similar. In preliminary docking studies, also the location of hA$_3$AR antagonist is similar, even if there are structural differences in the ligand binding sites between the models obtained from rhodopsin and the adrenergic receptor. The largest difference within the TM region between the two models occurs in helix 1, in which the adrenergic receptor-based model lacks the proline-kink found in rhodopsin-based model.

The recently published structure of hA$_{2A}$AR provides a new template for GPCR modeling and in particular for ARs. A new model of the hA$_3$AR was
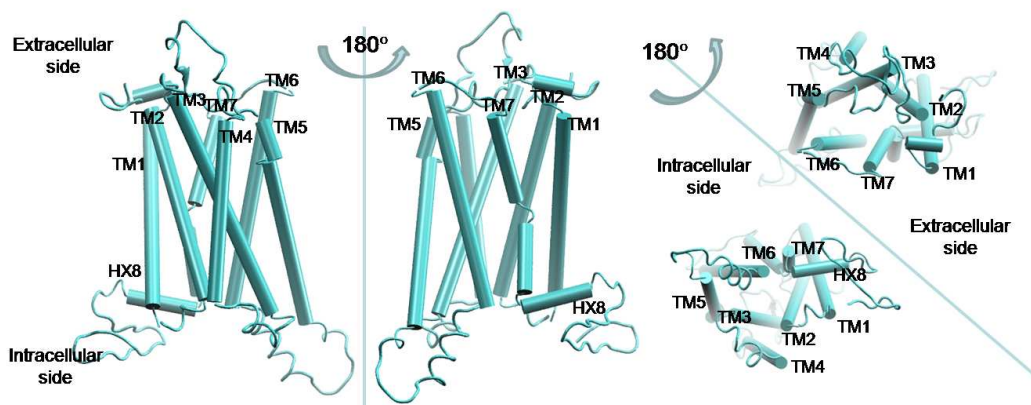
Figure 2.4: Topology of the hA$_3$AR built using A$_{2A}$AR as template.

built using this crystal structure as template (Figure 2.4). The helical arrangement is similar among the models. However, the helices are shifted, and the differences among their relative positions result in an RMSD around 2.50 Å. As observed for the model built using adrenergic receptors as templates, the main difference in the helical bundle is TM1 and in particular the N-terminal end of the helix. A detailed comparison of the superimposed models is in figure 2.5 and in table 2.2, in which values of RSMD for each TM helix are reported.

As it was seen for the templates, the main difference among the three models of the hA$_3$AR is in the loop region. The ligand binding pocket of the crystal structure of A$_{2A}$AR is shifted closer to TM6 and TM7, and the position of the A$_{2A}$AR antagonist ZM241385 is closer to these helices. Important

Table 2.2: Root mean square deviation (RMSD) of the backbone of the aligned models of hA$_3$AR. The main difference among the models is due to the loops, which represent the most variable region of the templates and consequently of the models. Particular attention has to be done to EL2 because it is part of the binding pocket and it can directly interact with ligands.

	all TM	all loops	TM1	TM2	TM3	TM4	TM5	TM6	TM7	HX8	EL2
RMSD in Å with respect to hA$_3$AR model from bovine rhodopsin (backbone)											
A$_3$-β_2	2.29	10.86	2.82	2.12	1.98	2.01	2.07	2.19	1.85	3.73	11.44
A$_3$-A$_{2A}$	2.43	10.06	2.55	2.40	2.78	2.45	2.85	2.02	2.04	1.64	14.30
RMSD in Å with respect to hA$_3$AR model from hβ_2-Adrenergic Receptor (backbone)											
A$_3$-rho	2.29	10.86	2.82	2.12	1.98	2.01	2.07	2.19	1.85	3.73	11.44
A$_3$-A$_{2A}$	2.57	7.46	3.84	1.89	2.02	1.73	2.09	2.71	2.23	3.66	6.18
RMSD in Å with respect to hA$_3$AR model from hA$_{2A}$AR (backbone)											
A$_3$-rho	2.43	10.06	2.55	2.40	2.78	2.45	2.85	2.02	2.04	1.64	14.30
A$_3$-β_2	2.57	7.46	3.84	1.89	2.02	1.73	2.09	2.71	2.23	3.66	6.18
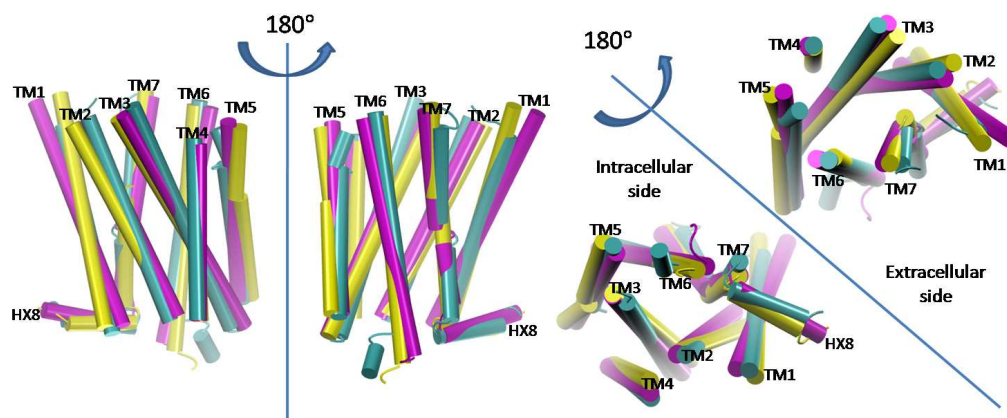
Figure 2.5: Topology of the superposed hA$_3$AR models. A$_3$AR from rhodopsin is in yellow, A$_3$AR from hβ_2-AR is in magenta and A$_3$AR from hA$_{2A}$AR is in cyan.
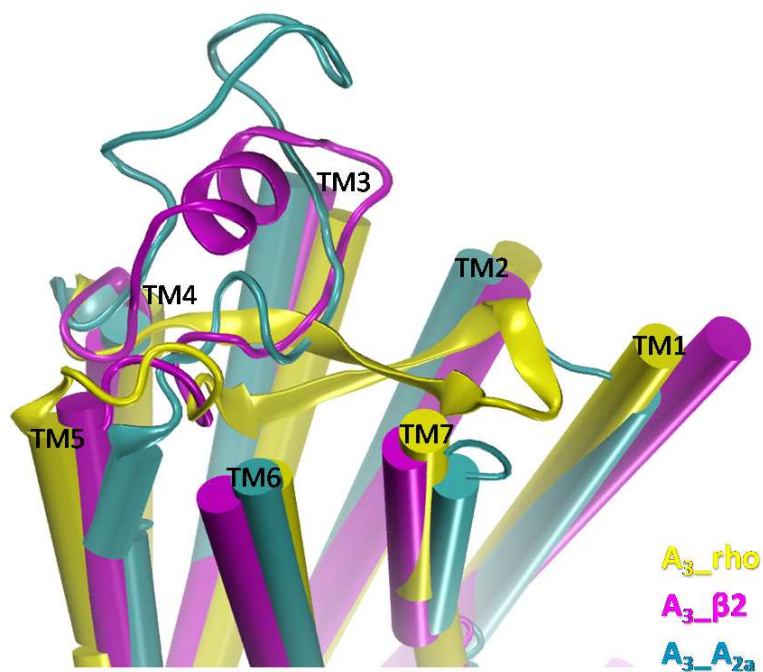

Figure 2.6: Representation of EL2 of A$_3$AR models: in yellow hA$_3$AR built from rhodopsin, in magenta hA$_3$AR built from β_2-AR and in cyan hA$_3$AR built from A$_{2A}$AR.

interactions are also established with EL2. The position of ZM241385 is significantly different from the one of retinal or carazolol. Even though GPCRs share a common topology, ligands may bind in a different fashion and interact with different positions of the receptor. The model built starting from the $A_{2A}AR$ template is different from the previous models of A_3AR: the binding pocket is closer to TM6 and TM7 and open to the extracellular side. The volume of the binding sites of A_3AR models built starting from $h\beta_2$-AR and $hA_{2A}AR$ is difficult to be measured because they present a binding site open to the extracellular side. The volumes were estimated as 1620 Å^3 and 1930 Å^3, respectively, but they cannot be compared with the volume of the binding site of the rhodopsin-based model, which is closed and has a volume of 660 Å3 (Figure 2.7).
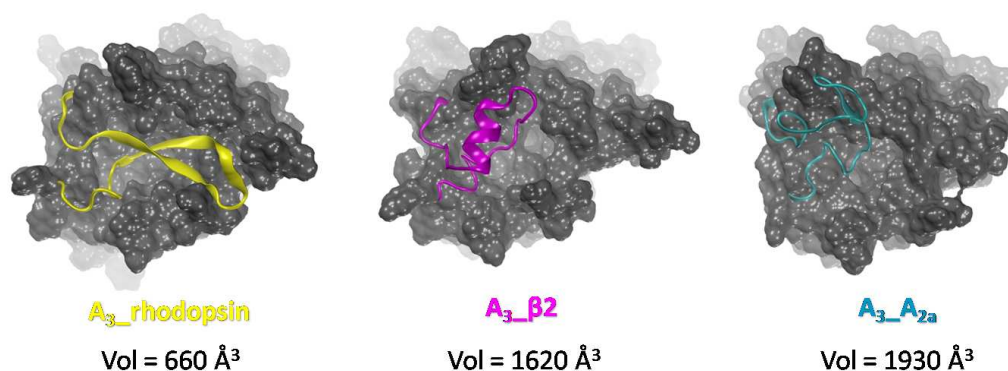

Figure 2.7: Extracellular side view of the hA$_3$AR models. A$_3$AR from rhodopsin is in yellow, A$_3$AR from hβ_2-AR is in magenta and A$_3$AR from hA$_{2A}$AR is in cyan.

Even if the percentage of identity of the hA$_3$AR is higher with respect to the $A_{2A}AR$ than with the previously reported structures, the conformation of the EL2 and consequently of the binding pocket of the hA$_3$AR might be different from the $A_{2A}AR$. The peculiarity of the $A_{2A}AR$ is the presence of three disulfide bridges on EL2, which are not conserved among ARs. Also, the particular conformation of EL2 and the binding pocket can be particular to this subtype, and use of the $A_{2A}AR$ as a template for modeling other AR subtypes is still imprecise. Also, mutagenesis data support the hypothesis of different roles of TM helices in different AR subtypes.

2.3.3 Ligand-Based Homology Modeling

We have revisited the rhodopsin-based model of the human A$_3$ receptor in its resting state (antagonist-like state), taking into account a novel strategy to simulate the possible receptor reorganization induce by the antagonist-binding. We called this new strategy ligand-based homology modeling and its
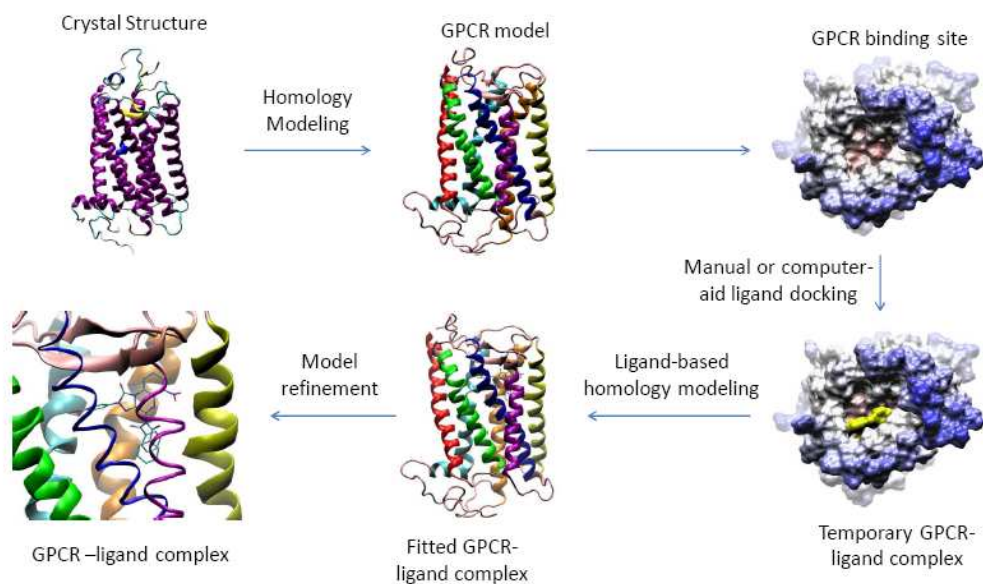
Figure 2.8: Flow chart of the ligand-based homology modeling technique considering an evolution of a conventional homology modeling algorithm implemented by Molecular Operating Environment modeling software.

schematic flow chart is summarized in Fig. 2.8. [88]

These specific homology modeling approach have been implemented into Molecular Operating Environment (MOE) software. [63] Succinctly, ligand-based homology modeling technique is an evolution of a conventional homology modeling algorithm based on a Boltzmann weighted randomized modeling procedure adapted from Levitt, combined with specialized logic for the proper handling of insertions and deletions; any selected atoms will be included in energy tests and in minimization stages of the modeling procedure. Ligand-based option is very useful when one wishes to build a homology model in the presence of a ligand docked to the primary template, or other proteins known to be complexed with the sequence to be modeled.

In this specific case both model building and refinement take into account the presence of the ligand in terms of specific steric and chemical features. To strictly validate this methodology as novel tool to address the multi-conformational space of GPCRs, we have analyzed many known human A_3 antagonists in the corresponding putative ligand binding site, which are reported in the Chapter 3. [82–87]

The analized compounds present a different chemical structures, with different molecular shape and volume. More detailed description of the models obtained with different classes of antagonists is in Chapter 3.

In general, considering the ligand recognition cavity of the receptor built from a rhodopsin-based model, we have estimated that its specific volume is

around 660 $Å^3$. However, even if this conventional rhodopsin based model of the human A_3 receptor is able to elucidate the observed activity of all derivatives bearing small substituents, the same model could not explain the observed activity when bulkier substituents are present. Independently from the used molecular docking algorithm, a strongly destabilizing van der Waals energy component avoided to sample reasonable antagonist-receptor complexes.

We interpret this fact as a clear indication that the rhodopsin based receptor cavity is not appropriated to guarantee a good complementarity among the topology of the receptor's cleft and the shape of these antagonists.

Starting from the conventional rhodopsin-based homology model and applying our ligand-based homology modeling implementation we have generated other antagonist-like conformational states of human A_3 receptor in which the ligand recognition cavity has been expanded. Using the new antagonist-like conformational states, we were able to rationalize the observed activities for all reported compounds. Many severe analysis concerning false-positive and false-negative situations have been conducted. For example, the less bulky compound that nicely fits into the conventional rhodopsin-based model, drastically reduces its interaction energy when it is docked into the other ligand-based models. Indeed, increasing of the TM cavity volume reduce both steric and chemical complementarities between ligand and receptor. Using this multi-conformational states approach, a consensus binding motif among all known antagonists has been found, and a novel "Y-shaped" 3D-pharmacophore model has been proposed. [89]

Molecular Docking of A$_3$ Adenosine Receptor Antagonists

3.1 Introduction

The GPCR models are theoretical structures whose reliability has to be checked. In order to evaluate the goodness of a GPCR model, "indirect" methods should be taken into consideration: some of these concern the computational procedure, others the accordance with the available experimental data (mainly mutagenesis and ligand activity), and finally the predictive ability of the model. A "structural" validation can be carried ahead through the inspection of experimental data: residues that mutagenesis studies had revealed to play a significant role, should be found involved in important ligand-receptor or inter-helices interactions in the GPCR model. A "functional" validation is the ability of the models to predict the activity of known ligands, to suggest the design of new ones or to suggest the mutation of residues that the model suggests as important for the ligand interaction or in the maintenance of the receptor folding.

Our theoretical model of hA$_3$AR based on rhodopsin has been used to evaluate and quantify the structure-activity relationship of new synthesized ligands, analyzing their interactions inside the binding sites and correlating them with their affinity and selectivity.

Later, the model has been used also with the purpose of synthesizing new ligands rationally designed on the basis of information obtained from the structure activity relationship analysis.

3.2 Materials and Methods

All the docking studies reported in this chapter were performed using the Molecular Operating Environment (MOE, version 2007.09) suite. [63]

3.2.1 Preparation of the Ligands

All docked structures were fully optimized without geometry constraints using RHF/AM1 semiempirical calculations. Vibrational frequency analysis was

used to characterize the minima stationary points (zero imaginary frequencies). The software package MOPAC (ver.7), [90] implemented in MOE suite, was utilized for all quantum mechanical calculations.

3.2.2   Model of Human A_3 Adenosine Receptor

The model that has been used for docking studies is the rhodopsin based model that was widely described in the chapter 2.

When this project started, only five crystal structures of Bovine Rhodopsin were available. These structural information were the starting point of homology modeling of human A_3 Adenosine Receptor and the crystal structure 1F88 [28] was used as template to built the first homology model of hA_3AR. Rhodopsin-based homology modeling has represented for many years a widely used and well-consolidated approach to create GPCR three dimensional models.

This model was used to describe structure activity relationship of more than 300 known human A_3 antagonists in the corresponding putative ligand binding site.

Moreover, our recently described ligand-based homology modeling (LBHM) approach has been used to simulate the conformational changes induced by ligand binding. [88] With LBHM technique it is possible to create different conformational states of the same receptor preserving the general rhodopsin based topology.

3.2.3   Docking Procedure

All antagonist structures were docked into the hypothetical TM binding site of the model of hA_3AR built using bovine rhodopsin as template by using the MOE-dock tool, part of the MOE suite. Searching is conducted within a user-specified 3D docking box, using the Tabu Search protocol [91] and the MMFF94 force field. [92] MOE-Dock performs a user-specified number of independent docking runs (50 in our specific case) and writes the resulting conformations and their energies in a molecular database file. The resulting docked complexes were subjected to MMFF94 energy minimization until the rms of conjugate gradient was <0.1 kcal mol^{-1} Å^{-1}. Charges for the ligands were imported from the MOPAC output files. To better refine all antagonist-receptor complexes, a rotamer exploration of all side-chain involved in the antagonist-binding was carried out. Rotamer exploration methodology is implemented in MOE suite.

3.3 Results and Discussion

3.3.1 4-Amido-2-aryl-1,2,4-triazolo[4,3-*a*]quinoxalin-1-one Derivatives

We used our improved model of the hA$_3$ receptor, obtained by a rhodopsin-based homology modeling approach to recognize the hypothetical binding motif of these newly synthesized 4-amino-2-phenyl-1,2,4-triazolo[4,3-*a*]quinoxalin-1-one antagonists. [82] All the docked compounds are listed in the Appendix A and are reported in figure 3.1.

1-5 R$_4$= Me
6-11 R$_4$= Ph
12-18 R$_4$= CHPh$_2$

19-23

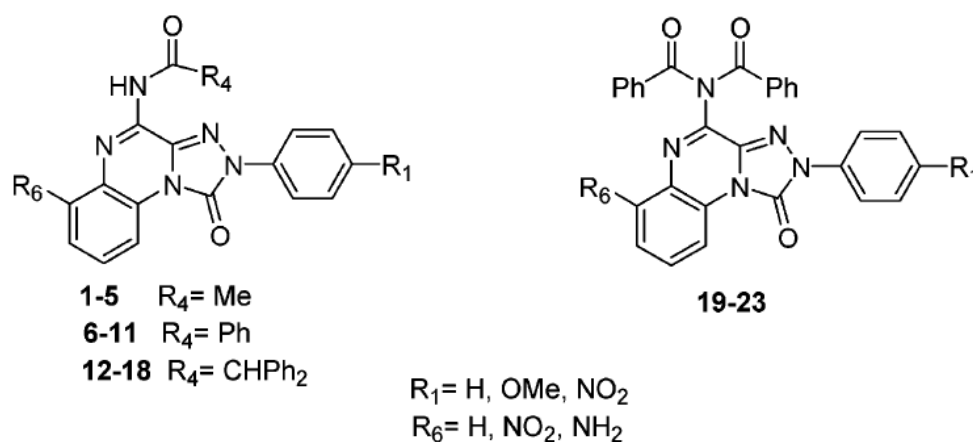R$_1$= H, OMe, NO$_2$
R$_6$= H, NO$_2$, NH$_2$

Figure 3.1: Reported 4-amido-2-aryl-1,2,4-triazolo[4,3-*a*]quinoxalin-1-one derivatives.

From analysis of docking simulation results, all triazoloquinoxalinone derivatives share a similar binding motif inside the transmembrane region of the hA$_3$ receptor, as previously described. [93] As shown in figure 3.2, we identified the hypothetical binding site of the triazoloquinoxalinone moiety surrounded by TMs 3, 5, 6, and 7 with the carbonyl group at 1-position pointing toward the EL2 and with the amide moiety in the 4-position oriented toward the intracellular environment. The phenyl ring at the 2-position is close to TMs 3, 6, and 7, whereas R$_6$ substituents are close to TM5. For a clear explanation of the observed structure-activity relationships, it is useful to immediately emphasize that the relative positions of the R$_6$ substituents are slightly different depending on the bulkiness of the R$_4$ substituent on the 4-amide moiety, as shown in figure 3.3.

However, the overall pharmacophore features are nicely consistent with our recently proposed receptor-based pharmacophore model [89,94,95].

From analysis of our model in detail, all triazoloquinoxalinone derivatives share at least two stabilizing hydrogen-bonding interactions inside the binding cleft as shown in figure 3.4.
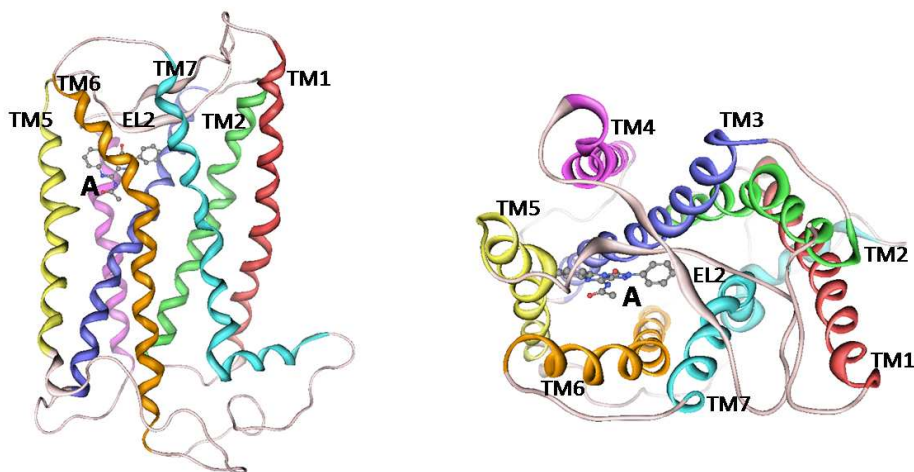
Figure 3.2: Human A_3 receptor model viewed from the membrane side (on the left) and from the extracellular side (on the right) showing the EL2 folded into the binding crevice. Compound **A** is the binding pocket according to his hypothetical binding pose.

The first hydrogen bonding is between the carbonyl group at the 1-position, pointing toward the EL2, and the NH of the Gln167-Phe168 amidic bond. This hydrogen-bonding distance is calculated to be around 2.8 Å for all docked compounds. Moreover, the 1-carbonyl group is also at the hydrogen-bonding distance with the amide moiety of Asn250 (6.55) side chain. This asparagine residue, conserved among all adenosine receptor subtypes, was found to be important for ligand binding. Second, the NH-CO moiety at the 4-position is surrounded by three polar amino acids: Thr94 (3.36), His95 (3.37), and Ser247 (6.52). This region seems to be very critical for the recognition of all antagonist structures. In fact, a major structural difference between the hypothetical binding sites in adenosine receptor subtypes is that the A_3 receptor does not contain the histidine residue (6.52) in TM6 common to all A_1 (His251 in hA_1) and A_2 (His250 in hA_{2A}) receptors. This histidine has been shown to participate in both agonist and antagonist binding to A_{2A} receptors. In the A_3 receptor this histidine in TM6 is replaced by a serine residue (Ser247 in hA_3). [96] The stabilizing interactions among the 4-carbamoyl moiety and these polar amino acids orient the adjacent R_4 substituent (methyl, **A** and **1-5**; phenyl, **B** and **6-11**; diphenylmethyl, **12-18**) in the middle of the TM bundle. In particular, the O-H of Ser247 (6.52) and the carbonyl oxygen of the amide group are separated by 2.4 Å and appropriately oriented to form a H-bonding interaction. Moreover, the side chain of His95 (3.37) is within dipole-dipole interaction distance of NH of the amide group, at around 2.9 Å. According to recently published mutagenesis results, both His95 and Ser247 seem to affect the binding of both agonists and antagonists. [96] Indeed, the receptor

region around R_4 substituents is mostly hydrophobic and characterized by five nonpolar amino acids: Ile98 (3.40), Ile186 (5.47), Leu190 (5.51), Phe239 (6.44), and Leu244 (6.49).

The effects of substituents in R_1-position and R_6-position are shown in figures 3.5 and 3.6.

Considering the observed structure-activity relationships in greater detail, methoxy substitution at the R_1-position is rather well tolerated among all newly synthesized triazoloquinoxalinone derivatives. This is consistent with its accommodation into a tiny hydrophobic pocket delimited by Leu90 (3.32) and Ile268 (7.39). Interestingly, the amino acid corresponding to Leu90 in the hA_{2A} receptor was found to be essential for the binding of both agonists and antagonists, and it is mutated in valine (Val87) in the human A_1 receptor. This mutation might play a role in the explanation of hA_3 versus hA_1 selectivity. In fact, even if the mutation Leu90 (hA_3)/Val87 (hA_1) can slightly enlarge the dimension of this hydrophobic cavity, at the same time it also notably decreases the shape and hydrophobic interaction complementarity (data
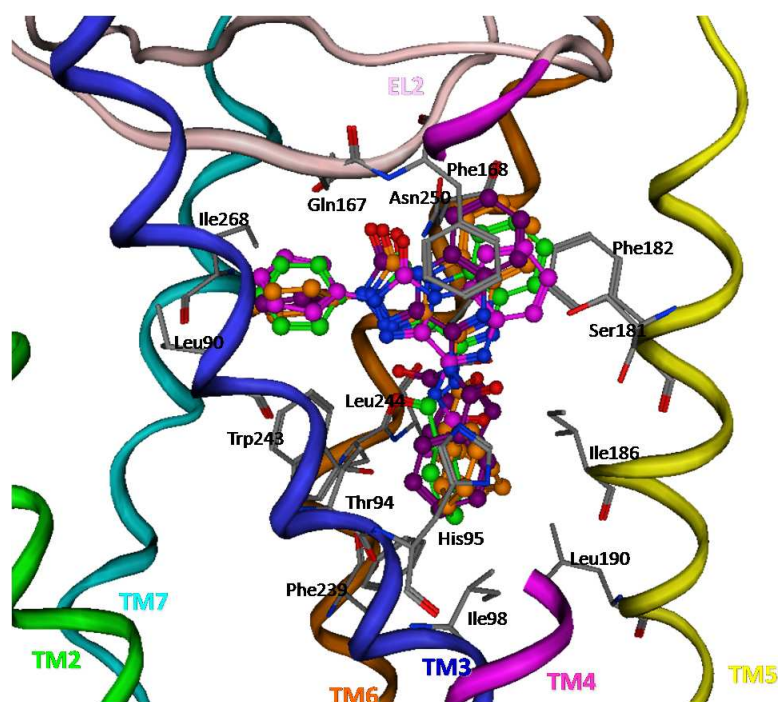
Figure 3.3: Hypothetical binding motif of the representative 4-amino-2-phenyl-1,2,4-triazolo[4,3-a]quinoxalin-1-one antagonists (derivative **A** in magenta, derivative **B** in green, derivative **14** in orange and derivative **19** in violet). All docked antagonists are viewed from the membrane side facing TM helices 3 and 4. To clarify the TM cavity, the view of TM4 from Leu136 to Pro145 has been voluntarily omitted. Side chains of some amino acids important for ligand recognition are highlighted. Hydrogen atoms are not displayed.

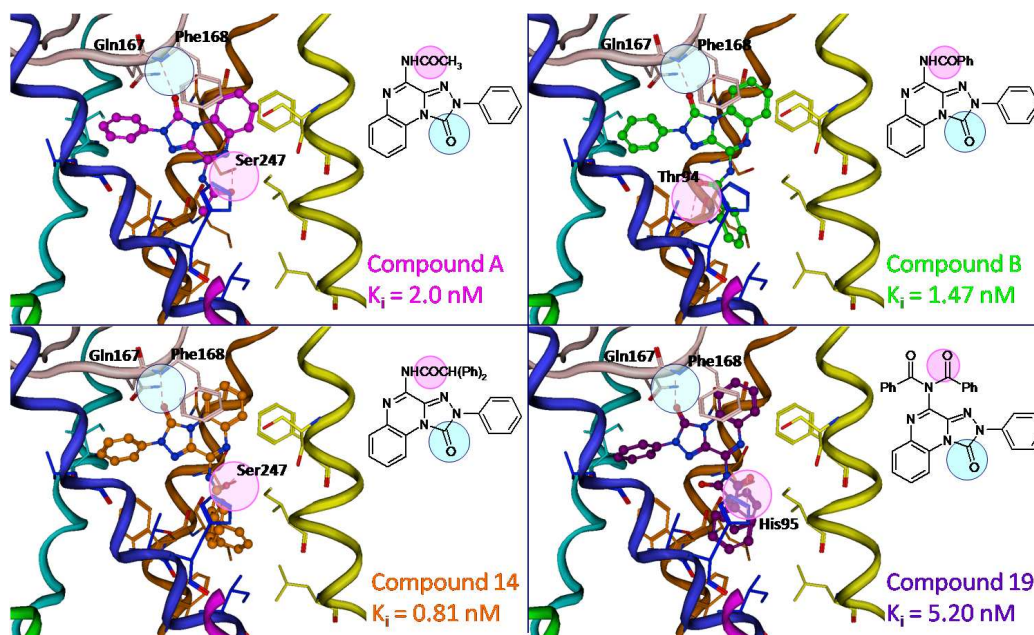Figure 3.4: Representative triazolo-quinoxalin-1-ones derivatives: two stabilizing hydrogen-bonding interactions inside the binding cleft that are conserved among all the derivatives.

not shown). Also, the mutation of Ser165 (EL2 of hA_3) with Lys168 in the hA_1 receptor could affect the recognition of the methoxy-substituted triazolo-quinoxalinone derivatives. Considering the same small pocket surrounded by Leu90 (3.32) and Ile268 (7.39), unfavorable steric and dipolar interactions are responsible for the reduction of affinity observed for derivatives **7** and **13**, whereas the methoxy substituent at R_1 is replaced by the nitro group.

On the other hand, the presence of the 6-nitro substituent does not always produce advantageous effects in terms of hA_3AR binding affinity. This phenomenon is particularly evident when derivatives **2** and **15** are compared with their unsubstituted compounds **A** and **14**. As already anticipated and clearly shown in figure 3.3, the relative positions of R_6 substituents are slightly different depending on the bulkiness of the R_4 substituent on the carbamoyl moiety at the 4-position. In particular, in the presence of a less bulky R_4 substituent such as a methyl group (derivative **A**), the triazoloquinoxalinone moiety binds more deeply in the middle of the TM bundle, positioning the 6-nitro substituent very close to TM5 (Figure 3.5). In this case, unfavorable steric and dipolar interactions are responsible for the remarkable reduction of affinity observed for derivatives **2** and **3**. In contrast, the smaller 6-amino substituent (derivatives **4** and **5**) is still well tolerated because of the favorable dipolar interaction with the carbonyl moiety of the Ser181-Phe182 amidic bond. When the bulkiness of the R_4 substituent is increased, the position of
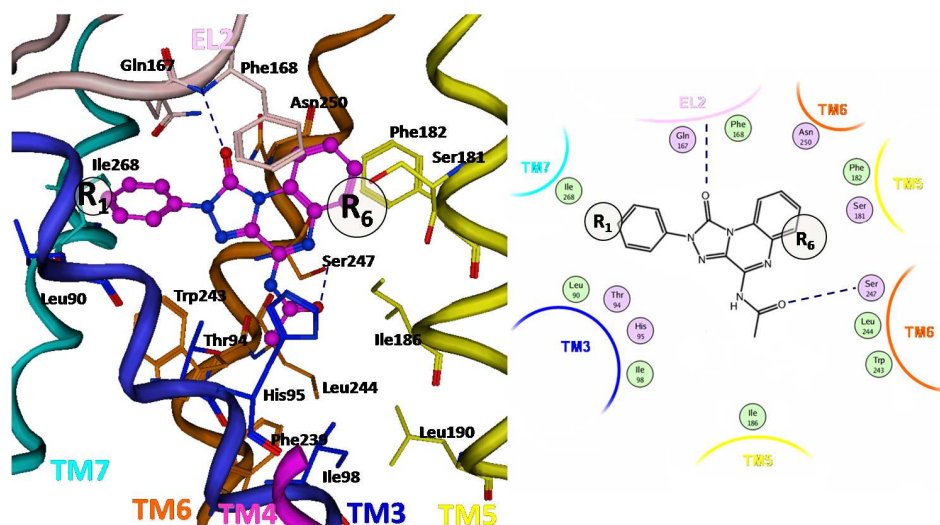
Figure 3.5: Compound **A** of triazolo-quinoxalin-1-ones derivatives in the binding pocket of hA$_3$AR. On the left: the antagonist is viewed from the membrane side facing TM helices 3 and 4. The positions of R$_1$ and R$_6$ are highlighted by two black circles. To clarify the TM cavity, the view of TM4 from Leu136 to Pro145 has been voluntarily omitted. Side chains of some amino acids important for ligand recognition are highlighted. Hydrogen atoms are not displayed. On the right: 2D scheme of the interactions.

Figure 3.6: Compound **14** of triazolo-quinoxalin-1-ones derivatives in the binding pocket of hA$_3$AR. On the left: the antagonist is viewed from the membrane side facing TM helices 3 and 4. The positions of R$_1$ and R$_6$ are highlighted by two black circles. To clarify the TM cavity, the view of TM4 from Leu136 to Pro145 has been voluntarily omitted. Side chains of some amino acids important for ligand recognition are highlighted. Hydrogen atoms are not displayed. On the right: 2D scheme of the interactions.
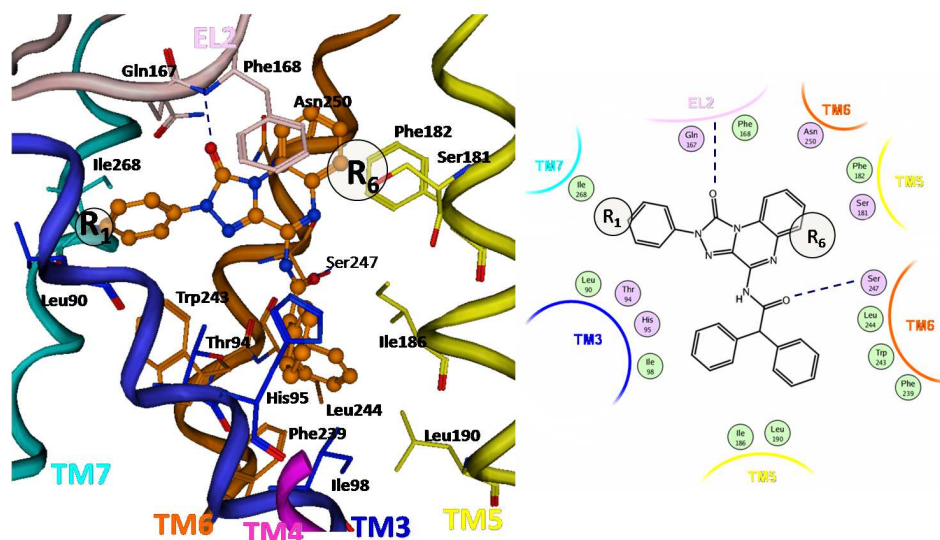
the R_6 group shifts away from TM5, and consequently, more empty space is available for the 6-nitro substituent, such as in derivatives **8**, **15**, and **16**.

In figure 3.6 is shown compound **14**, that does not present substituents in R_6, but the position of the scaffold is the same for the derivateves **12-18** with an R_6 group and a bulky substituent in R_4 position.
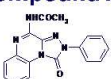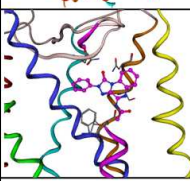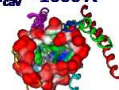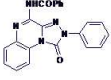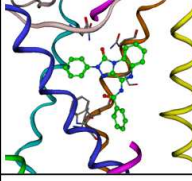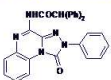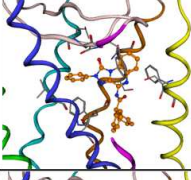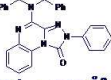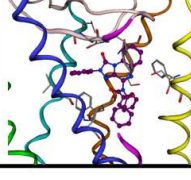

Reference derivatives	Docked derivatives	Rhodopsine based Homology model $Vol_{cav} = 660$ Å^3	Homology model 2 (using compound B as docked structure) $Vol_{cav} = 850$ Å^3	Homology model 3 (using compound C as docked structure) $Vol_{cav} = 1000$ Å^3
Compound A NHCOCH$_3$ $Vol_{eg} = 210$ Å^3 $K_i = 2.0$ nM	**Compounds A, 1 - 5**			
Compound B NHCOPh $Vol_{eg} = 501$ Å^3 $K_i = 1.47$ nM	**Compounds B, 6 - 11**	NO REASONABLE DOCKING POSES		
Compound 14 NHCOCH(Ph)$_2$ $Vol_{eg} = 641$ Å^3 $K_i = 0.81$ nM	**Compounds 12 - 18**	NO REASONABLE DOCKING POSES	NO REASONABLE DOCKING POSES	
Compound 19 $Vol_{eg} = 643$ Å^3 $K_i = 5.20$ nM	**Compounds 19 - 23**	NO REASONABLE DOCKING POSES	NO REASONABLE DOCKING POSES	

Figure 3.7: Ligand-based homology modeling (LBHM) data collection of triazolo-quinoxalin-1-ones derivatives. The reference derivatives **B** and **14** were used as ligand templates during the homology modeling process to built two new conformational states of A_3 model. Consequently, three different conformational states (rhodopsin based model and models 2 and 3) were selected as putative ambassadors of the conformational changes induced by different ligand binding. Depending on their different structure topologies, all other antagonists (docked derivatives) were docked into the most complementary receptor model.

Considering the 4-dibenzoyl derivatives **19-23**, the simultaneous presence of two bulky substituents at the 4-position forces a slight rearrangement of the triazoloquinoxalinone moiety inside the TM binding cavity (Figure 3.3). Curiously, while the position of the methoxy substitution at the R_1-position is relatively well conserved compared with all other triazoloquinoxalinone derivatives, the R_6 substituents are much closer to the R_6 position of derivative **2**

and consequently much closer to the TM5 domain. As already described for compound **2**, in this case the unfavorable steric and dipolar interactions are probably responsible for the remarkable reduction of affinity of derivatives **21** and **22**. To explain the different behavior of derivatives **21** (R_6 = NO2; I = 27% at 1 M) and **23** (R_6 = NH_2; K_i 1200 nM), we can apply the same argument already used for the comparison of derivatives **2** and **4**.

Starting from the rhodopsin based homology model of AR and applying the LBHM approach, we obtained 3 different conformational states of the hA_3 model. These conformational states preserve the conventional rhodopsin-like receptor topology and they were used in the SAR study of the reported triazoloquinoxalinone derivatives. The results are summarized in figure 3.7.

3.3.2 2-Arylpyrazolo[3,4-*c*]quinoline Derivatives

Molecular modeling studies were performed on the pyrazoloquinoline derivatives **1-36** (reported in Appendix B and in figure 3.8) in order to identify the hypothetical binding motif of this class of 2-arylpyrazolo[3,4-c]quinoline derivatives and rationalize the observed SAR. [83]

1-4 R= H, 3-Me, 4-Me, 4-OMe
5-6 R= 3-OMe, 4-NO_2

7

8-11 R= H, 3-Me, 4-Me, 4-OMe
12 R= 3-OMe

13-28 R_4= Me, Ph, CH_2Ph
CHPh$_2$
29-32 R_4= NHCH$_2$Ph

R= H, 3-Me, 4-Me, 4-OMe

33-36
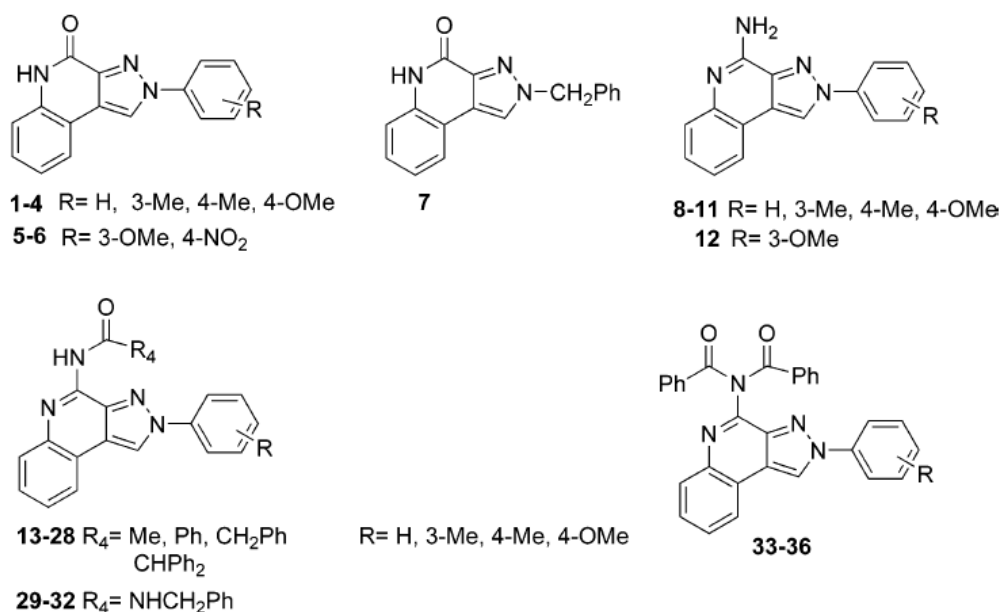
Figure 3.8: Reported 2-Arylpyrazolo[3,4-*c*]quinoline Derivatives

The main issues to be addressed were:

- to clarify the different role of the R substituent on hA_3 affinity and selectivity of the 4-oxo/4-amino compounds **1-12** and 4-acylamino/4-benzylureido derivatives **13-36**;

- to interpret the advantageous effect of the 4-acylamino moieties both for hA$_3$ affinity and selectivity.

Following our previously reported modeling studies, [82,93,97,98] we have constructed a refined model of the hA$_3$ receptor by using a rhodopsin-based homology modeling (RBHM) approach. [94,95,99,100] Moreover, our recently described ligand-based homology modeling (LBHM) approach has been used to simulate the conformational changes induced by ligand binding. [88]

Figure 3.9: Ligand-based homology modeling (LBHM) data collection. Each "reference derivative" (compounds **8**, **21**, **25**, and **29**) was used as ligand template during the homology modeling process. Consequently, four different conformational states (models 1-4) were selected as putative ambassadors of the conformational changes induced by different ligand binding. Depending on their different structure topologies, all other antagonists (docked derivatives) were docked into the most complementary receptor model.

As reported in figure 3.9, depending on the topological properties of the different ligands, we found four different conformational models of the human A$_3$ receptor reverse agonist-like state in which both shape and chemical complementarities have been specifically optimized around each ligand. In this specific case, with the varying of ligand structure, the molecular volume of the transmembrane (TM) binding cavity changes from the 660 Å^3 of the standard RBHM-driven model to the 1120 Å^3 of the largest LBHM-driven model, without altering the conventional rhodopsin-like receptor topology. The modifications of both shape and volume of the human A$_3$ TM binding cavity are the

Figure 3.10: Compound **17** of arylpyrazolo-quinoline derivatives in the binding pocket of hA₃AR. On the left: the antagonist is viewed from the membrane side facing TM helices 3 and 4. To clarify the TM cavity, the view of TM4 from Leu136 to Pro145 has been voluntarily omitted. Side chains of some amino acids important for ligand recognition are highlighted. Hydrogen atoms are not displayed. On the right: 2D scheme of the interactions.

most important receptor modeling perturbations obtained by the application of the LBHM technique. The binding cavity reorganization induced by ligand binding is due to the conformational change in several amino acid side chains, such as Leu90 (3.32), Leu91 (3.33), Thr94 (3.36), His95 (3.37), Ile98 (3.40), Gln167 (EL2), Phe168 (EL2), Phe182 (5.43), Ile186 (5.47), Leu190 (5.51), Phe239 (6.44), Trp243 (6.48), Leu244 (6.49), Leu264 (7.35), and Ile268 (7.39). However, molecular docking studies carried out for all the pyrazoloquinoline antagonists, using the appropriate conformational states of the receptor as listed in figure 3.9, have shown a similar binding motif, indicating that a common receptor-driven pharmacophore model can be depicted. This finding is consistent with our previously reported studies. [82,88,93–95,97–100]

Interestingly, none of the new pyrazoloquinoline antagonists found an energetically stable docking pose in the conventional RBHM-driven A₃ model. This is mainly due to the unfavorable topological complementarity among these antagonists and corresponding RBHM-driven TM binding cavity. In particular, highly destabilizing van der Waals interactions (steric conflicts) seem to be the reason for a lack of topological complementarities. These steric conflicts are drastically reduced or completely eliminated after application of the LBHM approach.

The ligand recognition occurs in the upper region of the TM bundle, and the pyrazoloquinoline moiety is surrounded by TMs 3, 5, 6, 7 with the substituent in the 4-position oriented toward the intracellular environment. As shown in figure 3.10, the phenyl ring at the 2-position is close to TMs 3, 6, and 7. Interestingly, an important hydrogen-bonding network can be observed in all energetically stable docked conformations of all pyrazoloquinoline

antagonists; in particular, Thr94 (3.36), His95 (3.37), and Ser247 (6.52) are able to interact through hydrogen bonding with the 4-carbonyl oxygen of compounds **1-7**, with the 4-amino group of compounds **8-12**, or with the 4-acylamino group of compounds **13-36**. These polar amino acids seem to be critical for the recognition of all antagonist structures and for receptor selectivity. In particular, Ser247 (6.52) of the hA_3 receptor subtype is not present in the corresponding position of A_1 and A_2 receptors, where the residue is replaced by a histidine (His251 in hA_1, His250 in hA_{2A}, and His251 in hA_{2B}). The histidine side chain is bulkier than serine and, possibly for this reason, large substituents at the 4-position of the pyrazoloquinoline framework are not well-tolerated by A_1 and A_2 receptor subtypes. Indeed, 4-acylamino and 4-benzylureido analogs (**13-36**) are inactive or modestly active on hA_1 and hA_{2A}ARs. On the contrary, the hydroxyl group of Ser247 (6.52) of the hA_3 receptor is appropriately positioned to form a hydrogen-bonding interaction with the carbonyl oxygen of the 4-amide/ureide group of compounds **13-36**. These observations support the importance of a 4-N-acyl/carbamoyl group in modulating receptor selectivity.

Specifically referring to 4-N-acylated derivatives, hA_3 receptor affinity increases with the bulkiness of the R_4 substituent (compare the 4-acetylamino compounds **13-16** with the 4-benzoyl compounds **17-20**). The hydrophobic environment of the five nonpolar amino acids, Ile98 (3.40), Ile186 (5.47), Leu190 (5.51), Phe239 (6.44), and Leu244 (6.49), can justify this affinity trend. Moreover, substituents bulkier than phenyl (compounds **21-36**) are also tolerated. In fact, compounds **21-36** maintain their hA_3 receptor affinities in the low nanomolar range ($K_i < 30$ nM). Both hydrogen-bonding interactions and shape/hydrophobicity complementarity of this region of the binding pocket are crucial for the anchoring of all compounds with a hydrophobic substituent at the R_4 position. Indeed, the introduction of a hydrophobic R substituent, such as a methyl group, on the 2-phenyl ring (compounds **14, 15, 18, 19, 22, 23, 26, 27, 30, 31, 34, 35**) does not play any special role even if this ring is surrounded by a hydrophobic pocket delimited by Leu90 (3.32) and Ile268 (7.39).

The effect of a hydrophobic substituent at the R-position is significantly different for the 4-oxo- and 4-aminopyrazoloquinoline derivatives **1-12** with respect to compounds **13-36**. Both the 4-oxo (**1-7**) and 4-amino (**8-12**) derivatives interact only with the upper part of the binding pocket, and the introduction of a methyl group in meta or para position of the phenyl ring (compounds **2, 3, 9, 10**) increases affinity versus the hA_3 receptor. The 2-(4-methoxyphenyl)pyrazoloquinoline derivatives, either 4-oxo or 4-amino substituted (compounds **4** and **11**, respectively), can favorably interact with Ser165 of the second extracellular loop (EL2). The hydroxyl group of Ser is separated by 3 Å from the p-methoxy group and correctly oriented to form a weak

hydrogen bond. The displacement of the methoxy substituent from the para to the meta position (derivatives **5** and **12**) causes the loss of interaction with Ser165. The replacement of the 4-methoxy with a nitro group leads to unfavorable steric and dipolar interactions with Leu90 (3.32) and Ile268 (7.39) that are responsible for the reduction of affinity observed for compound **6**.

In contrast, introduction of the 4-methoxy at the R-position on the 4-acylamino/4-benzylureido derivatives does not produce considerable effects on hA_3 receptor affinity: when the bulkiness of the R_4 substituent is increased, the position of 2-phenyl shifts away from EL2 and, in particular, the hydrogen-bonding interaction with the residue of Ser165 (EL2) is lost.

Finally, the pyrazoloquinoline moiety does not present any specific hydrogen-bonding interaction with Gln167 (EL2), Phe168 (EL2), or Asn250 (6.55) as previously reported for other classes of antagonists, signifying that these interactions are ancillaries with respect to all others mentioned above.

3.3.3   4-modified-2-aryl-1,2,4-triazolo[4,3-a]quinoxalin-1-one Derivatives

Molecular modeling studies were performed on the 2-aryl-1,2,4-triazolo[4,3-a]quinoxalin-1-one derivatives **A** and **1-21** in order to identify the hypothetical binding motif of the new hA_3 antagonists and rationalize the observed SAR. [86] All the reported compounds are listed in the Appendix C and in figure 3.11.

Following our previously reported modeling studies, [82,83,93,97,98] we have constructed a refined model of hA_3 receptor by using a rhodopsin-based homology modeling (RBHM) approach. [94,95,99,100] Moreover, our recently described ligand-based homology modeling (LBHM) approach has been used to simulate the conformational changes induced by ligand binding. [88]

As reported in figure 3.12, depending on the topological properties of the different ligands, we found four different conformational models of the human A_3 receptor reverse agonist-like state in which both shape and chemical complementarities have been specifically optimized around each ligand. In this specific case, with varying ligand structure, the molecular volume of the transmembrane (TM) binding cavity changes from the 660 $\mathring{A}^3$ of the standard RBHM-driven model to the 1120 $\mathring{A}^3$ of the largest LBHM-driven model, without altering the conventional rhodopsin-like receptor topology. The modifications of both 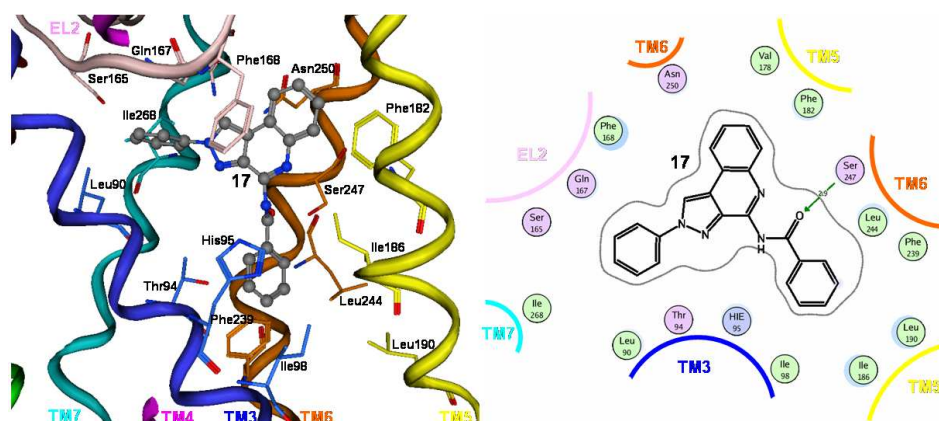shape and volume of the human A_3 TM binding cavity are the most important receptor modeling perturbations obtained by the application of the LBHM technique. The binding cavity reorganization induced by ligand binding is due to the conformational change in several amino acid side chains, such as Leu90 (3.32), Leu91 (3.33), Thr94 (3.36), His95 (3.37), Ile98 (3.40), Gln167 (EL2), Phe168 (EL2), Phe182 (5.43), Ile186 (5.47), Leu190 (5.51), Phe239 (6.44), Trp243 (6.48), Leu244 (6.49), Leu264 (7.35), and Ile268
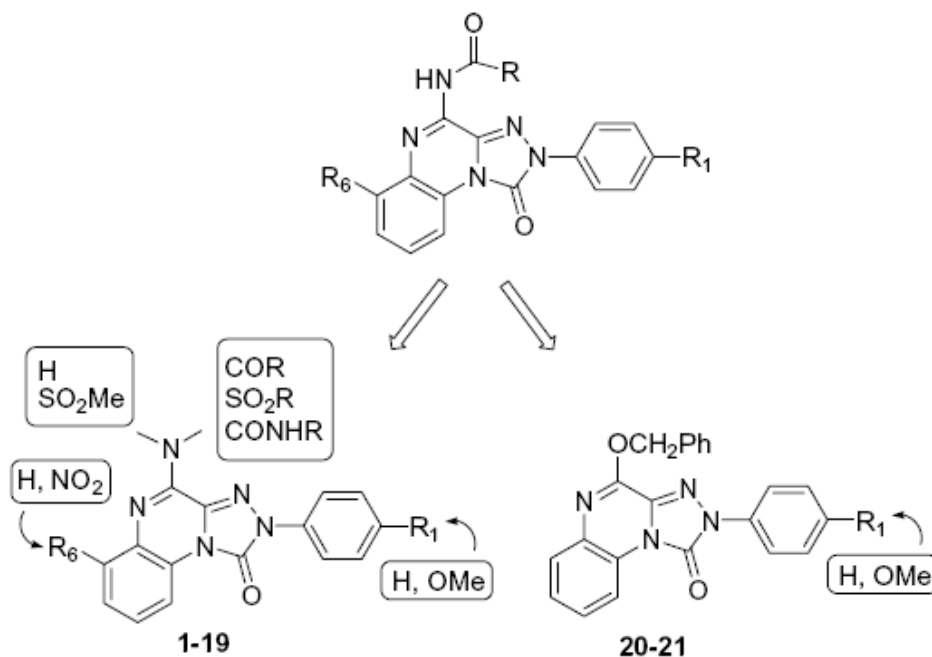
Figure 3.11: Reported 4-modified-2-aryl-1,2,4-triazolo[4,3-a]quinoxalin-1-one derivatives

(7.39).

However, molecular docking studies carried out for all the triazoloquinoxaline antagonists, using the appropriate conformational states of the receptor as listed in figure 3.12, have shown a similar binding motif indicating that a common receptor-driven pharmacophore model can be depicted. This finding is coherent with our previously reported studies. [82,94,95,99,100] Interestingly, none of the new triazoloquinoxaline antagonists found an energetically stable docking pose in the conventional RBHM-driven A_3 model. This is mainly due to the unfavorable topological complementarity among these antagonists and corresponding RBHM-driven TM binding cavity. In particular, highly destabilizing van der Waals interactions (steric conflicts) seem to be the reason for lacking topological complementarities. These steric conflicts are drastically reduced or completely eliminated after the application of the LBHM approach.

As previously described, [82,83,93,97,98] ligand recognition occurs in the upper region of the TM bundle, and the triazoloquinoxaline moiety is surrounded by TMs 3, 5, 6, and 7 with the substituent in the four-position oriented toward the intracellular environment. Furthermore, this hypothetical binding cleft has also been recently suggested by other authors. [101,102]

As shown in figure 3.13, the phenyl ring at the two-position is close to TMs

Figure 3.12: Ligand-based homology modeling (LBHM) data collection of 4-modified-triazoloquinoxalin-1-one derivatives. Each "reference derivative" (compounds **A**, **13** and **15**) was used as ligand template during the homology modeling process. Consequently, four different conformational states (rhodopsin based model and models 1-3) were selected as putative ambassadors of the conformational changes induced by different ligand binding. Depending on their different structure topologies, all other antagonists (docked derivatives) were docked into the most complementary receptor model.

3, 6, and 7. Analyzing our model in detail, all triazoloquinoxaline derivatives share at least two stabilizing hydrogen-bonding interactions inside the binding cleft. The first hydrogen bond is between the carbonyl group at one-position, that points toward the EL2, and the NH2 of the Gln167. This hydrogen-bonding distance is calculated around 2.8 Å for all docked compounds. Moreover, the 1-carbonyl group is also at the hydrogen-bonding distance (ca. 3.2 Å) with the amide moiety of Asn250 (6.55) side chain. This asparagine residue, conserved among all adenosine receptor subtypes, was found to be important for ligand binding. [103,104]

An important hydrogen-bonding network can be observed in all energetically stable docked conformations of all the triazoloquinoxaline antagonists; in particular, Thr94 (3.36), His95 (3.37), and Ser247 (6.52) are able to interact through hydrogen bonds with the 4-carbonyl oxygen of compounds **1-19** and with the ether oxygen of derivatives **20-21**. These polar amino acids seem to be critical for the recognition of all antagonist structures and for receptor selectivity. In particular, Ser247 (6.52) of the hA$_3$ receptor subtype is not present in the corresponding position of hA$_1$ and hA$_2$ receptors, where the residue is
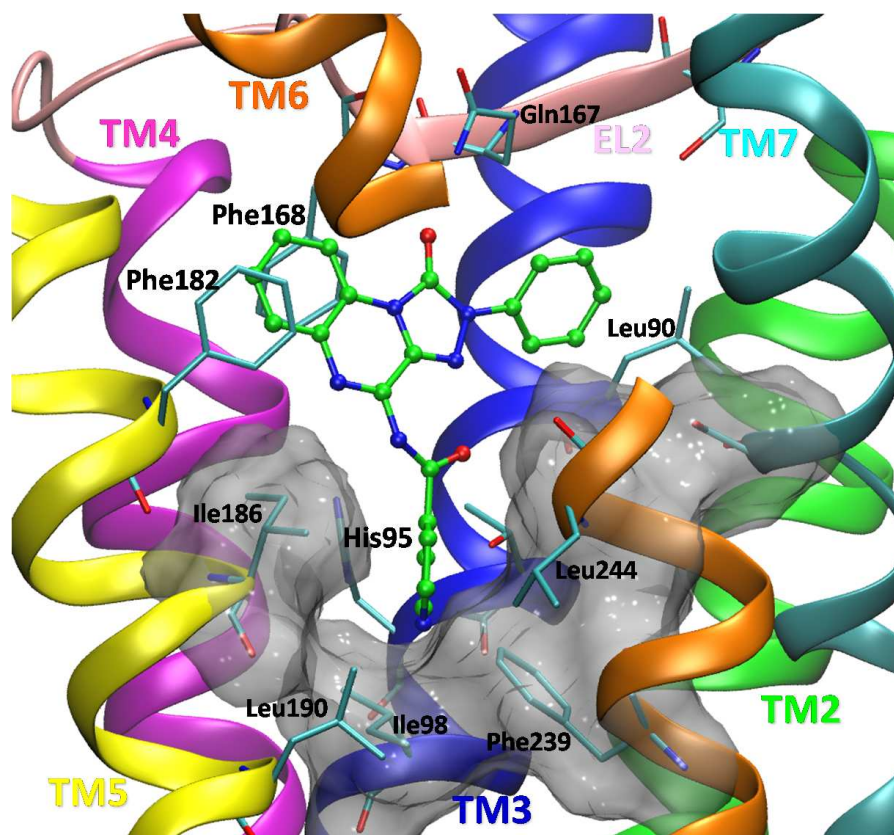
Figure 3.13: Hypothetical binding motif of the representative newly synthesized triazoloquinoxa-line antagonists. The most energetically favorable docked conformation of derivative **4** into LBHM-model 1 is viewed from the membrane side facing TM6. Side chains of some amino acids important for ligand recognition are highlighted. Hydrogen atoms are not displayed. Moreover, the receptor region around R4-substituents characterized by five non-polar amino acids, Ile98 (TM3), Ile186 (TM5), Phe239 (TM6), Phe243 (TM5), and Ser271 (TM7), has been represented by its Connolly's molecular surface.

replaced by a histidine (His251 in hA$_1$, His250 in hA$_{2A}$, and His251 in hA$_{2B}$). The histidine side chain is bulkier than serine and, possibly for this reason, large substituents at the four-position of the triazoloquinoxaline framework are not well tolerated by hA$_1$ and hA$_2$ receptor subtypes. Indeed, 4-acylamino, 4-sulfonamido and 4-benzylureido derivatives are inactive or modestly active on hA$_1$ and hA$_{2A}$ ARs. On the contrary, the hydroxyl group of Ser247 (6.52) of the hA$_3$ receptor is appropriately positioned to form a hydrogen-bonding interaction with the carbonyl oxygen of the 4-amido/sulfonamido/ureido group of compounds **A, 1-19**. In particular, the 4-sulfonamido derivatives **13** and **14** interact simultaneously through hydrogen bonds with all three polar amino acids Thr94 (3.36), His95 (3.37), and Ser247 (6.52). Interestingly, also the 4-benzyloxy analogs **20-21** are selectively accommodated into the hA$_3$ binding cavity. These observations support the importance of the group at the four-

position in modulating receptor selectivity. Indeed, the receptor region around the R_4-substituent is mostly hydrophobic and characterized by five non-polar amino acids: Ile98 (3.40), Ile186 (5.47), Leu190 (5.51), Phe239 (6.44), and Leu244 (6.49), as shown in figure 3.13.

Considering the observed structure-activity relationships in greater detail, methoxy substitution at R_1 position is rather well tolerated among all newly synthesized triazoloquinoxaline derivatives. This is consistent with its accommodation into a tiny hydrophobic pocket delimited by Leu90 (3.32) and Ile268 (7.39). Interestingly, the amino acid corresponding to Leu90 in the hA_3 receptor was found to be essential for the binding of both agonists and antagonists, and it is mutated in valine (Val87) in the human A_1 receptor. This mutation might explain the hA_3 versus hA_1 selectivity. In fact, even if the mutation Leu90 (hA_3)/Val87 (hA_1) can slightly enlarge the dimension of this hydrophobic cavity, simultaneously it also sensibly decreases both shape and hydrophobic complementarities (data not shown). Also the mutation of Ser165 (EL2 of hA_3) with Lys168 in the hA_1 receptor could affect the recognition of the methoxy-substituted triazoloquinoxaline derivatives.

As previously described in the Section 3.3.1 [82] the presence of the 6-nitro substituent has not always produced advantageous effects in terms of hA_3 AR binding affinity. This phenomenon is particularly evident comparing derivatives **6** and **12** with respect to their unsubstituted compounds **4** and **10**. As already anticipated and clearly shown in figure 3, the relative position of R_6-substituent is slightly different depending on the bulkiness of the R_4-substituent on the carbamoyl moiety at the four-position. In particular, in the presence of a less bulky R_4 substituent, the triazoloquinoxaline moiety binds more deeply in the middle of the TM bundle, positioning the 6-nitro substituent very close to TM5. In this case, unfavorable steric and dipolar interactions are responsible for the remarkable reduction of affinity observed for derivatives **6** and **12**. Increasing the bulkiness of the R_4-substituent, the position of the R_6 group shifts away from TM5 and, consequently, more empty space is available for the 6-nitro substituent such as in derivatives **9** and **17**.

3.3.4 Pyrido[2,3-e]-1,2,4-triazolo[4,3-a]pyrazin-1-one Derivatives

Following our recently reported modeling investigations, we used our improved model of the hA_3 receptor, obtained by a rhodopsin-based homology modelling (RBHM) approach, to recognize the hypothetical binding motif of these newly synthesized Pyrido[2,3-e]-1,2,4-triazolo[4,3-a]pyrazin-1-one (PTP) derivatives. [87]

All the pyrido-triazolo-pyrazine derivatives are reported in the Appendix D and in figure 3.14.

Our recently described ligand-based homology modeling (LBHM) approach has been used to simulate the conformational changes induced by ligand bind-
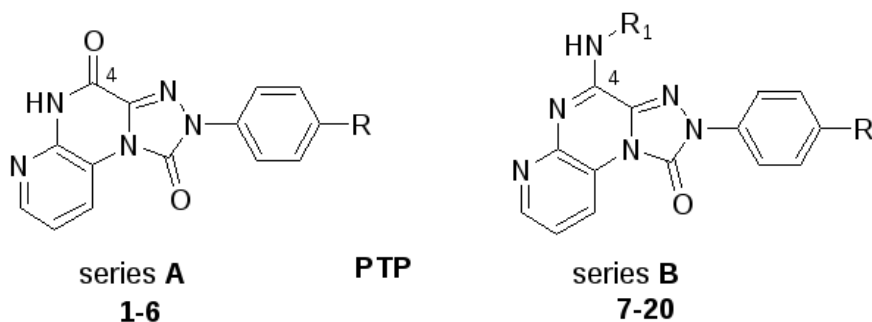
Figure 3.14: Reported pyrido[2,3-e]-1,2,4-triazolo[4,3-a]pyrazin-1-one derivatives

ing. [88] The topological properties of the ligands change depending on the bulkiness of the R1 substituents. According to the volumes, shapes and chemical complementarities of the analyzed compounds we obtained three different conformational models of the human A_3 receptor using the LBHM approach. [88] The volume of the transmembrane (TM) binding cavity changes from 660 Å^3 of the standard RBHM-driven model to 850 Å^3 and 1000 Å^3 of the LBHM-driven models. The conventional rhodopsin-based model was used to detect the atomic level specific interaction of this class of compounds. This model is suitable to rationalize the structure-activity relationships of compounds **1-11**, **14** and **17**. The first ligand-based homology model was built by using compound **15** as reference, and the binding pocket of this model has a volume of 850 Å^3. The model was used to describe the receptor-ligand interactions of compounds **12, 13, 15, 16** and **18**. For compounds **19** and **20**, the volume of the cavity was expanded to 1000 Å^3. The most important receptor modeling perturbation, obtained by the application of the LBHM technique, is the modification of both shape and volume of the human A_3 TM binding cavity, without altering the conventional rhodopsin-like receptor topology. The binding cavity reorganization induced by ligand binding is due to the conformational change in several amino acid side chains, such as: Leu90 (3.32), Leu91 (3.33), Thr94 (3.36), His95 (3.37), Ile98 (3.40), Gln167 (EL2), Phe168 (EL2), Ser181 (5.42), Phe182 (5.43), Ile186 (5.47), Leu190 (5.51), Phe239 (6.44), Trp243 (6.48), Leu244 (6.49), Leu264 (7.35), Ile268 (7.39).

From the docking simulation analysis resulted that all the **PTP** derivatives share a similar binding pose in the TM region of the hA$_3$ adenosine receptor. As shown in figure 3.15, the ligand recognition occurs in the upper region of the TM bundle and the PTP scaffold is surrounded by the TMs 3, 5, 6, 7 with the 1-carbonyl group pointing toward the EL2 and the substituent in the 4-position oriented toward the intracellular environment. The phenyl ring at the 2-position is close to TMs 3, 6 and 7.

As observed for the **TQX** derivatives, the **PTP** antagonists present a π-π
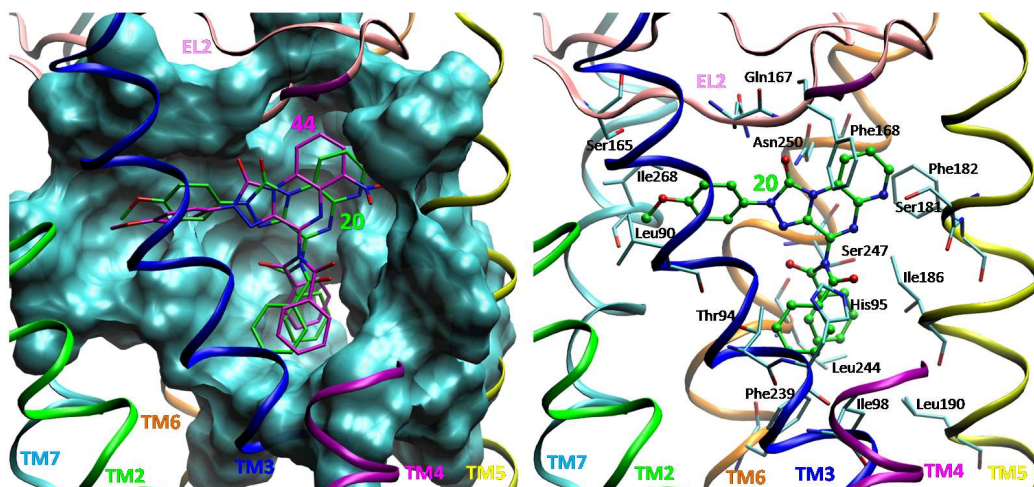
Figure 3.15: (left) Structure superimposition: hypotethical binding motif of a representative newly synthesized Pyrido[2,3-e]-1,2,4-triazolo[4,3-a]pyrazin-1-one antagonists (in green, compound **20**, K_i hA$_3$AR= 7.75 ± 0.8) and a representative compound of 4-Amido-2-aryl-1,2,4-triazolo[4,3-a]quinoxalin-1-ones antagonists (in magenta, compound **44**, K_i hA$_3$AR= 342 ± 21). The most energetically favorable docked conformations of derivatives **20** and **44** are viewed from the membrane side facing TM helices 3 and 4. To clarify the TM cavity, the view of TM4 from Ser138 to Thr144, has been voluntarily omitted. The surface show the shape of the binding pocket that correspond to residues of TM5, 6 and 7. (right) Hypothetical binding motif of compound **20**. Side chains of some amino acids important for ligand recognition are highlighted. Hydrogen atoms are not displayed.

stacking interaction with both side chains of Phe168 (EL2) and Phe182 (5.43) and a hydrogen bonding network in the most energetically stable docked conformations. The first hydrogen bond is between the 1-carbonyl group and the NH of the Glu167 (EL2) and Phe168 (EL2) amidic bond. A second important hydrogen bond involve the side chains of Thr94 (3.36), His95 (3.37) and Ser247 (6.52) that interact with the 4-carbonyl oxygen of compounds **1-6**, the 4-amino group of compounds **7-13** or the 4-acylamino group of compounds **14-20**. This region seems to be critical both for the recognition of all antagonist structures and for receptor selectivity. In particular, Ser247 (6.52) of hA$_3$ receptor subtype is not present in the corresponding position of A$_1$ and A$_2$ receptors, where this amino acid is replaced by histidine (His251 in hA$_1$, His250 in hA$_{2A}$ and His251 in hA$_{2B}$). Histidine side chain is bulkier than serine, and probably for this reason, large substituents at the 4-position of PTP framework are not well tolerated by hA$_1$ and hA$_{2A}$ receptor subtypes. On the contrary, the hydroxyl group of Ser247 (6.52) of hA$_3$ receptor is appropriately positioned to form a hydrogen bonding interaction with the carbonyl oxygen of the 4-amido group of compounds **14-20**. These observations support the importance of a N-acyl group in modulating receptor selectivity. Specifically referring to 4-N-acylated derivatives, the hA$_3$ receptor affinity increases with

the bulkiness of the R_1 substituent (compare the 4-amino compounds **7** and **8** to the 4-acetylamino derivatives **14** and **17** and to the 4-benzoylamino compounds **15** and **18**). Finally, as shown in figure 3.15 by the comparison of the best docking poses of both the pyrido[2,3-e]-1,2,4-triazolo[4,3-a]pyrazin-1-one (in green, compound **20**, K_i hA$_3$AR= 7.75 ± 0.8) and the 4-amido-6-nitro-2-phenyl-1,2,4-triazolo[4,3-a]quinoxalin-1-one antagonists (in magenta, compound **44**, K_i hA$_3$AR= 342 ± 21), the described hydrogen bond interaction between the 6-nitro group of **44** with the side chain of Ser181 (5.42) is now replaced by the interaction with the same aminoacid side chain and the endocyclic nitrogen atom of **20**.

The hydrophobic environment of the five non polar amino acids, Ile98 (3.40), Ile186 (5.47), Leu190 (5.51), Phe239 (6.43) and Leu244 (6.49), can justify this trend of the observed binding affinity. To support this theory, hydrophobic substituents were introduced on 4-amino derivatives: cyclohexyl (compound **12**) and cyclopentyl (compound **13**) interact with this hydrophobic pocket increasing the hA$_3$ receptor affinity (compare compounds **12** and **13** to the unsubstituted 4-amino derivatives **7**).

Considering the substituent on the 2-phenyl ring, the methoxy group turned out advantageous in all the **PTP** derivatives, either 4-oxo, 4-amino or 4-amido substituted. The 2-(4-methoxyphenyl)-derivatives **2, 8, 17, 18, 20** possess higher A$_3$AR affinities than the corresponding 2-phenyl derivatives **1, 7, 14, 15** and **19** because the methoxy substituent can favourably interact with Ser165 (EL2). The hydroxyl group of Ser165 is separated by 2 Å from the p-methoxy group and correctly oriented to form a weak H-bond. The side chains of Leu90 (3.32) and Ile268 (7.39) delimit a small hydrophobic pocket that can accommodate the methoxy substituent, but create unfavourable steric and dipolar interaction with the other groups (OH, F, COOH/COOEt) introduced on the 2-phenyl ring (derivatives **3-6**, **9-11**). Compounds **3** and **9** present a hydroxyl group that looses the hydrophobic interactions with Leu90 (3.32) and Ile268 (7.39) and decreases the hA$_3$AR affinity. The bulkiness of the carboxy acid/ester groups of compounds **5, 6, 11** determine the lack of affinity of these derivatives. The fluorine atom seems to have no effect because the 2-(4-fluorophenyl) derivatives **4** and **10** display comparable affinities to those of the 2-phenyl compounds **1** and **7**. The fluorine atom could interact as hydrogen bond acceptor but, in the most energetically stable conformations of compounds **4** and **10**, the distance between the fluorine and the hydroxy group of Ser165 is more than 3 Å.

In summary, it has to be pointed out that the nitrobenzene moiety of the triazoloquinoxaline-1-one derivatives can be conveniently replaced by the pyridine ring to afford a new class of AR antagonists, the pyrido[2,3-e]-1,2,4-triazolo[4,3-a]pyrazin-1-one derivatives. The electrostatic effect is conserved but the steric clashes created by the nitrobenzene with the backbone of TM5,

and in particular with the peptide bond of Ser181 (5.42) and Phe182 (5.43), have been overcome. Moreover, the endocyclic nitrogen atom can favourably interact with the side chain of Ser181 (5.42) through an hydrogen-bond interaction. This structural modification turned out particularly beneficial in the 4-amino series B when the volume of the molecule is increased by the presence of cicloalkyl and acyl substituents on the 4-amino group. In fact, as it appears by comparing the binding data of some new derivatives to those of the corresponding triazoloquinoxalines [82] (Appendix D, Table D.2), the hA$_3$ affinities of the **PTP** derivatives **12-15**, **17-20** are significantly higher than those of the corresponding **TQX** [82,93,105,106] with the only exception being the 4-benzoylamino derivative **15** that shows a three-fold reduced A$_3$ receptor affinity, compared to the triazoloquinoxaline analogue **40**.

3.3.5   N-5 Substitured Pyrazolo-triazolo-pyrimidine Derivatives

Molecular modeling studies were performed on the pyrazolo-triazolo-pyrimidine derivatives **2-5** in order to identify the hypothetical binding motif of these N-5 analogues and to rationalize their structure-activity relationship. [85] All the docked compounds are listed in the Appendix E and in figure 3.16.
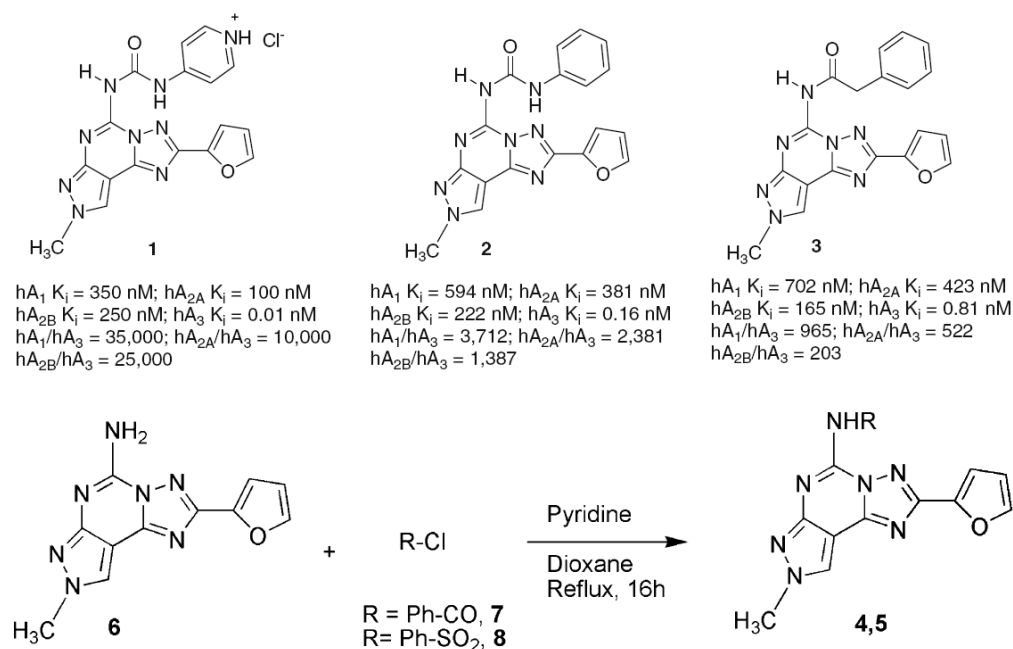
Figure 3.16: Structures and binding profiles of some representative pyrazolo-triazolo-pyrimidines as human A$_3$ adenosine receptor antagonists.

Following our previously reported modeling studies, [89,107–112] we built up a refined model of human A_3 receptor by using a rhodopsin-based homology modeling (RBHM) approach [94,95,99,100]. Moreover, our recently described ligand-based homology modeling methodology (LBHM) has been used to simulate the conformational changes induced by ligand binding. [88]

Using this methodology, we found an "expanded" conformational model of the human A_3 receptor reverse agonist-like state, in which both shape and chemical complementarities have been specifically optimized around each ligand. Considering these new N-5 analogues, the molecular volume of transmembrane (TM) binding cavity has been changed from 660 $Å^3$ (A_3 model obtained by the conventional rhodopsin-based homology modeling) to 840 $Å^3$ (expanded A_3 model obtained by ligand-based homology modeling) without altering the conventional rhodopsin-like receptor topology. The binding cavity reorganization induced by ligand binding is due to the conformational change in several amino acid side chains, such as Leu90 (3.32), Leu91 (3.33), Thr94 (3.36), His95 (3.37), Ile98 (3.40), Gln167 (EL2), Phe168 (EL2), Phe182 (5.43), Ile186 (5.47), Leu190 (5.51), Phe239 (6.44), Trp243 (6.48), Leu244 (6.49), Leu264 (7.35), and Ile268 (7.39).

Interestingly, none of the new pyrazoloquinoline antagonists found an energetically stable docking pose in the conventional RBHM-driven A_3 model. This is mainly due to the unfavorable topological complementarity among these antagonists and corresponding RBHM-driven TM binding cavity. In particular, highly destabilizing van der Waals interactions (steric conflicts) seem to be the reason for absent topological complementarities. These steric conflicts are drastically reduced or completely eliminated after the application of the LBHM approach.

Molecular docking studies were carried out for the pyrazolo-triazolo-pyrimidine antagonists **2-4**, using the "expanded" conformational state of the receptor. As shown in figure 3.17 , we found a similar binding motif indicating that a common receptor-driven pharmacophore model can be depicted. This finding is in agreement with our previously reported studies. [89,107–112]

Indeed, ligand recognition occurs in the upper region of the TM bundle, and the pyrazolo-triazolo-pyrimidine moiety is surrounded by TMs 3, 5, 6, 7 with the substituent in the N5 position oriented toward the intracellular environment. As shown in figure 3.17, the furan ring at the 2-position is close to TMs 3 and 7. Interestingly, an important hydrogen bonding network can be observed in all energetically stable docked conformations of pyrazolo-triazolopyrimidine antagonists. In particular His95 (3.37) and Ser247 (6.52) are able to interact through hydrogen bonding with the N5-carbonyl oxygen of compounds **2-4** ($2_{C=O}$· His95 ca. 3.0 Å; $3_{C=O}$· Ser247 ca. 2.8 Å; $4_{C=O}$·His95 ca. 2.9 Å).

These polar amino acids seem to be critical for the recognition of all an-
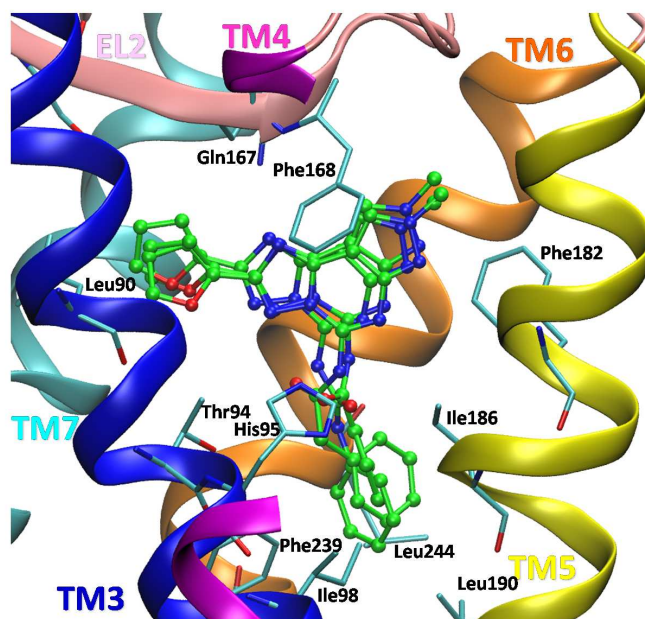
Figure 3.17: Hypothetical binding motif of the newly synthesized pyrazolo-triazolo-pyrimidine antagonists **2-4**. The most energetically favorable docked conformation of each derivative is viewed from the membrane side facing TM helices 4 and 5. To clarify the TM cavity, the view of TM4 was omitted. Side chains of some amino acids important for ligand recognition are highlighted. Hydrogen atoms are not displayed.

tagonist structures and for receptor selectivity. In particular, Ser247 (6.52) of human A_3 receptor subtype is not present in the corresponding position of A_1 and A_2 receptors, where the residue is replaced by a histidine (His251 in human A_1, His250 in human A_{2A} and His251 in human A_{2B}). Histidine side chain is bulkier than serine, and possibly for this reason, large substituents at the N5 position of pyrazolo-triazolopyrimidine framework are not well tolerated by A_1 and A_2 receptor subtypes. In contrast, the hydroxyl group of Ser247 (6.52) of human A_3 receptor is appropriately positioned to form a hydrogen-bonding interaction with the carbonyl oxygen of the N5-amide/ureide group of compounds **2-4**. These observations support the importance of an N5-acyl/carbamoyl group in modulating receptor selectivity. The hydrophobic environment of the five nonpolar amino acids Ile98 (3.40), Ile186 (5.47), Leu190 (5.51), Phe239 (6.44), and Leu244 (6.49) can comfortably accommodate the phenyl ring of all N5-acyl/carbamoyl derivatives.

In contrast, the introduction of the N5-sulfonamido moiety, as present in derivative **5**, drastically reduces the affinity at the human A_3 receptor. Interestingly, in this specific case molecular docking is not able to find an antagonist pose comparable to those described for the other N5-acyl/carbamoyl derivatives. As shown in figure 3.18, the rigid tetrahedral configuration associated
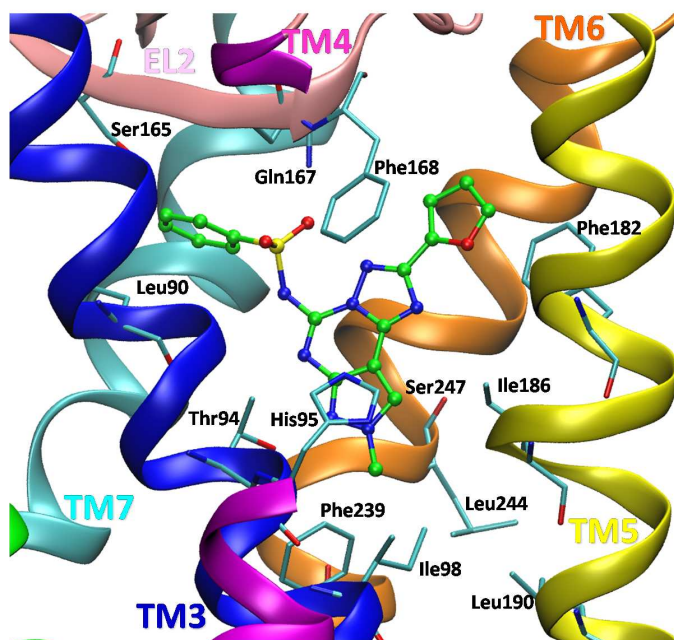
Figure 3.18: Hypothetical binding motif of the newly synthesized N5-sulfonamido pyrazolo-triazolo-pyrimidine antagonist **5**. The most energetically favorable docked conformation of each derivative is viewed from the membrane side facing TM helices 4 and 5. To clarify the TM cavity, the view of TM4 was omitted. Hydrogen atoms are not displayed.
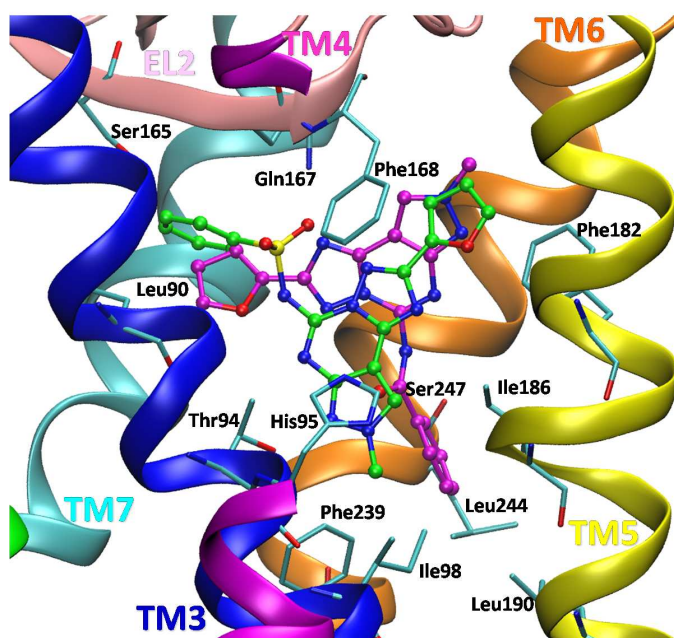
Figure 3.19: Structure superimposition of compounds **4** (in magenta) and **5** (in green) inside the receptor binding site.

with the N5-sulfonamido moiety avoids the sampling of energetically favorable antagonist poses in which the phenyl ring is linked to the N5 position in the hydrophobic pocket delimited by Ile98 (3.40), Ile186 (5.47), Leu190 (5.51), Phe239 (6.44), and Leu244 (6.49).

The most stable docking pose of compound **5** presents the N5-sulfonamido moiety close to TM3 and TM7, and the phenyl ring linked to N5 position is surrounded by a hydrophobic pocket delimited by Leu90 (3.32) and Ile268 (7.39). This antagonist pose is energetically less stable (ca. 15 kcal/mol) with respect to those found for derivatives **2-4**, due to the absence of the stabilizing interactions among the polar residues Thr94 (3.36), His95 (3.37), and Ser247 (6.52) and the N5-sulfonamido moiety. Structure superimposition of compounds **4** and **5** is shown in figure 3.19.

This severe steric constriction might explain the drastic reduction in affinity of derivative **5** at the human A_3 receptor.

3.3.6 Molecular Simplification Approach: From Triazoloquinoxaline to a Pyrimidine Skeleton

Our past research on the study of AR antagonists had been focused for many years on classes of tricyclic compounds. [82,93,97,98,105,113,114] One of these classes is represented by the 2-aryl-1,2,4-triazolo[4,3-a]quinoxalin-1-one derivatives (**TQX** series), either 4-amino or 4-oxo-substituted, which were intensively investigated by evaluating the effect of different substituents on the 2-phenyl ring and on the 4-amino group (Figure 3.20). [82,93,105,113,114]

A R₁=H
C R₁= OMe

B R₁=H
D R₁= OMe

E R₄= Me
F R₄= Ph
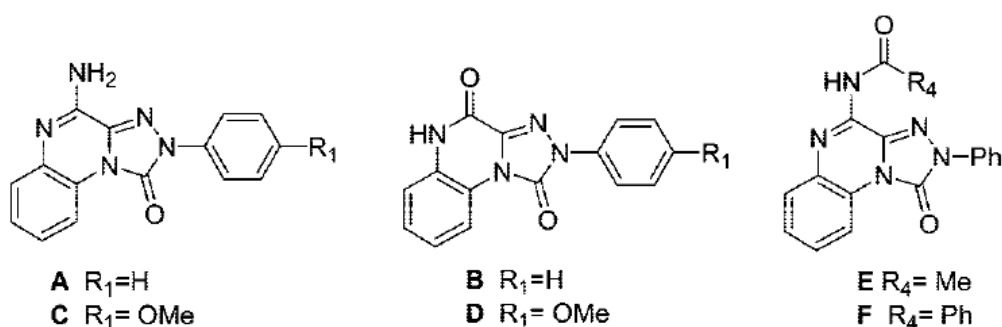
Figure 3.20: Previously reported 2-Aryl-1,2,4-triazolo[4,3-a]quinoxalin-1-ones (TQX Series).

These studies led to the identification of some groups which, introduced one by one in a suitable position of the parent compounds 4-amino-2-phenyl-1,2,4-triazolo[4,5-a]quinoxalin-1-one **A** and 2-phenyl-1,2,4-triazolo[4,5-a]quinoxalin-1,4-dione **B**, afforded high hA₃AR affinity and good selectivity. These groups are the para-methoxy substituent on the 2-phenyl ring (compounds **C** and **D**)

and acyl residues, such as the acetyl or benzoyl groups, on the 4-amino group (compounds **E** and **F**). However, besides potency and selectivity, the straightforward synthesis and pharmacokinetic profile represent crucial requirements in developing new possible therapeutic agents.

Structural simplification represents a drug design strategy to shorten synthetic routes while keeping or enhancing the biological activity of the original candidate. Following this strategy, we have carried out an in silico molecular simplification approach to identify a suitable fragmentation route and explore which of the structural features are essential to guarantee an efficient ligand-receptor recognition. In this context, three series of triazoloquinoxalin-1-one analogues were prepared (Figure 3.21) and, among them, the easily synthesizable 2-amino/2-oxoquinazoline-4-carboxamido derivatives **1-11** (**QZ** series) proved to be highly potent and selective against the hA_3AR.
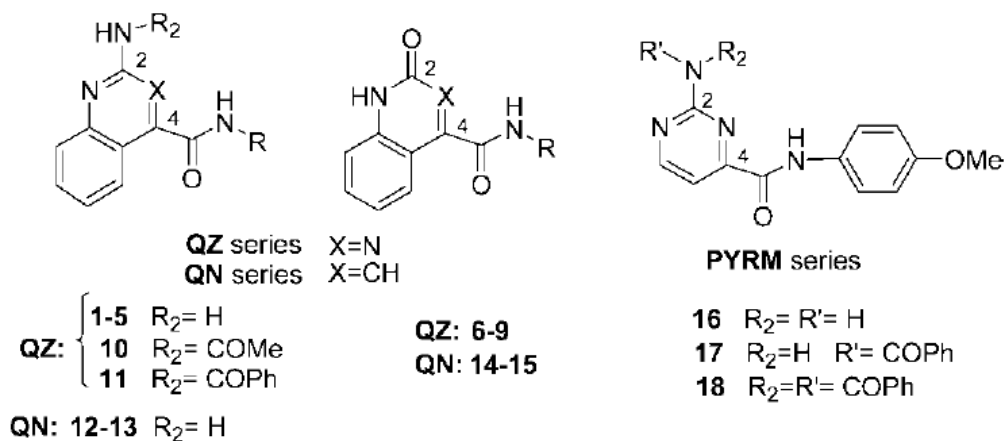
Figure 3.21: Reported 1,2,4-triazoloquinoxalin-1-one simplified analogues.

As previously reported, 4-amino-2-aryl-1,2,4-triazolo[4,3-a]quinoxalin-1-one derivatives nicely bind to hA_3AR. We recognized the hypothetical binding site of the triazolo-quinoxalinone moiety surrounded by transmembrane (TM) regions 3, 5, 6, and 7, with the carbonyl group at the 1-position pointing toward the second extracellular loop (EL2) and the amide moiety in the 4-position oriented toward the intracellular environment. The phenyl ring at the 2-position is positioned close to TM3, TM6, and TM7. The asymmetric topology of the binding cavity is characterized by a major axis (measured from TM1 toward TM5) of about 17 Å and by a minor axis (measured from TM3 toward TM6) of about 6 Å. The peculiar geometric properties of the hA_3 AR binding pocket effortlessly rationalize the experimental evidence that planar polyaromatic systems are usually suitable scaffolds to design potent and se-

lective hA$_3$ AR antagonists. Moreover, planar polyaromatic systems seem to interact through π-π stacking interactions at least with one of the two side chains of Phe168 (EL2) and Phe182 (5.43), as shown in figure 3.22. This interaction has already been described as a crucial pharmacophoric feature in the hA$_3$ AR recognition.
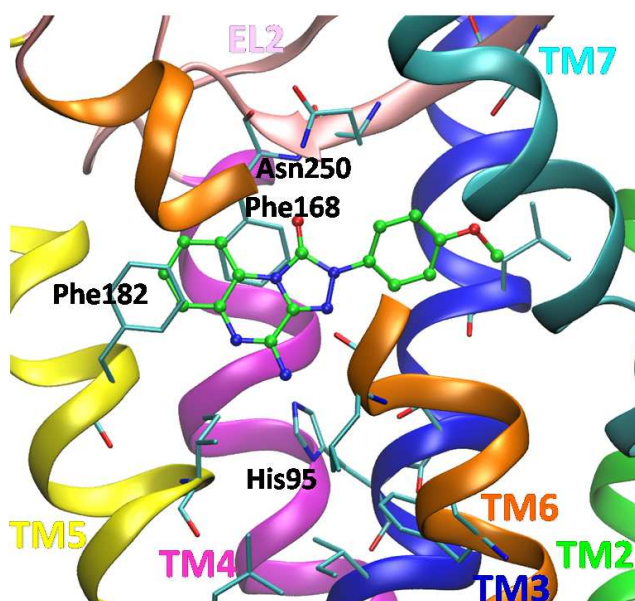
Figure 3.22: Hypothetical binding motif of the reference derivative **C**. The most energetically favorable docked conformation is viewed from the membrane side facing TM helices 5, 6, and 7. To clarify the TM cavity, the view of TM6 from Pro245 to Cys251, was voluntarily omitted. Side chains of some amino acids, important for ligand recognition, are highlighted. Hydrogen atoms are not displayed.

All triazolo-quinoxalinone derivatives also share at least two stabilizing hydrogen-bonding interactions inside the binding cleft (Figure 3.22). The first hydrogen bonding is between the carbonyl group at the 1-position, which points toward the EL2, and the NH of the Gln167. This hydrogen-bonding distance is calculated around 2.7 Å for all docked compounds. Moreover, the 1-carbonyl group is also at the hydrogen-bonding distance with the amide moiety of Asn250 (6.55) side chain. This asparagine residue, conserved among all adenosine receptor subtypes, was found to be important for ligand binding. Second, the NH2 or NHR moiety at the 4-position is surrounded by three polar amino acids: Thr94 (3.36), His95 (3.37), and Ser247 (6.52). This region seems to be very critical for the recognition of all antagonist structures. In fact, a major structural difference between the hypothetical binding sites in these receptor subtypes is that the hA$_3$ receptor does not contain the histidine residue in TM6 (6.52), common to all A$_1$ (His251 in hA$_1$) and A$_2$ (His250

in hA$_2$A and His251 in hA$_2$B) receptors. This histidine has been shown to participate in both agonist and antagonist binding to A$_2$A receptors. In the A$_3$ receptor, this histidine in TM6 is replaced by a serine residue (Ser247 in hA$_3$).

Starting from these binding requirements, we decided to perform an in silico molecular simplification approach to identify a suitable fragmentation route of the 4-amino-triazoloquinoxalin-1-one scaffold and explore which of the structural features were essential to guarantee an efficient ligand-receptor recognition. A schematic representation of our molecular simplification is shown in figure 3.23.
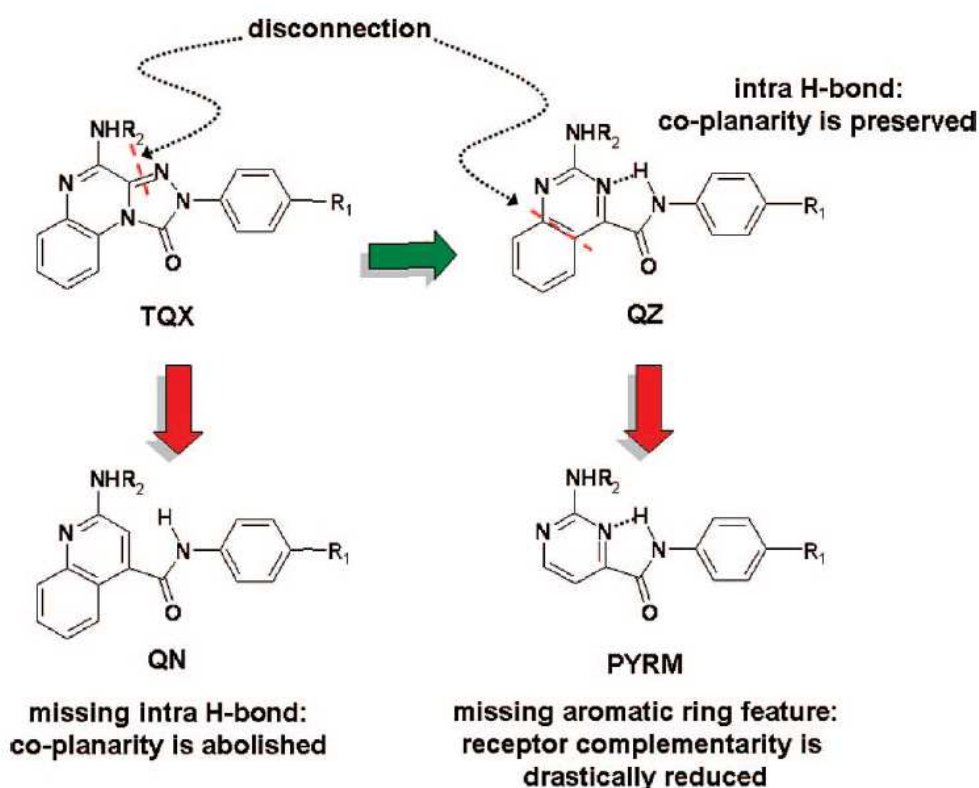
Figure 3.23: Flowchart of the simplification approach.

The first step was to verify the effects of the 4-aminotriazoloquinoxalin-1-one (**TQX** series) replacement with the 2-amino-quinazoline scaffold bearing a CO-NH-C$_6$H$_4$-R$_1$ moiety at the 4-position (**QZ** series). Interestingly, the formation of an intramolecular H-bond between the nitrogen at the 3-position of the quinazoline system and the NH of the amide moiety at the 4-position simulates the presence of a planar tricycle with similar steric properties with respect to the original triazoloquinoxalinone analog. Quantum chemistry calculations support the crucial role of the intramolecular H-bond in stabilizing
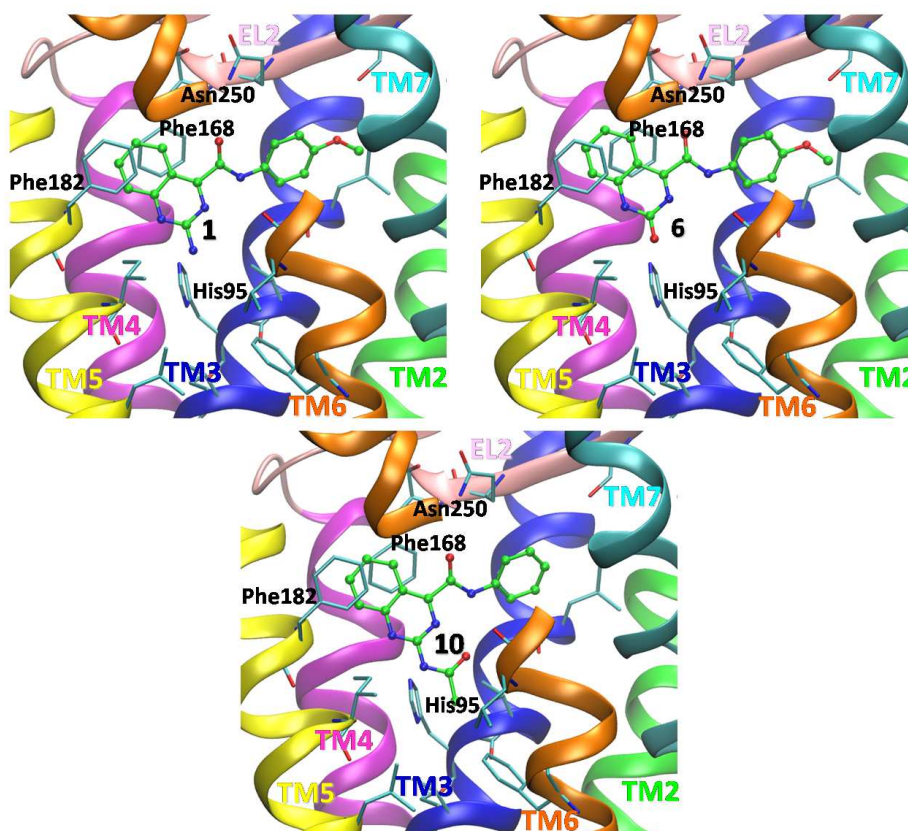
Figure 3.24: Hypothetical binding motif of the newly synthesized A_3 antagonists: 1 (top on the left), 6 (top on the right), and 10 (bottom). The most energetically favorable docked conformations are viewed from the membrane side facing TM helices 5, 6, and 7. To clarify the TM cavity, the view of TM6 from Pro245 to Cys251 was voluntarily omitted. Side chains of some amino acids, important for ligand recognition, are highlighted. Hydrogen atoms are not displayed.

the cyclic conformer.

Indeed, it is worth noting that a different entropy contribution, between the **TQX** and the **QZ** series, could differently affect the total free energy of binding. Unfortunately, in our docking simulations, the entropy effect could not accurately be taken into account. In particular, using the 4-amino-2-(4-methoxyphenyl)-triazoloquinoxalin-1-one derivative **C** (Figure 3.20) as primary reference compound, the corresponding 2-aminoquinazoline-4-carboxyamide derivative **1** (Figure 3.21, R=C_6H_4-p-OMe) was investigated. As shown in figure 3.24, molecular docking simulations confirm that the new compound **1** is efficiently accommodated in the TM binding cavity, maintaining all crucial interactions above-mentioned (π-π stacking interactions at least with both side chains of Phe168 (EL2) and Phe182 (5.43), two H-bonds with Gln167 (EL2) and Asn250 (6.55), and a H-bond interaction with His95 (3.37). In particular, His95 (3.37) is involved in a H-bond interaction with the amino

group at the 2-position of the quinazoline-4-carboxyamide moiety.

Analogously, we decided to extend our investigation, also considering the corresponding 2-oxo analogue of **1**, that is, the 2-oxoquinazoline-4-carboxyamide derivative **6** (Figure 3.21, R=C_6H_4-p-OMe), which can also be considered the simplified analogue of the triazoloquinoxalin-1,4-dione derivative **D** (Figure 3.20). As is clearly shown in figure 3.24, the 2-oxo derivative **6** assumes a binding conformation very similar to that of the 2-aminoquinazoline derivative **1**. In compound **6**, the 2-oxo group interacts through a H-bond interaction with His95 (3.37).

Subsequently, docking studies were also carried out to evaluate whether the presence of acyl residues on the 2-amino group of the new quinazoline-4-carboxamido series (Figure 3.21, **QZ** series, R_2=acyl) was tolerated. The docking simulations, performed on the 2-acetylaminoquinazoline-4-carboxyanilide **10** (Figure 3.21, R=Ph) showed that the acetyl substituent is not only well tolerated, but it might reinforce the binding to the hA$_3$ AR (Figure 3.24). Indeed, consistently with that observed in the triazoloquinoxaline series, an additional H-bond interaction takes place between the carbonyl moiety of the 2-acylamino group and the side chain of Ser247 (6.52).
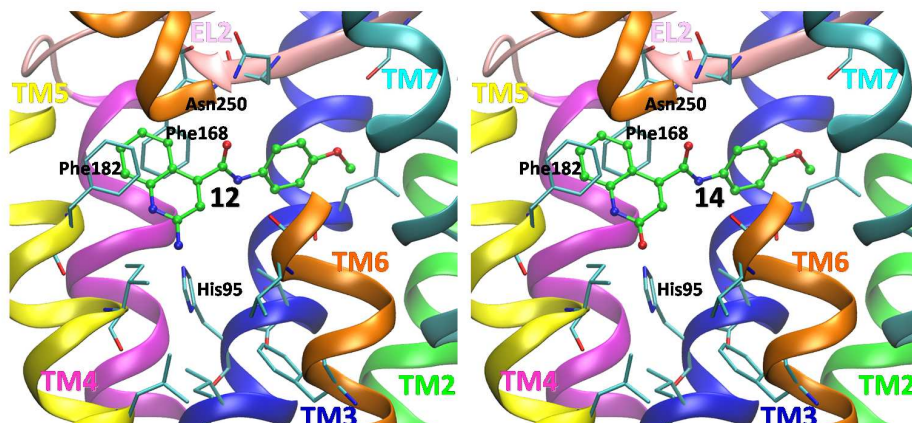
Figure 3.25: Hypothetical binding motif of the newly synthesized analogs 12 (on the left) and 14 (on the right). The most energetically favorable docked conformations are viewed from the membrane side facing TM helices 5, 6, and 7. To clarify the TM cavity, the view of TM6 from Pro245 to Cys251 was voluntarily omitted. Side chains of some amino acids, important for ligand recognition, are highlighted. Hydrogen atoms are not displayed.

To demonstrate the important role of the intramolecular H-bond interaction in maintaining the coplanarity of both 2-amino- and 2-oxo-quinazoline scaffolds and the CO-NHC$_6$H$_4$-R$_1$ moiety at the 4-position, we decided to design a new class of analogs: the 2-aminoquinoline-4-carboxamides (Figure 3.21, **QN** series) and the corresponding 2-oxo derivatives. In fact, in these quinoline derivatives, the formation of the intramolecular H-bond is not al-

lowed and, consequently, the CO-NHC$_6$H$_4$-R$_1$ is twisted with respect to the quinoline ring of about 135°, as suggested by the systematic conformational analysis of the corresponding dihedral angle (data not shown). The impossibility of both 2-amino- and 2-oxo-quinoline systems to adopt a planar conformation is also confirmed by the docking simulations. In fact, as shown in figure 3.25, for the 2-aminoquinoline-4-carboxyamide derivative **12** (Figure 3.21, R=C$_6$H$_4$-p-OMe) and its 2-oxo analogue **14** (Figure 3.21, R=C$_6$H$_4$-p-OMe), the corresponding energetically more stable docking pose is still twisted (of about 121°) and, in this conformation, the 2-amino-quinoline derivatives completely missed some of the most important interactions (in particular, the two H-bonds with Gln167 and Asn250) and drastically reduced their cavity-shape complementarity.
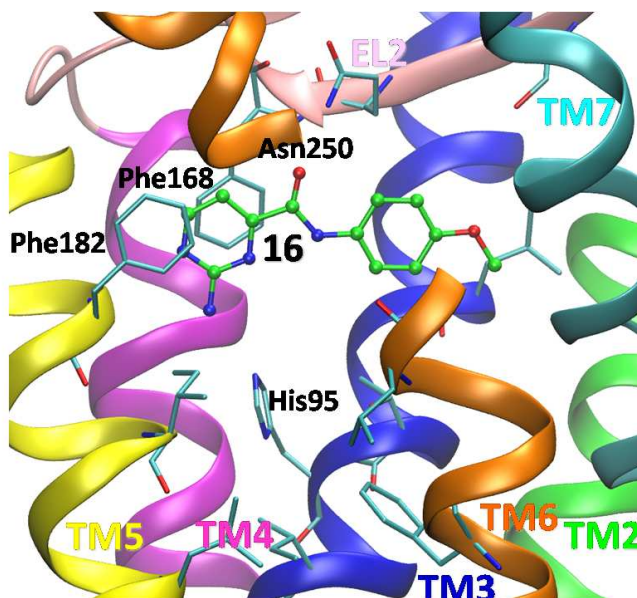
Figure 3.26: Hypothetical binding motif of the newly synthesized analog 16. The most energetically favorable docked conformations are viewed from the membrane side facing TM helices 5, 6, and 7. To clarify the TM cavity, the view of TM6 from Pro245 to Cys251 was voluntarily omitted. Side chains of some amino acids, important for ligand recognition, are highlighted. Hydrogen atoms are not displayed.

Finally, to explore how reducible was the extension of the planar aromatic ring, starting from a 2-aminoquinazoline scaffold, we designed the corresponding 2-aminopyrimidines bearing a CO-NHC$_6$H$_4$-R$_1$ moiety at the 4-position (Figure 3.21, **PYRM** series). As above-described for the quinazoline derivatives, also in this series, we can observe the formation of the intramolecular H-bond between the 3-nitrogen atom of the pyrimidine system and the NH of the 4-amide moiety, which allows a simulation of the presence of a planar bicycle with a missing benzene ring, with respect to the original triazolo-

quinoxalinone analogs. As shown in figure 3.26 for the 2-aminopyrimidine-4-carboxy-(4-methoxyphenyl)amide **16** (Figure 3.21), molecular docking simulations indicate that 2-amino-pyrimidine skeleton maintains the stabilizing π-π stacking interactions with both Phe168 and Phe182. However, the shift of the ligand position into the binding cleft abolishes the possibility an interaction through the H-bond with His95, Gln167, and Asn250, reducing the stability of the corresponding antagonist/receptor complex.

From these theoretical hypotheses, we synthesized and pharmacologically characterized some derivatives belonging to the three designed classes of triazolo-quinoxalinone simplified analogs (see figure 3.21 and Appendix D), that is, the 2-amino/2- oxoquinazoline-4-carboxamides **1-11** (**QZ** series), the 2-amino/ 2-oxoquinoline-4-carboxamides **12-15** (**QN** series), and the 2-aminopyrimidine-4-carboxyamides **16-18** (**PYRM** series).

Among these compounds, there are the above cited and theoretically investigated quinazolines **1**, **6**, and **10**, quinolines **12** and **14**, and pyrimidines **16**, all except one (**10**) bearing the 4-carboxy-(4-methoxyphenyl)amide function. To perform a preliminary structure-affinity relationship (SAR) study, in the first two series, we synthesized derivatives lacking the methoxy group on the 4-carboxyamide moiety, that is, the 4-carboxyanilide compounds **2, 7, 13**, and **15**. In the quinazoline series, the methoxy group was also replaced by lipophilic substituents, such as methyl (compounds **3** and **8**) or bromine (compounds **4** and **9**). In addition, to evaluate the importance of the aromatic phenyl ring on the carboxyamide function, the 2-aminoquinazoline-4-carboxy-cyclohexylamide 5 was synthesized. The effect of a benzoyl residue on the 2-amino function was evaluated both in the quinazoline (compound **11**) and in the pyrimidine (compound **17**) series, and in the latter, the 2-dibenzoylamino derivative **18** was also prepared.

Molecular Docking Protocols Validation

4.1 Introduction

One of the main problem in computational chemistry is the ability to predict the binding mode and estimate the binding affinity for each ligand, given the structure of a protein active site and a list of potential small molecule ligands.

The first step of this problem is the application of computational methods to try to reproduce the bound conformation of a ligand in a high-resolution X-ray crystal structure. This step allows researchers to select the most accurate molecular docking protocol to analyse the ligands.
For many years it has not been possible to validate the molecular docking protocols for GPCR family because no 3D structures of complexes were available. Rodopsin presents his natural ligand in the binding pocket, but retinal represents a particular case because it is covalently bound to the receptor. The release of $A_{2A}AR$, β_2 and β_1 Adrenergic Receptors provided not only new information about the structural conformation of GPCRs, but also information about ligands binding.

We used the new available information to test different molecular docking software and to evaluate the results that we obtained before with SAR studies of antagonists of hA_3AR.

4.2 Materials and Methods

Molecular Docking studies were performed using the following crystal structures:

- human β_2-Adrenergic Receptors (PDB ID: 2RH1) [38]

- turkey β_1-Adrenergic Receptors (PDB ID: 2VT4) [43]

- human A_{2A} Adenosine Receptor (PDB ID: 3EML) [45]

Structures of ligands and proteins were prepared using MOE. Ligands were built using MOE builder and MOPAC (ver.7), [90] was utilized for all quantum

mechanical calculations. Proteins were prepared starting from the crystallographic structures and adding hydrogen atoms, which were minimized until the rms gradient of the potential energy was less than 0.1 kcal mol^{-1} Å^{-1}.

4.2.1   MOE Docking Protocol

Each ligand was docked into the hypothetical TM binding site of the respective receptor by using the MOE-dock tool, part of the MOE suite. [63] Searching is conducted within a user-specified 3D docking box, using one of the three available search protocols:

- Tabu Search

- Genetic Algorithm

- Simulated Annealing

and the MMFF94 force field. [92] MOE-Dock performs a user-specified number of independent docking runs (25 in our specific case) and writes the resulting conformations and their energies in a molecular database file. The resulting docked complexes were subjected to MMFF94 energy minimization until the rms of conjugate gradient was <0.1 kcal mol^{-1} Å^{-1}. Charges for the ligands were imported from the MOPAC output files.

Docking poses were rescored using predicted pKi, that was calculated using MOE. The scoring function is based upon a Bohm-like empirical scoring function consisting of a directional hydrogen-bonding term (direct bonds, water-mediated contacts, transition metals), a directional hydrophobic interaction term, and an entropic term (ligand atoms immobilized in binding).

4.2.2   Glide Docking Protocol

Glide [65] searches for favorable interactions between one or more ligand molecules and a receptor molecule, usually a protein. Shape and properties of the receptor are represented on a grid by several different sets of fields that provide progressively more accurate scoring of the ligand poses. Ligand docking jobs cannot be performed until the receptor grids have been generated.

Receptor grid generation requires a "prepared" structure: an all-atom structure with appropriate bond orders and formal charges. Proteins were prepared with Protein Preparation Wizard of Schrödinger.

Receptor grid was centered at the centroid of the defined ligand molecule, that is the cocristallized molecule. The size of the grid was set as default (20Å x 20Å x 20Å). No constraints were defined.

Glide ligand docking jobs require a set of previously calculated receptor grids and one or more ligand structures.

Extra-precision (XP) docking and scoring were chosen as procedure. Docking is flexible: this is the default option, and directs Glide to generate conformations internally during the docking process. No constraints were defined.

Final scoring is then carried out on the energy-minimized poses. By default, GlideScore multi-ligand scoring function is used to score the poses. GlideScore is based on ChemScore, but includes a steric-clash term and adds buried polar terms devised by Schrödinger to penalize electrostatic mismatches.

25 independent docking poses were written in the output.

4.2.3 Gold Docking Protocol

The binding site was defined starting from a point and the size was defined as a sphere. This is respectively 19.9700 6.7110 1.4950 (radius: 13Å) for β_1-adrenergic receptor, -38.1410 10.3080 4.4190 (radius: 13Å) for β_2-adrenergic receptor and -7.6208 -7.8614 52.6288 (radius: 14Å) for $A_{2A}AR$. The search algorithm is based on Genetic Algorithm. All the options are set as defaults values.

Two different scoring functions were used to perform two separeted docking runs: ChemScore, that is an empirical scoring function and Goldscore that is a force-field-based scoring function. 25 independent docking poses for each scoring function were written in the output.

4.2.4 Plants Docking Protocol

All ligand structures were docked using Plants version 1.08 [66]. The binding site was defined with the same parameters that were used for Gold Docking Protocols (central point and radius). The search algorithm considered 15 ants with an evaporating factor of 0.30. Chemplp scoring funcion was used.

25 structures were generated by the cluster algorithm and the RMSD similarity threshold was set at 1Å.

4.2.5 Autodock Docking Protocol

The compound were docked using Autodock 4. [64] Ligands were considered flexible and no constraints were defined. The grid box was centered on the ligand and the size was defined as 60 points per dimension (x,y,z).
It was used a semi-flexible docking in which only the ligand can explore the conformational space available. The search algorithm that was used is the Lamarckian Genetic Algorithm (LGA) docking also known as a Genetic Algorithm-Local Search (GA-LS). 25 independent docking poses were written in the output.

4.2.6   FlexX Docking Protocol

FlexX was used as an implementation in MOE. FlexX uses an incremental fragment growth strategy to find the poses. Default parameters were used to obtain 25 docking poses.

4.2.7   Clustering

Clustering is the classification of data objects into similarity groups (clusters) according to a defined distance measure.

After the collection of the objects, which are the docking poses in our case, one or more properties has to be calculated to be used for the clustering. The property that define the distance among the poses is the Root Mean Square Deviation (RMSD) and the measures of RMSD are collected in a dissimilarity matrix (or distance matrix). It is a square symmetrical MxM matrix with the ijth element equal to the value of a RMSD between the ith and the jth pose. Distance matrix is calculate using VMD [115] and the iTrajComp plugin. RMSD values are calculated considering all the atoms in the structures.
RMSD distance matix is processed with the software R [116] and a hierarchical clustering is constructed with a agnes-algorithm (Agglomerative Nesting) [117] and the Ward's Method. [118] At first, each observation is a small cluster by itself. Clusters are merged until only one large cluster remains, which contains all the observations. At each stage the two nearest clusters are combined to form one larger cluster.

To analyze the membership of each structure to the clusters, one can cuts the hierarchical structure at a user defined level. The cutting level is defined by the final number of clusters that one wishes to obtain or by the RMSD value that defines the maximum difference between two members of the same cluster.

4.3   Results and Discussion

The availability of crystallographic structures of GPCRs with a ligand cocrystalized allowed us to validate the docking protocol that we had been using before.

We collected the docking results obtained with different search algorithms and scoring functions. We compared the best docking pose, according to the scoring function that was used, with the crystallographic pose in term of root mean square deviation (RMSD) measured in Å and the conformational sampling, that is the number of poses in a docking result that present an

RMSD lower than 2,5 Å in comparison to the crystallografic pose of the ligand in the complex.

Comparison of different docking protocol was conducted using the following crystal structures of complexes with ligand and protein:

- Carazolol on human β_2-Adrenergic Receptor (PDB ID: 2RH1) [38]

- Cyanopindolol on turkey β_1-Adrenergic Receptor (PDB ID: 2VT4) [43]

- ZM241385 on human A_{2A} Adenosine Receptor (PDB ID: 3EML) [45]

4.3.1   Carazolol on human β_2-Adrenergic Receptor

Docking results obtained with different docking protocols are summarized it the table 4.1 and the best results are in figure 4.1.

The majority of the protocols is able to reproduce the crystallographic pose of Carazolol with an RMSD lower than 1 Å.

The docking protocol that can better reproduce the conformation of Carazolol in the crystal structure is Gold with the scoring function Goldscore: the RMSD between the best ranked pose and the crystallographic pose is 0,59 Å and 20 out of 25 poses present a docking pose with an RMSD value lower than 2,5 Å.

Another protocol that reproduces the crystallographic pose with good results is FlexX: 1,02 Å of RMSD for the best pose and all the poses have an an RMSD value lower than 2,5 Å.

Table 4.1: Carazolol - human β2-Adrenergic Receptor

Docking Protocol	RMSD (Å)	Sampling
moe-GA-EnTot	1,58	5/25
moe-GA-pKi	5,16	5/25
moe-SA-EnTot	0,78	3/25
moe-SA-pKi	6,85	3/25
moe-TS-EnTot	0,93	7/25
moe-TS-pKi	0,93	7/25
glide	1,05	18/25
gold-chemscore	0,74	5/25
gold-goldscore	0,59	20/25
plants	0,68	13/25
autodock	0,55	13/25
flexX	1,02	25/25
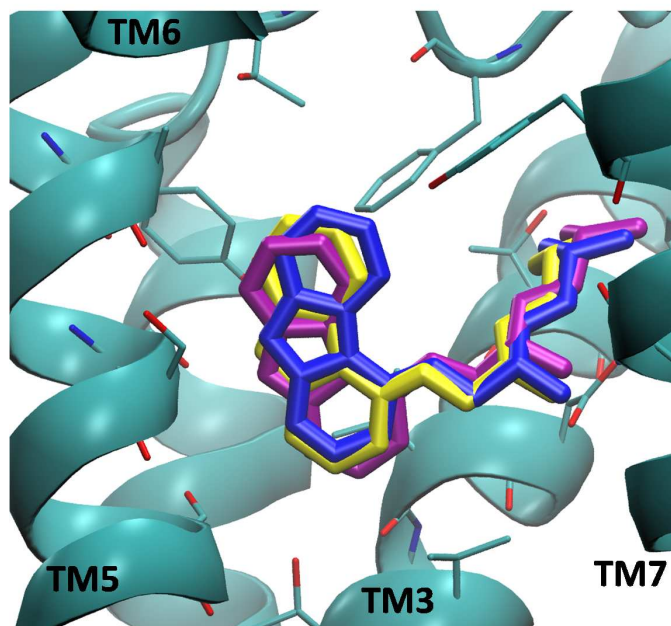
Figure 4.1: Docking results of carazolol on β_2-AR. Crystallographic pose is represented in blu, the best pose of Gold protocol is represented in yellow, that is the best docking protocol according to our analysis. In magenta the best docking pose obtained with the protocol Tabu Search of MOE, the protocole used for the SAR studies of antagonists of hA$_3$AR. Antagonists are viewed from the membrane side facing TM6, that has been voluntarily omitted. Side chains of some amino acids important for ligand recognition are highlighted. Hydrogen atoms are not displayed.

Similar results are available using Glide: the best pose according to the software presents 1,05 Å of RMSD compared to the crystallographic pose and 18 out of 25 poses have an an RMSD value lower than 2,5 Å.

The best poses of carazolol according to the protocols of Plants and Autodock have 0,68 and 0,55 Å of RMSD with the crystallographic pose and in both cases more than half of the poses (13 out of 25) has an RMSD value lower than 2,5 Å.

4.3.2   Cyanopindolol on turkey β_1-Adrenergic Receptor

Docking results obtained with different docking protocols are summarized it the table 4.2 and the best results are in figure 4.2.

The docking protocol of Glide is the best protocol in this case and it reproduced the crystallographic pose with an RMSD of 0,28 Å and 23 out of 25 poses have an an RMSD value lower than 2,5 Å.

Gold with the scoring function goldscore is among the best protocols also in this case. All the poses of the output have an an RMSD value lower than 2,5 Å and the RMSD between the best ranked pose and the crystallographic structure is lower than 1 Å.
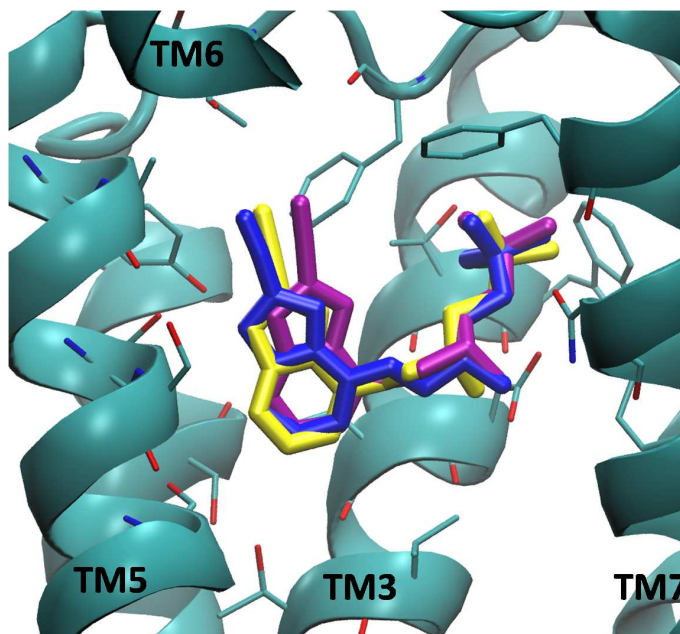
Figure 4.2: Docking results of cyanopindolol on β_1-adrenergic receptor. Crystallographic pose is represented in blu, the best pose of Glide protocol is represented in yellow, that is the best docking protocol according to our analysis. In magenta the best docking pose obtained with the protocol Tabu Search of MOE, the protocole used for the SAR studies of antagonists of hA$_3$AR. Antagonists are viewed from the membrane side facing TM 6, that has been voluntarily omitted. Side chains of some amino acids important for ligand recognition are highlighted. Hydrogen atoms are not displayed.

Table 4.2: Cyanopindolol - turkey β1-Adrenergic Receptor

Docking Protocol	RMSD (Å)	Sampling
moe-GA-EnTot	2,26	4/25
moe-GA-pKi	1,65	4/25
moe-SA-EnTot	4,22	3/25
moe-SA-pKi	5,79	3/25
moe-TS-EnTot	3,25	2/25
moe-TS-pKi	0,98	2/25
glide	0,28	23/25
gold-chemscore	3,93	5/25
gold-goldscore	0,67	25/25
plants	1,15	15/25
autodock	1,13	16/25
flexX	1,63	1/25

Also Autodock and Plants reproduce the crystallographic pose with good results.

The best ranked pose according to TS algorithm and pki scoring funcion of MOE has an RMSD value lower than 1 Å if compared with the crystallographic ligand, but the sampling is very poor, only 2 poses out of 25.

4.3.3 ZM241385 on human A_{2A} Adenosine Receptor

Docking results obtained with different docking protocols are summarized it the table 4.3 and the best results are in figure 4.3.

Docking results of ZM241385 are less accurate than the previous ones. Ligand is bigger and the binding pocket is more open: conformational search can explore more empty space and the crystallographic pose is reproduced with lower precision.

In this case FlexX is the protocol that works better: all the 25 poses of the output have an an RMSD value lower than 2,5 Å and between the best ranked pose and the crystallographic pose there is the lowest RMSD for this analysis.

Other protocols that give fairly good results are, also in this case, Glide, Gold with Goldscore, Autodock and Plants.

Table 4.3: ZM241385 - human A_{2A} Adenosine Receptor

Docking Protocol	RMSD (Å)	Sampling
moe-GA-EnTot	6,07	1/25
moe-GA-pKi	1,41	1/25
moe-SA-EnTot	1,77	6/25
moe-SA-pKi	1,87	6/25
moe-TS-EnTot	2,15	5/25
moe-TS-pKi	2,16	5/25
glide	2,86	10/25
gold-chemscore	3,93	9/25
gold-goldscore	3,05	11/25
plants	2,00	15/25
autodock	2,95	16/25
flexX	1,39	25/25
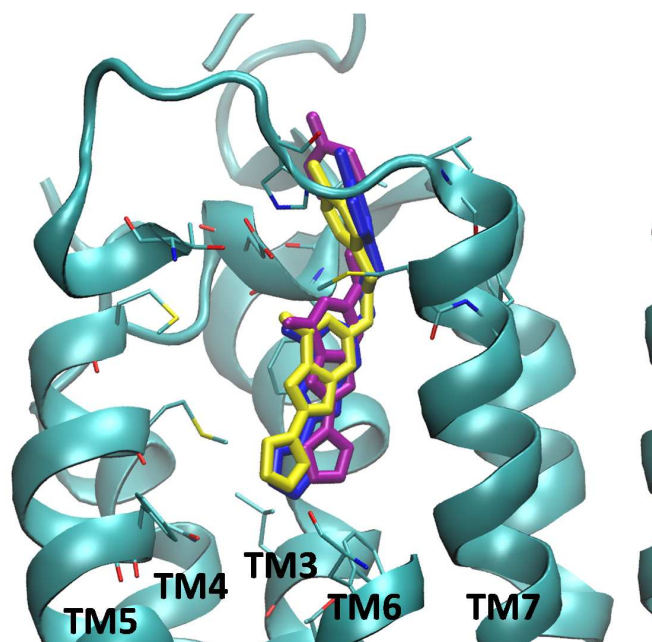
Figure 4.3: Docking results of ZM241385 on human A_{2A} adenosine receptor. Crystallographic pose is represented in blu, the best pose of FlexX protocol is represented in yellow, that is the best docking protocol according to our analysis. In magenta the best docking pose obtained with the protocol Tabu Search of MOE, the protocole used for the SAR studies of antagonists of hA_3AR. Antagonists are viewed from the membrane side facing TM6, that has been voluntarily omitted. Side chains of some amino acids important for ligand recognition are highlighted. Hydrogen atoms are not displayed.

The protocol that was used for the SAR studies of antagonists of hA_3AR is Tabu Search algorithm implemented in MOE software and we used the scoring function that predict the pKi to rescore the poses.

The sampling of poses for the crystal structures using this protocol is poor: 7 poses out of 25 for Carazolol, 2 poses out of 25 for Cyanopindolol and 5 poses out of 25 for ZM241385. Anyway the scoring function is able to select good poses among all the results. If we rank the results according to the score of pKi, the best poses have an RMSD value with the crystal structures of 0.93 Å, 0.98 Å and 2.16 Å respectively for Carazolol, Cyanopindolol and ZM241385.

In general we can say that the protocol that we used (MOE software, Tabu Search algorithm and pKi as scoring function) is acceptable. Before we didn't have a basis for comparison for GPCRs, for this reason the protocol was chosen among the available protocols and according to the one that better described the SAR among the analyzed antagonists.

4.3.4   Analysis of Previously Reported Docking Results with Different Docking Protocols

To verify the results that we obtained previously with the SAR analysis of antagonists of human A_3AR, we considered some compounds that present affinity for hA_3AR and we performed a molecular docking study with the available docking protocols. We clusterized the results and selected the most populated groups as representative binding poses.

4.3.4.1   4-Amido-2-aryl-triazolo-quinoxalin-1-one Derivative

4-Amido-2-aryl-triazolo-quinoxalin-1-one derivatives are the compounds that have been analysed more extensively and we are now considering derivative **A** reported in Appendix A and in figure 4.4.
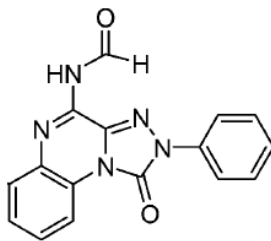
Figure 4.4: 4-Amido-2-aryl-triazolo-quinoxalin-1-one derivative used for validation of docking protocols.

From the cluster analysis, two clusters were selected: *poseA* and *poseB* (Figure 4.5). Three of the docking protocols (Gold with the scoring function Goldscore, Plants and Glide) present docking poses in both selected clusters, all the other protocols present only poses that belong to cluster *poseB*.
Poses that were not selected in one of the two clusters were condidered outliers: they were not part of any of the two selected most populated clusters (*poseA* and *poseB*) and they were not enough similar among them, in term of RMSD value, to form a new cluster.

The most populated cluster is the cluster *poseB*, to which belong the previously reported pose, obtained with the docking protocol used for all SAR study published before the release of the crystallographic structures of GPCRs with a ligand cocrystallized (Figure 4.5 and 4.6).
In the table in figure 4.6 RMSD value are reported: these values are calculated using as references two average conformations, one of the cluster *poseA* and one of the cluster *poseB*.

Anyway, the selection of the best docking pose is usually not limited to the most populated cluster, or the best pose in term of score value, but it is

selected with an accurate SAR study, considering a series of derivatives with similar chemical structures and available affinity data.
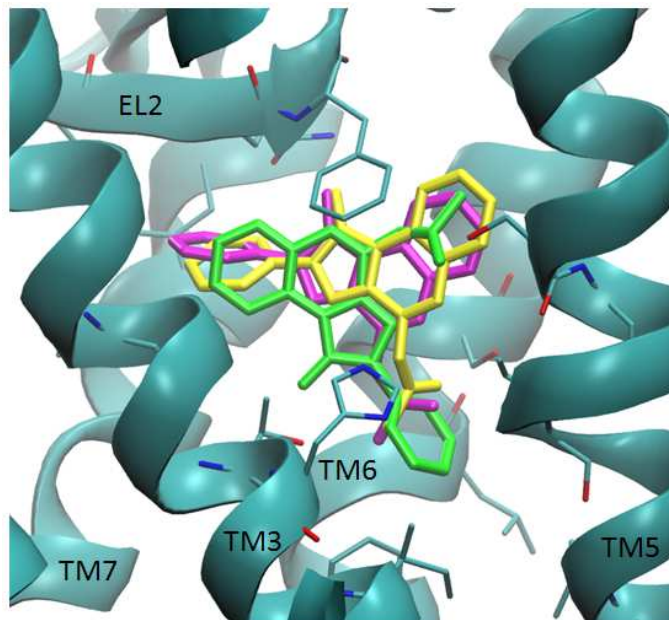
Figure 4.5: Docking results of compound **A** of triazolo-quinoxalin-1-one derivatives. In yellow: poseB, this is the best docking pose according to GA of Gold as search algorithm and Goldscore as scoring function; in green: poseA; in magenta: best docking pose obtained with TS algorithm of MOE and pKi as scoring function, this is the pose reported in SAR studies of triazolo-quinoxalin-1-one derivatives. Docking poses are viewed from the membrane side facing TM helices 3 and 4. To clarify the TM cavity, the view of TM4 from Leu136 to Pro145 has been voluntarily omitted. Side chains of some amino acids important for ligand recognition are highlighted. Hydrogen atoms are not displayed.

	plants_A	plants_B	goldscore_A	goldscore_B	gold_chemscore	glide	autodock
plants_poseA	0.00	6,63	0,66	6,39	6,4	6,45	6,62
plants_poseB	6,63	0,00	6,66	1,06	0,78	0,77	0,81
goldscore_A	0,66	6,66	0,00	6,46	6,46	6,52	6,67
goldscore_B	6,39	1,06	6,46	0,00	0,71	0,72	1,13
gold_chemscore	6,4	0,78	6,46	0,71	0,00	0,65	0,94
glide	6,45	0,77	6,52	0,72	0,65	0,00	1,02
autodock	6,62	0,81	6,67	1,13	0,94	1,02	0,00
moe_GA_Etot	6,63	0,53	6,68	0,86	0,67	0,73	0,85
moe_GA_pKi	6,63	0,49	6,67	0,86	0,67	0,66	0,89
moe_TS_Etot	6,54	0,56	6,6	1,18	0,77	0,97	0,94
moe_TS_pKi	6,64	0,36	6,68	0,94	0,7	0,63	0,91
moe_SA_Etot	6,56	0,54	6,62	1,07	0,73	0,94	0,85
moe_SA_pKi	6,63	0,36	6,68	0,89	0,68	0,73	0,75
flexX_poseA	0,61	6.53	0,60	6.31	6.31	6.38	6.52
flexX_poseB	6,61	1.21	6,64	1.14	1.17	1.05	1.21

	moe_GA_Etot	moe_GA_pKi	moe_TS_Etot	moe_TS_pKi	moe_SA_Etot	moe_SA_pKi	flexX_A	flexX_B
plants_poseA	6,63	6,63	6,54	6,64	6,56	6,63	0,61	6,61
plants_poseB	0,53	0,49	0,56	0,36	0,54	0,36	6,53	1,21
goldscore_A	6,68	6,67	6,6	6,68	6,62	6,68	0,60	6,64
goldscore_B	0,86	0,86	1,18	0,94	1,07	0,89	6,31	1,14
gold_chemscore	0,67	0,67	0,77	0,7	0,73	0,68	6,31	1,17
glide	0,73	0,66	0,97	0,63	0,94	0,73	6,38	1,05
autodock	0,85	0,89	0,94	0,91	0,85	0,75	6,52	1,21
moe_GA_Etot	0,00	0,42	0,72	0,43	0,59	0,26	6,52	1,15
moe_GA_pKi	0,42	0,00	0,88	0,23	0,81	0,47	6,53	0,95
moe_TS_Etot	0,72	0,88	0,00	0,78	0,23	0,56	6,44	1,60
moe_TS_pKi	0,43	0,23	0,78	0,00	0,72	0,43	6,54	0,99
moe_SA_Etot	0,59	0,81	0,23	0,72	0,00	0,43	6,45	1,52
moe_SA_pKi	0,26	0,47	0,56	0,43	0,43	0,00	6,52	1,24
flexX_poseA	6.52	6.53	6.44	6.54	6.45	6.52	0.00	6.54
flexX_poseB	1.15	0.95	1.60	0.99	1.52	1.24	6.54	0.00

Figure 4.6: Comparison of docking results, in terms of RMSD in Å, of compound A on hA$_3$AR using different docking protocols.

Molecular Dynamics of Adenosine Receptors

5.1 Introduction

Homology models represent a rigid conformation of a protein, but proteins are known to be dynamic molecules that show rapid, small-scale structural fluctuations. [119]

A simple two-state model can describe a receptor: a conformation that binds the agonist and transfers the signal and a conformation that binds the antagonist. It is well known that GPCRs behave in a more complex way. Efficacy can be explained by a simple model of receptor activation, but evidence from both functional and biophysical studies supports the existence of multiple, ligand specific conformational states. [68]

Our models of human A_3AR were built using homology modeling technique. As it was deeply analyzed in Chapter 2, there are differences among the models that have to be considered when one wants to use them for drug design.

We consider that our models correspond to the antagonist-like state of hA_3AR, but this pharmacological state can be described by more than one conformational state. Which one of these models better characterizes the antagonist-like state of hA_3AR, if one of these model can evolve to another one, if the models can converge to a common conformation are questions that remain to be answered. We investigated the molecular dynamic behaviour of the models in a lipid bilayer to try to answer to these questions.

5.2 Materials and Methods

MD simulations were carried out starting from the models of hA_3AR inserted into a lipid bilayer environment. The lipid bilayer was built starting from an existing bilayer as described by C. Kandt at all. [120] Water was added using an initial box and redundant water was deleted based on their z position. Ten chlorine ions were added to neutralize the system. The membrane was equilibrated for 10 ns.

MD simulation was carried out using the GROMACS 3.3.1 MD package [121,122] applying periodic boundary conditions. The simulation was carried

out for 30 ns (time step = 2 fs), with a constant temperature of 300 K, using a Berendsen (τT = 0.1 ps) thermostat, [123] while coupling the protein, lipid and water/ions separately. The pressure was maintained at 1 bar using a Berendsend coupling algorithm [123] with a coupling constant of 1.0 ps and a compressibility of 4.6 x 10^{-5} bar^{-1}. Electrostatic interactions were evaluated using the PME (particle mesh Ewald) methods [124,125] with a cutoff of 1.0 nm. The long-range electrostatic interactions were calculated with fourth-order B-spline interpolation and a Fourier spacing of 0.14 nm. The Lennard-Jones interactions were evaluated using a twin-range cutoff (1 and 1.4 nm) with the neighbor list updated every ten steps. All bonds in the system were constrained using LINCS [126].

5.3 Results and Discussion

Molecular dynamics simulations were performed starting from the following models of hA_3AR:

- hA_3AR built from bovine rhodopsin (PDB ID 1F88);

- hA_3AR built from $h\beta_2$-adrenergic receptor (PDB ID 2RH1);

- hA_3AR built from $hA_{2A}AR$ (PDB ID 3EML).

MD was carried out with the same protocol for the three models.

In the following pages are represented some preliminary results of the simulations.

The graphs that are reported in figures 5.2, 5.5 and 5.8 show the RMSD per residue of the backbone. On the x axis is reported the number of the aminoacids and on the y axis is reported the time of MD simulation, expressed in nanoseconds. Colors symbolize the RMSD value in Å calculated using as reference the conformation of the protein at the beginning or the MD run, after the equilibration step.

In these graphs it is easy to visualize which are the regions of the protein that are more flexible, because they are colored in red, orange, yellow or green, from the more to the less flexible. Residues that belong to loops, N-term and C-term are the more flexible, while residues of TM regions are caracterized by blue or white color, that means that the RMSD is always lower than 4.

Similar analysis is shown in the graphs reported in figures 5.3, 5.6 and 5.9. In these graphs are reported the values of RMSD in function of time (in ns). In all the graphs in black is reported as reference the RMSD of the whole backbone. In the upper parts there are the RMSD values of the backbone of TM regions, that are always lower than the RMSD of the backbone of the whole structure. The RMSD of the backbone of the loops may vary: very

short loop like IL1 and EL1 have low values of RMSD, bigger loop are more flexible, together with N-term and C-term.

The loop that presents the biggest change of conformation is IL3. This loop is known to vary considerably among GPCRs, and probably the flexibility and variability of this region may be critical for the functionality and specificity of G-protein activation. The conformational change of the loop does not affect the conformation of the binding pocket, but further investigations of this loop may be interesting to understand its role in the transmission of the signal. In the structure of hA_3AR built using bovine rhodopsin as template some residues reach an RMSD higher than 20 Å and the average RMSD of this loop reaches 10 Å after 11 ns of MD simulation, than the conformation is more stable and the loop oscillates around that position. In the model of hA_3AR built from $h\beta_2AR$, IL3 is less flexible, but it seems that the conformation is not stable even after 20 ns. In the third model, the one built from $hA_{2A}AR$, there is a fast conformational change of this loop in the first nanosecond of simulation, but, after this change, the conformation seems stable and the RMSD value doesn't change any more.

N-term and C-term are also very flexible. This is probably due to the fact that these domains are more exposed and connected to the protein with only one end.

The conformational change of EL2 is of particular interest, because EL2 constitutes one of the main differences among the templates used in homology modeling and it influences the conformation of the binding pocket. For structure-based drug design the conformation of the binding site is crucial.

In the templates, the conformation of EL2 is influenced by the presence of disulfide links that create constraints that keep the loop in a particular conformation. As we discussed before in Section 2.3.1, hA_3AR does not have the same cysteine residues that are present in $h\beta_2AR$ and $hA_{2A}AR$. The conformation of the EL2 of the models of hA_3AR follow the conformation of the templates, but it presents only one disulfide bridge that is the one conserved among family A GPCRs. It is interesting to analyse the behaviour of this loop in an environment that mimics the membrane.

Starting and final conformations of EL2 in the three models are in figures 5.1, 5.4 and 5.7. Starting conformations are in yellow (hA_3AR from rhodopsin), in magenta (hA_3AR from $h\beta_2AR$) and in cyan (hA_3AR from $hA_{2A}AR$); final conformations are in blue. Red arrows represent the displacements of $C\alpha$ of EL2 in 30 ns of MD.
EL2 confomational change is stronger in the models built starting from $h\beta_2AR$ and $hA_{2A}AR$ than in the model built from rhodopsin. It may be interesting to compare the conformational changes of EL2 in the models and in the crystal structures that were used as templates to check the importance of the disulfide links in preserving the conformation of the loop.
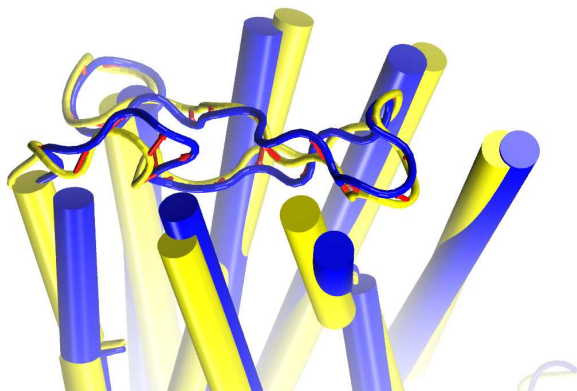
Figure 5.1: Representation of the second extracellular loop of the hA$_3$AR model built using bovine rhodopsin as template before (in yellow) and after (in blue) 30 ns of molecular dynamics in a lipid bilayer.
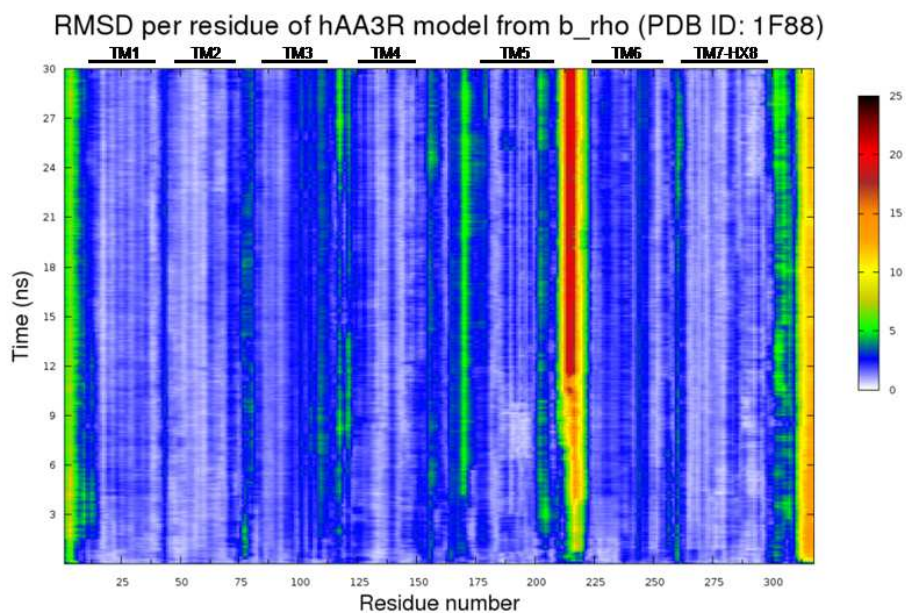
Figure 5.2: RMSD per residue of the backbone of the hA$_3$AR model built using bovine rhodopsin as template.
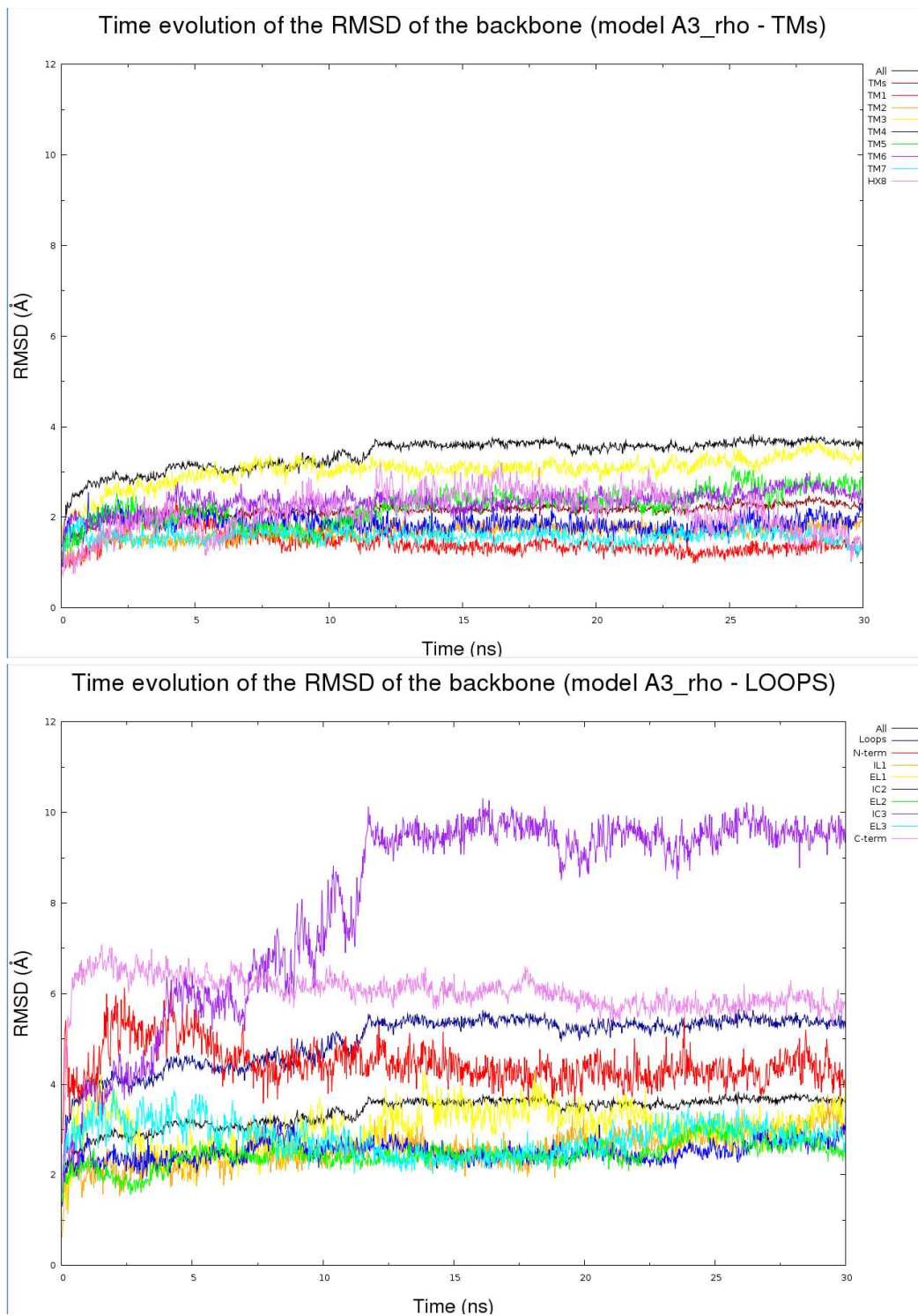
Figure 5.3: Time evolution of the RMSD of C α of the hA$_3$AR model built using bovine rhodopsin as template. On the top, RMSD of the TM regions; on the bottom, rmsd of the loops, N-term and C-term.
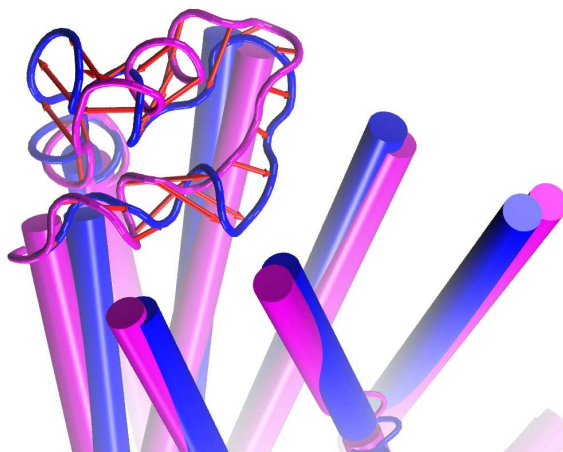
Figure 5.4: Representation of the second extracellular loop of the hA$_3$AR model built using β_2-adrenergic receptor as template before (in magenta) and after (in blue) 30 ns of molecular dynamics in a lipid bilayer.
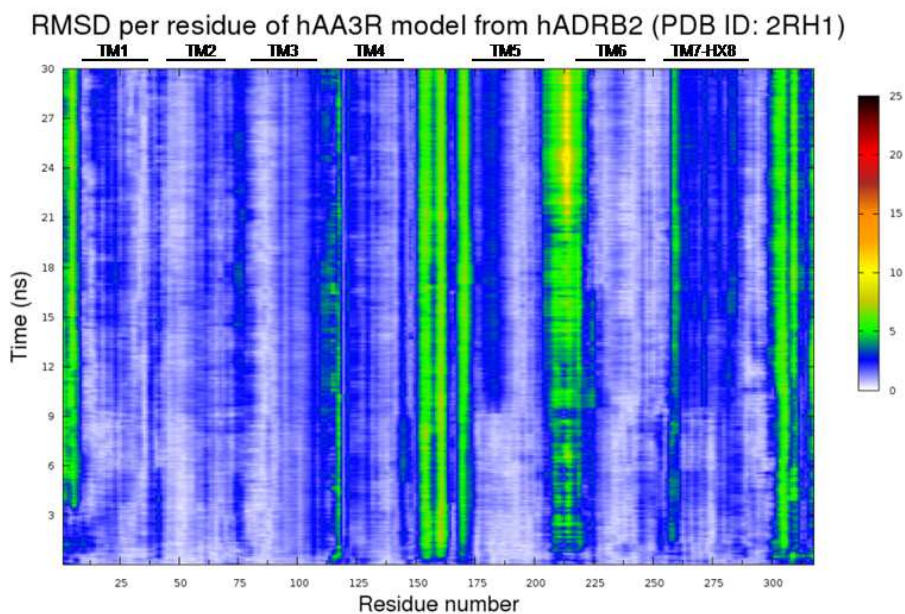
Figure 5.5: RMSD per residue of the backbone of the hA$_3$AR model built using β_2-adrenergic receptor as template.
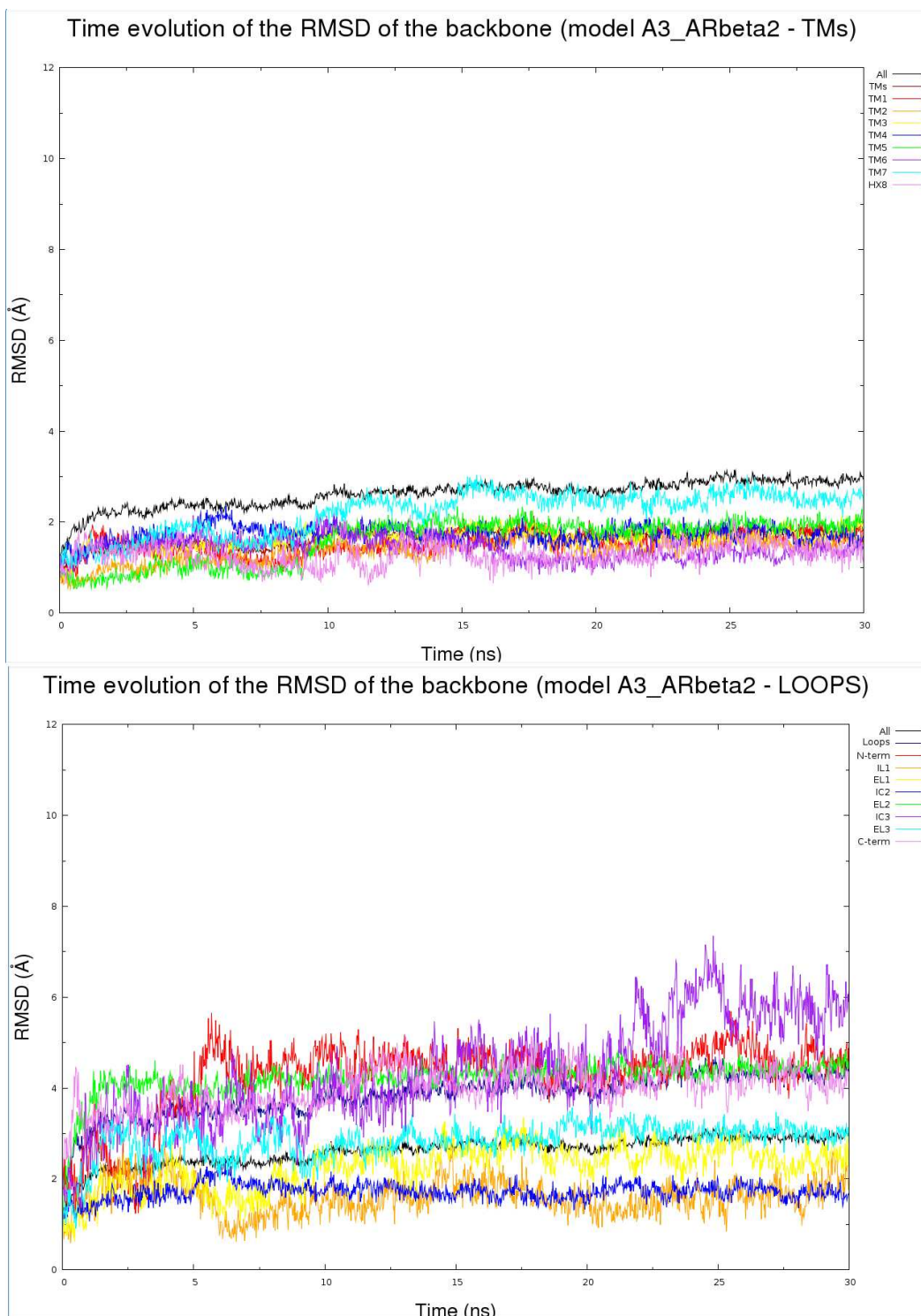
Figure 5.6: Time evolution of the RMSD of C α of the hA$_3$AR model built using β_2-adrenergic receptor as template. On the top, RMSD of the TM regions; on the bottom, rmsd of the loops, N-term and C-term.
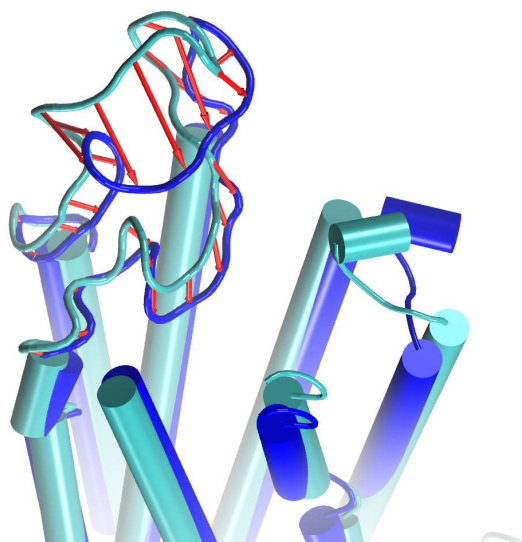
Figure 5.7: Representation of the second extracellular loop of the hA$_3$AR model built using hA$_{2A}$AR as template before (in cyan) and after (in blue) 30 ns of molecular dynamics in a lipid bilayer.
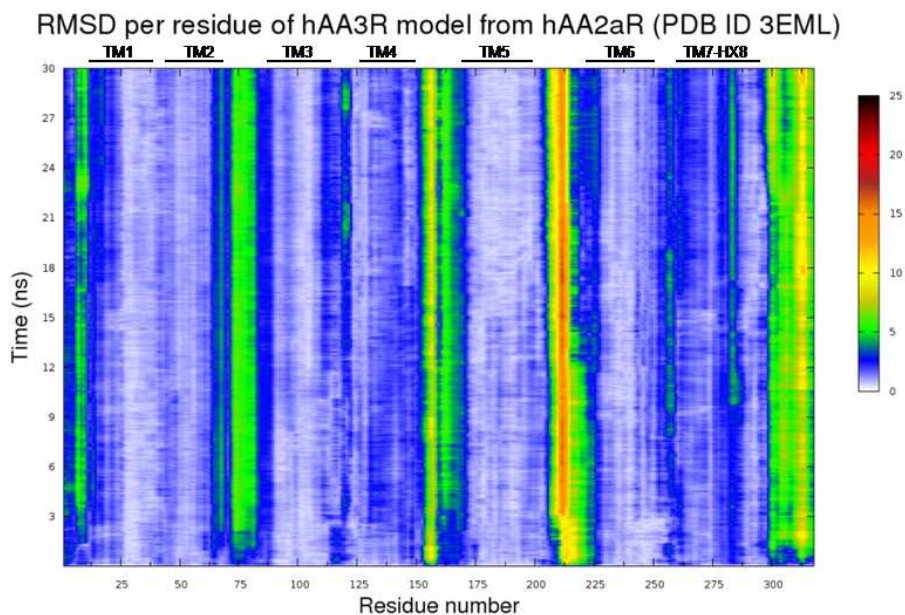
Figure 5.8: RMSD per residue of the backbone of the hA$_3$AR model built using hA$_3$ adenosine receptor as template.
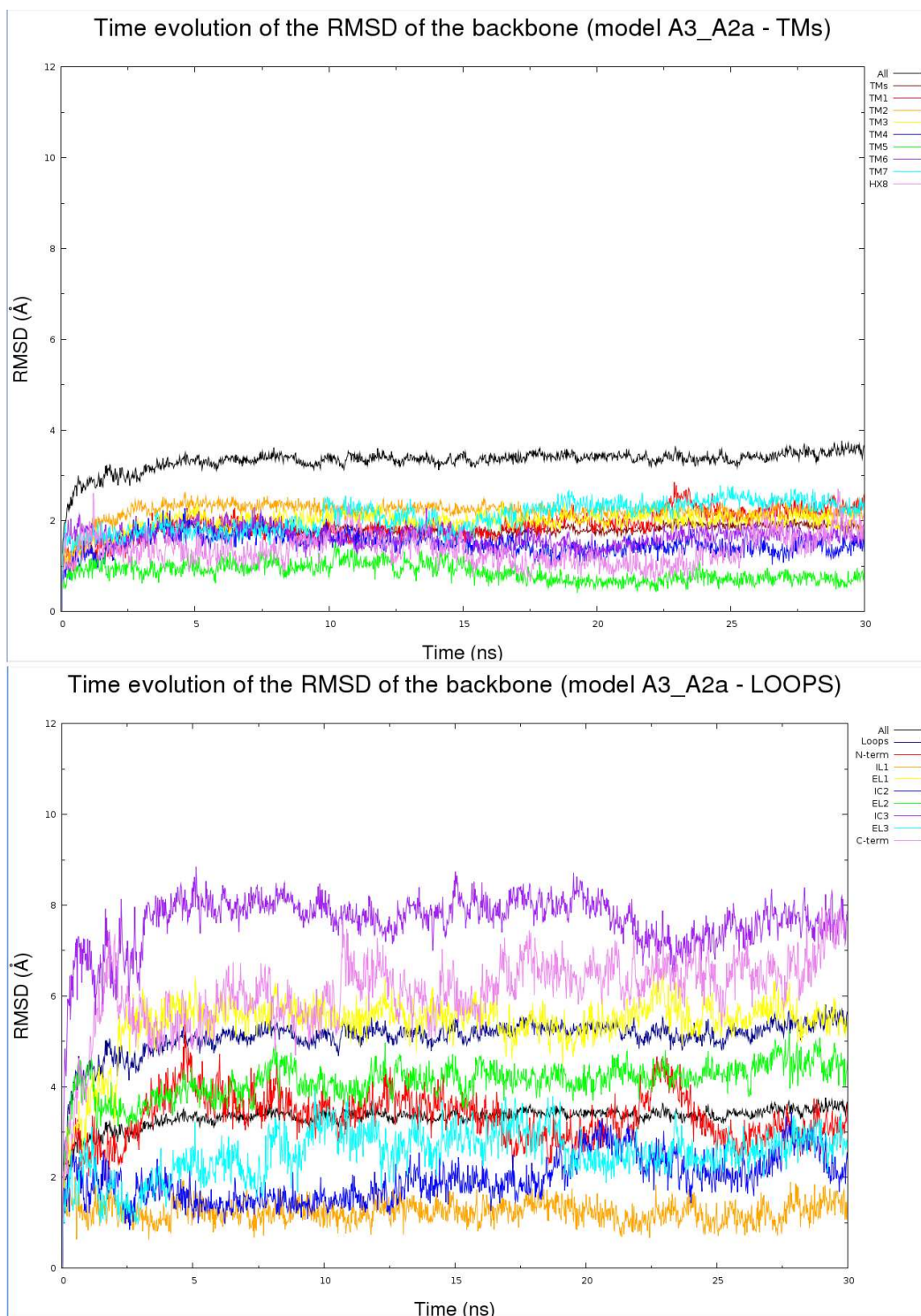
Figure 5.9: Time evolution of the RMSD of C α of the hA$_3$AR model built using hA$_3$adenosine receptor as template. On the top, RMSD of the TM regions; on the bottom, rmsd of the loops, N-term and C-term.

4-Amido-2-aryl-1,2,3-triazolo[4,3-a]quinoxalin-1-one Derivatives

Table A.1: Binding Activity at Human A_1, A_{2A}, A_3 and Bovine A_1, A_{2A} ARs.

	R_4	R_1	R_6	hA$_3^b$	hA$_1^c$	hA$_{2A}^d$	bA$_1^e$	bA$_{2A}^f$
				K_i^a(nM) or I%
A^g	CH$_3$	H	H	2.0 ± 0.11	2000 ± 140	22%	4.3 ± 0.38	70%
1	CH$_3$	OMe	H	35.7 ± 2.40	34%	6%	245 ± 23.1	0%
2	CH$_3$	H	NO$_2$	18%			6 ± 0.55	36%
3	CH$_3$	OMe	NO$_2$	36%			0%	7%
4	CH$_3$	H	NH$_2$	48 ± 2.10	32%	367 ± 24	1 ± 0.09	6250 ± 410
5	CH$_3$	OMe	NH$_2$	5.5 ± 0.23	2700 ± 150	1100 ± 10	363 ± 24	20%
B^g	Ph	H	H	1.47 ± 0.11	87.8 ± 6.30	88.2 ± 5.80	89.6 ± 7.20	53%
6	Ph	OMe	H	2.9 ± 0.30	37%	3585 ± 224	1010 ± 112	23%
7	Ph	NO$_2$	H	100 ± 9.60			55%	26%
8	Ph	H	NO$_2$	22 ± 2.60	15%	25%	32%	0%
9	Ph	OMe	NO$_2$	217 ± 20.40			35%	15%
10	Ph	H	NH$_2$	22 ± 1.70	98 ± 7.4	4850 ± 330	42 ± 3.1	27.8%
11	Ph	OMe	NH$_2$	1 ± 0.30	45%	24%	393 ± 27	16%
12	CHPh$_2$	OMe	H	44 ± 3.10	25%	27%	7.2 ± 0.41	28.5%
13	CHPh$_2$	NO$_2$	H	13%			30%	0%
14	CHPh$_2$	H	H	0.81 ± 0.03	18.8 ± 1.20	58%	10.2 ± 1.60	1160 ± 97.40
15	CHPh$_2$	H	NO$_2$	14.9 ± 1.10	12%	49%	3.9 ± 20.2	29.5%
16	CHPh$_2$	OMe	NO$_2$	0.8 ± 0.04	11%	2%	260 ± 11	0%
17	CHPh$_2$	H	NH$_2$	8.65 ± 0.61	2.5%	627 ± 34	1.6 ± 0.05	12%
18	CHPh$_2$	OMe	NH$_2$	2.58 ± 0.15	0%	31%	77.5 ± 0.52	0%
19		H	H	5.2 ± 0.31	1%	43%	30 ± 2.40	19%
20		OMe	H	3.29 ± 0.15	2%	26%	174.5 ± 11.40	6570 ± 460

Continued on next page

	R_4	R_1	R_6	$K_i{}^a$ (nM) or I%				
				$hA_3{}^b$	$hA_1{}^c$	$hA_{2A}{}^c$	$bA_1{}^e$	$bA_{2A}{}^f$
21	H	H	NO_2	27%			39%	0%
22	OMe	OMe	NO_2	343 ± 21.0			20%	0%
23	H	H	NH_2	1243 ± 115			79 ± 5.10	36%

[a] The K_i values are mean $\pm$ SEM of four separated assays, each performed in triplicate. [b] Displacement of specific [^{125}I]AB-MECA binding at human A_3 receptors expressed in CHO cells or percentage of inhibition (I) of specific binding at 1 μM concentration. [c] Displacement of specific [^{3}H]DPCPX binding at hA_1 receptors expressed in CHO cells or percentage of inhibition (I) of specific binding at 10 μM concentration. [d] Displacement of specific [^{3}H]NECA binding at hA_{2A} receptors expressed in CHO cells or percentage of inhibition (I) of specific binding at 10 μM concentration. [e] Displacement of specific [^{3}H]DPCPX binding in bovine brain membranes or percentage of inhibition (I) of specific binding at 10 μM concentration. [f] Displacement of specific [^{3}H]CGS binding from bovine striatal membranes or percentage of inhibition (I) of specific binding at 10 μM concentration. [g] bA_1, bA_{2A}, hA_3 AR binding data were reported in [127].

2-Arylpyrazolo[3,4-*c*]quinoline Derivatives

Table B.1: Binding Activity at Human A_1, A_{2A}, A_3ARs.
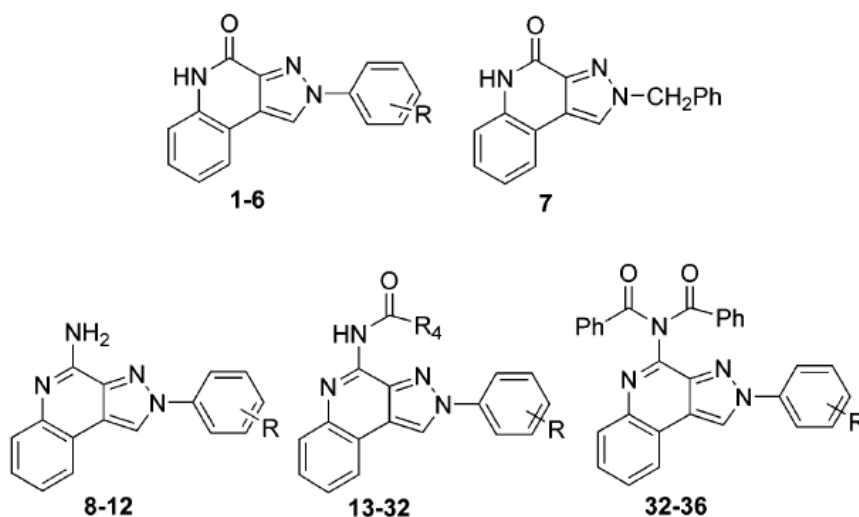
	R	R_4	$K_i{}^a$(nM) or I%		
			$hA_3{}^b$	$hA_1{}^c$	$hA_{2A}{}^c$
1[d]	H		30.8 ± 2.6	203 ± 12	43%
2[d]	3-Me		5.0 ± 0.4	12 ± 1	46%
3[d]	4-Me		3.2 ± 0.2	29 ± 0.5	44%
4[d]	4-OMe		3.2 ± 0.2	176.4 ± 8.8	25%
5	3-OMe		7.3 ± 0.1	14 ± 0.4	52%
6	4-NO_2		85.5 ± 4	357 ± 35	0%
7			74.5 ± 5.3	8%	32%
8[d]	H		551 ± 34	659 ± 43	91 ± 7.3
9[d]	3-Me		99.3 ± 7.8	21 ± 1.6	228 ± 12.3
10[d]	4-Me		188 ± 15	45 ± 3.4	329 ± 22
11[d]	4-OMe		90.2 ± 7.3	40 ± 3.1	1060 ± 96
12	3-OMe		228.5 ± 19	32 ± 3.0	486 ± 34
13[d]	H	Me	48.2 ± 3.5	0%	3%
14	3-Me	Me	31 ± 2.4	203 ± 15	10%
15	4-Me	Me	123 ± 10	455 ± 41	1500 ± 130
16	4-OMe	Me	101.5 ± 7.4	2875 ± 110	0%

	R	R_4	$K_i{}^a$(nM) or I%		
			$hA_3{}^b$	$hA_1{}^c$	$hA_{2A}{}^c$
17^d	H	Ph	2.1 ± 0.1	0%	9%
18	3-Me	Ph	4.3 ± 0.5	57 ± 4.2	2860 ± 224
19	4-Me	Ph	4.4 ± 0.2	629 ± 51	26%
20	4-OMe	Ph	3.4 ± 0.2	250 ± 13	39%
21^d	H	CH_2Ph	9.9 ± 0.8	5%	15%
22	3-Me	CH_2Ph	3.9 ± 0.3	60 ± 4.5	24%
23	4-Me	CH_2Ph	5.6 ± 0.4	55%	21%
24	4-OMe	CH_2Ph	4.5 ± 0.6	201 ± 12	51%
25	H	$CHPh_2$	9.9 ± 0.8	5%	15%
26	3-Me	$CHPh_2$	3.9 ± 0.3	60 ± 4.5	24%
27	4-Me	$CHPh_2$	5.6 ± 0.4	55%	21%
28	4-OMe	$CHPh_2$	4.5 ± 0.6	201 ± 12	51%
29^d	H	$NHCH_2Ph$	8.3 ± 0.7	0%	3%
30	3-Me	$NHCH_2Ph$	3.35 ± 0.2	6800 ± 510	20%
31	4-Me	$NHCH_2Ph$	257 ± 21	5%	39%
32	4-OMe	$NHCH_2Ph$	40%	43%	0%
33	H		6.1 ± 0.5	0%	0%
34	3-Me		23.25 ± 2.1	42%	20%
35	4-Me		30 ± 2.3	32%	0%
36	4-OMe		17.2 ± 1.4	25%	7%

^{a}The K_i values are mean $\pm$ SEM of four separated assays, each performed in triplicate. ^{b}Displacement of specific [^{125}I]AB-MECA binding at human A_3 receptors expressed in CHO cells or percentage of inhibition (*I%*) of specific binding at 1 μM concentration. ^{c}Displacement of specific [^{3}H]DPCPX and [^{3}H]NECA binding at, respectively, hA_1 and hA_{2A} receptors expressed in CHO cells or percentage of inhibition (*I%*) of specific binding at 10 μM concentration. ^{d}The hA_3 AR binding affinity was reported in [113].

4-modified-2-aryl-1,2,4-triazolo[4,3-a]quinoxalin-1-one Derivatives

Table C.1: Binding Affinity at Human A_1, A_{2A}, A_3 and Bovine A_1, A_{2A} ARs.

A, 1-21

	R_4	R_1	R_6	$K_i{}^a$ (nM) or I%				
				$hA_3{}^b$	$hA_1{}^c$	$hA_{2A}{}^d$	$bA_1{}^e$	$bA_{2A}{}^f$
A^g	NHCOPh	H	H	1.47 ± 0.06	87.8 ± 6.3	89.6 ±6.7	89.6 ± 7.2	53%
1	NHCOC$_6$H$_4$-4COOMe	H	H	41%			106 ± 2.1	36%
2	NHCOC$_6$H$_4$-4COOMe	OMe	H	1370 ± 121			30.5%	41%
3	NHCOC$_6$H$_4$-3I	H	H	36%			473 ± 34	35%
4	NHCO-4-Pyridyl	H	H	6.1 ± 0.5	2379 ± 191	188 ± 9.4	57 ± 4.3	812 ± 71
5	NHCO-4-Pyridyl	OMe	H	68 ± 5.2	779 ± 53	397 ± 39	236 ± 15	44%
6	NHCO-4-Pyridyl	H	NO$_2$	0%			37.5%	22%
7	NHSO$_2$Ph	H	H	32.2 ± 2.8	0%	27%	157 ± 1.4	35%
8	NHSO$_2$Ph	OMe	H	2.2 ± 0.11	2700 ± 142	23%	4700 ± 260	16%
9	NHSO$_2$Ph	H	NO$_2$	100 ± 7.2			210 ± 12	25%
10	NHSO$_2$CH$_3$	H	H	1427 ± 125			164 ± 11.3	32%
11	NHSO$_2$CH$_3$	OMe	H	493 ± 33			6%	0%
12	NHSO$_2$CH$_3$	H	NO$_2$	37%			36 ± 1.3	56%
13	N(SO$_2$CH$_3$)$_2$	H	H	5.5 ± 0.4	36%	32%	36± 1.3	56%
14	N(SO$_2$CH$_3$)$_2$	H	H	387 ± 24			6.2%	17%
15	NHCONHCH$_2$Ph	H	H	83.5 ± 4.9	12.3 ± 1.2	158.3 ± 15	4.1 ± 0.2	172.6 ± 12
16	NHCONHCH$_2$Ph	OMe	H	65 ± 5.1	4215 ± 350	23%	20.8 ± 1.2	12%
17	NHCONHCH$_2$Ph	H	NO$_2$	63 ± 4.4	4%	20%	4.6 ± 0.3	46.5%
18	NHCONHCOPh	H	H	1300 ± 115			100.6 ± 8.9	379 ± 24
19	NHCONH-Ph-3I	H	H	953 ± 61			359 ± 25	1800 ± 150
20	OCH$_2$Ph	H	H	21 ± 1.8	46%	10%	55 ± 3.6	19%
21	OCH$_2$Ph	OMe	H	6.4 ± 0.4	54%	4%	53%	41%

^{a}The K_i values are mean ± SEM of four separated assays, each performed in triplicate. ^{b}Displacement of specific [^{125}I]AB-MECA binding at human A_3 receptors expressed in CHO cells or percentage of inhibition (I) of specific binding at 1 μM concentration. ^{c}Displacement of specific [^{3}H]DPCPX binding at hA_1 receptors expressed in CHO cells or percentage of inhibition (I) of specific binding at 10 μM concentration. ^{d}Displacement of specific [^{3}H]NECA binding at hA_{2A} receptors expressed in CHO cells or percentage of inhibition (I) of specific binding at 10 μM concentration. ^{e}Displacement of specific [^{3}H]DPCPX binding in bovine brain membranes or percentage of inhibition (I) of specific binding at 10 μM concentration. ^{f}Displacement of specific

[^{3}H]CGS binding from bovine striatal membranes or percentage of inhibition (I) of specific binding at 10 μM concentration. ^{g}bA$_1$, bA$_{2A}$, hA$_3$ AR binding data were reported in [82].

Pyrido[2,3-e]-1,2,4-triazolo[4,3-a]pyrazin-1-one Derivatives

Table D.1: Binding Affinity at Human A_1, A_{2A}, A_3 and Bovine A_1, A_{2A} ARs.

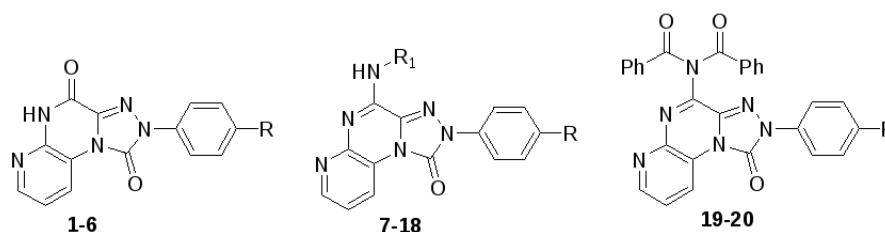|	R_1	R	$hA_3{}^b$	$bA_1{}^c$	$bA_{2A}{}^d$	$hA_1{}^e$	$hA_{2A}{}^e$
				$K_i{}^a$(nM) or I%			
1		H	251 ± 16	145 ± 11 12%			
2		4-OMe	3.3 ± 0.2 26%	0%	114 ± 8	0%	
3		4-OH	32%	449 ± 25	0%		
4		4-F	590 ± 42	305.5 ± 25	26%		
5		4-COOEt	0%	16%	0%		
6		4-COOH	0%	30%	7%		
7	H	H 656 ± 41	3.1 ± 0.28	92.6 ± 5.6			
8	H	4-OMe	158 ± 9.8	1102 ± 81	413 ± 34		
9	H	4-OH	1335 ± 112	112 ± 8.1	832 ± 62		
10	H	4-F	490 ± 36	181 ± 15	1508 ± 130		
11	H	4-COOEt	0%	39%	17%		
12	C_6H_11	H	15.5 ± 1.2	0.38 ± 0.029	199 ± 13	37%	211 ± 8.4
13	C_5H_9	H	8.4 ± 0.9	0.47 ± 0.047	510 ± 36	36%	208 ± 10
14	COMe	H	138 ± 12	14 ± 1.1	59%		
15	COPh	H	70.3 ± 6	152 ± 10	7100 ± 550	8%	
16	$COCH_2Ph$	H	11.7 ± 1	7.15 ± 0.5	414 ± 32	37%	208 ± 6.2
17	COMe	4-OMe	41 ± 3.2	56%	19%	48%	29%
18	COPh	4-OMe	4.54 ± 0.2	355 ± 22	7%	38%	27%
19		H	335 ± 28	70.7 ± 6.5	12%		
20		4-OMe	7.75 ± 0.8	17%	0%	0%	0%

[a] The K_i values are mean $\pm$ SEM of four separated assays, each performed in triplicate. [b] Displacement of specific [^{125}I]AB-MECA binding at human A_3 receptors expressed in CHO cells or percentage of inhibition (I%) of specific binding at 1 μM concentration. [c] Displacement of specific

[^{3}H]DPCPX binding in bovine brain membranes or percentage of inhibition (I) of specific binding at 10 μM concentration. ^{d}Displacement of specific [^{3}H]CGS binding at bovine striatal membranes or percentage of inhibition ($I\%$) of specific binding at 10 μM concentration. ^{e}Displacement of specific [^{3}H]DPCPX and [^{3}H]NECA binding at, respectively, hA$_1$ and hA$_{2A}$ receptors expressed in CHO cells or percentage of inhibition ($I\%$) of specific binding at 10 μM concentration.

Table D.2: Comparison between the hA$_3$ AR affinities of the Pyridotriazolopyrazin-1-ones (X= N) and the corresponding 6-Nitro-triazoloquinoxalin-1-ones (X= C-NO$_2$).

X= N   **1, 2**
X= C-NO$_2$ **33, 34**

X= N   **7, 10, 14, 15, 17-20**
X= C-NO$_2$ **35-44**

			X= N^a		X= C-NO$_2$^{b}
R	R$_4$		K$_i$(nM) hA$_3$ or I% (1 μM)
H		**1**	251 ± 16	**33**	279 ± 16
OMe		**2**	3.3 ± 0.2	**34**	4.7 ± 0.52
H	H	**7**	656 ± 41	**35**	4.75 ± 0.3
OMe	H	**10**	58 ± 9.8	**36**	47 ± 1.2
H	NHC$_6$H$_1$1	**12**	15.5 ± 1.2	**37**	281 ± 24
H	NHC$_5$H$_9$	**13**	8.4 ± 0.9	**38**	116 ± 24
H	NHCOMe	**14**	138 ± 12	**39**	18%
H	NHCOPh	**15**	70.3 ± 6	**40**	22 ± 2.60
OMe	NHCOMe	**17**	41 ± 3.6	**41**	36%
OMe	NHCOPh	**18**	4.54 ± 0.2	**42**	217 ± 20
H	N(COPh)$_2$	**19**	335 ± 22	**43**	27%
OMe	N(COPh)$_2$	**20**	7.75 ± 0.8	**44**	343 ± 21

^{a}Data from previous table. ^{b}Data from AppendixA.

N-5 Substituted Pyrazolo-triazolo-pyrimidine Derivatives

Table E.1: Biological profile of synthesized (**4,5**) and reference (**2,3**) compounds at Human A_1, A_{2A}, A_3 and Bovine A_1, A_{2A} ARs.
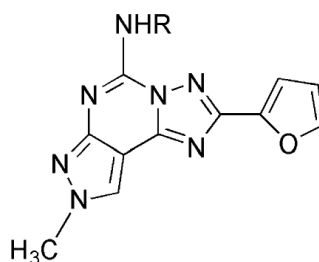
	R	$hA_1{}^a$(K_i nM)	$hA_{2A}{}^b$(K_i nM)	$hA_{2B}{}^c$(IC_{50} nM)	$hA_3{}^d$(K_i nM)
2	CONHPh	310(295-327)	27.7(13.3-57.8)	3440(2880-4110)	1.80(0.88-3.68)
3	COCH−2Ph	1040(864-1260)	282(201-375)	12320(9730-16400)	0.92(0.80-1.06)
4	COPh	2030(1710-2400)	879(643-1200)	>30000	15.7(7.85-31.5)
5	SO−2Ph	20700(16700-25700)	6060(5170-7110)	>30000	744(534-1040)

Data are expressed as geometric means, with 95% confidence limits ^{a}Displacement of specific [^{3}H]-CCPA binding at human hA_1 receptors expressed in CHO cells, ($n = 3 - 6$) ^{b}Displacement of specific [^{3}H]-NECA binding at human hA_{2A} receptors expressed in CHO cells ^{c}IC_{50} values of the inhibition of NECA-stimulated adenylyl cyclase activity in CHO cells expressing hA_{2A} receptors ^{d}Displacement of specific [^{3}H]-NECA binding at human hA_3 receptors expressed in CHO cells

Quinazoline, Quinoline and Pyrimidine Derivatives

Table F.1: Binding Affinity at hA_1, hA_{2A}, hA_3 and Potency (IC_{50}) at hA_{2B} and hA_3 ARs.
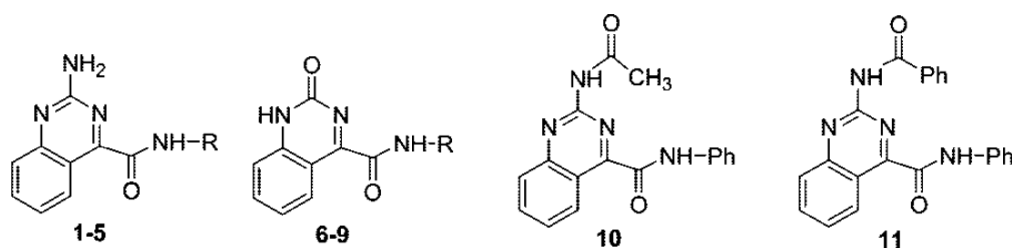
| | R | K_i(nM) or I% | | | IC_{50}(nM) or I% cAMP | |
		$hA_3{}^a$	$hA_1{}^b$	$hA_{2A}{}^c$	$hA_{2B}{}^d$	$hA_3{}^e$
1	C_6H_4-4-OMe	87.5 ± 6.6	8%	6%	23%	
2	C_6H_5	350 ± 40	40%	17%	5%	
3	C_6H_4-4-Me	98.3 ± 7.3	3%	5%	4%	
4	C_6H_4-4-Br	550 ± 47	1%	1%	2%	
5	C_6H_{11}	21%	2%	3%	1%	
6	C_6H_4-4-OMe	19.5 ± 2.2	4%	1%	9%	125 ± 10
7	C_6H_5	50 ± 4	22%	1%	4%	238 ± 21
8	C_6H_4-4-Me	26.7 ± 3.3	21%	2%	2%	
9	C_6H_4-4-Br	27.2 ± 3.1	3%	1%	2%	
10		25.3 ± 2.8	25%	7%	5%	140 ± 13
11		182 ± 10	7%	10%	3%	

[a] Displacement of specific [^{125}I]AB-MECA binding to hA_3 CHO cells. K_i values are means $\pm$ SEM of four separate assays, each performed in duplicate. [b] Percentage of inhibition in [^{3}H]DPCPX competition binding assays to hA_1 CHO cells at 1 μM concentration of the tested compounds. [c] Percentage of inhibition in [^{3}H]ZM241385 competition binding assays to hA_{2A} CHO cells at 1 μM concentration of the tested compounds. [d] Percentage of inhibition on cAMP experiments in hA_2B CHO cells, stimulated by 200 nM NECA, at 1 μM concentration of the examined compounds. [e] IC_50 values are expressed as means $\pm$ SEM of four separate cAMP experiments in hA_3 CHO cells, inhibited by 100 nM Cl-IB-MECA.

Table F.2: Inibition of Specific Binding at hA_1, hA_{2A}, hA_3 AR
and of cAMP Production at hA_{2B} and hA_3 ARs.

| | R_1 | R_2 | binding experiments | | | cAMP assays |
			$hA_3{}^a$	$hA_1{}^b$	$hA_{2A}{}^c$	$hA_{2B}{}^d$
12	OMe		1%	2%	3%	3%
13	H		6%	1%	2%	5%
14	OMe		26%	4%	1%	3%
15	H		38%	1%	1%	2%
16		H	15%	3%	6%	3%
17		COPh	22%	6%	6%	2%
18			14%	5%	1%	3%

$_a$ Percentage of inhibition in [^{1}25I]AB-MECA competition binding assays to hA_3 CHO cells at 1 μM concentration of the tested compounds. $_b$ Percentage of inhibition in [^{3}H]-DPCPX competition binding assays to hA_1 CHO cells at 1 μM concentration of the tested compounds. ^{c} Percentage of inhibition in [^{3}H]-ZM 241385 competition binding assays to hA_2ACHO cells at 1 μM concentration of the tested compounds. ^{d} Percentage of inhibition on cAMP experiments in hA_2B CHO cells, stimulated by 200 nM NECA, at 1 μM concentration of the tested compounds.

Bibliography

[1] K.L. Pierce, R.T. Premont, and R.J. Lefkowitz. Seven-transmembrane receptors. *Nat. Rev. Mol. Cell Biol.*, 3:639–650, Sep 2002.

[2] A.E. Brady and L.E. Limbird. G protein-coupled receptor interacting proteins: emerging roles in localization and signal transduction. *Cell. Signal.*, 14:297–309, Apr 2002.

[3] R.J. Lefkowitz. Historical review: a brief history and personal retrospective of seven-transmembrane receptors. *Trends Pharmacol. Sci.*, 25:413–422, Aug 2004.

[4] J. Bockaert and J.P. Pin. Molecular tinkering of G protein-coupled receptors: an evolutionary success. *EMBO J.*, 18:1723–1729, Apr 1999.

[5] G.R. Post and J.H. Brown. G protein-coupled receptors and signaling pathways regulating growth responses. *FASEB J.*, 10:741–749, May 1996.

[6] T.B. Patel. Single transmembrane spanning heterotrimeric g protein-coupled receptors and their signaling cascades. *Pharmacol. Rev.*, 56:371–385, Sep 2004.

[7] P.H. McDonald and R.J. Lefkowitz. Beta-Arrestins: new roles in regulating heptahelical receptors' functions. *Cell. Signal.*, 13:683–689, Oct 2001.

[8] R.J. Lefkowitz and S.K. Shenoy. Transduction of receptor signals by beta-arrestins. *Science*, 308:512–517, Apr 2005.

[9] B. Brone and J. Eggermont. PDZ proteins retain and regulate membrane transporters in polarized epithelial cell membranes. *Am. J. Physiol., Cell Physiol.*, 288:C20–29, Jan 2005.

[10] T. Metaye, H. Gibelin, R. Perdrisot, and J.L. Kraimps. Pathophysiological roles of G-protein-coupled receptor kinases. *Cell. Signal.*, 17:917–928, Aug 2005.

[11] G. Milligan. A day in the life of a G protein-coupled receptor: the contribution to function of G protein-coupled receptor dimerization. *Br. J. Pharmacol.*, 153 Suppl 1:S216–229, Mar 2008.

[12] S.C. Prinster, C. Hague, and R.A. Hall. Heterodimerization of g protein-coupled receptors: specificity and functional significance. *Pharmacol. Rev.*, 57:289–298, Sep 2005.

[13] G. Milligan. G-protein-coupled receptor heterodimers: pharmacology, function and relevance to drug discovery. *Drug Discov. Today*, 11:541–549, Jun 2006.

[14] S. Takeda, S. Kadowaki, T. Haga, H. Takaesu, and S. Mitaku. Identification of G protein-coupled receptor genes from the human genome sequence. *FEBS Lett.*, 520:97–101, Jun 2002.

[15] M. Eilers, V. Hornak, S.O. Smith, and J.B. Konopka. Comparison of class A and D G protein-coupled receptors: common features in structure and activation. *Biochemistry*, 44:8959–8975, Jun 2005.

[16] R. Fredriksson, M.C. Lagerstrom, L.G. Lundin, and H.B. Schioth. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.*, 63:1256–1272, Jun 2003.

[17] L.F. Kolakowski. GCRDb: a G-protein-coupled receptor database. *Recept. Channels*, 2:1–7, 1994.

[18] D.T. Chalmers and D.P. Behan. The use of constitutively active GPCRs in drug discovery and functional genomics. *Nat Rev Drug Discov*, 1:599–608, Aug 2002.

[19] L. Arvanitakis, E. Geras-Raaka, and M.C. Gershengorn. Constitutively signaling g-protein-coupled receptors and human disease. *Trends Endocrinol. Metab.*, 9:27–31, 1998.

[20] A. Shenker. G protein-coupled receptor structure and function: the impact of disease-causing mutations. *Baillieres Clin. Endocrinol. Metab.*, 9:427–451, Jul 1995.

[21] B.K. Rana, T. Shiina, and P.A. Insel. Genetic variations and polymorphisms of G protein-coupled receptors: functional and therapeutic implications. *Annu. Rev. Pharmacol. Toxicol.*, 41:593–624, 2001.

[22] M.D. Thompson, W.M. Burnham, and D.E. Cole. The G protein-coupled receptors: pharmacogenetics and disease. *Crit Rev Clin Lab Sci*, 42:311–392, 2005.

[23] T. Klabunde and G. Hessler. Drug design strategies for targeting G-protein-coupled receptors. *Chembiochem*, 3:928–944, Oct 2002.

[24] D.P. Behan and D.T. Chalmers. The use of constitutively active receptors for drug discovery at the G protein-coupled receptor gene pool. *Curr Opin Drug Discov Devel*, 4:548–560, Sep 2001.

[25] J.A. Ballesteros and H. Weinstein. 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.

[26] D. Mustafi and K. Palczewski. Topology of class A G protein-coupled receptors: insights gained from crystal structures of rhodopsins, adrenergic and adenosine receptors. *Mol. Pharmacol.*, 75:1–12, Jan 2009.

[27] Helen Berman, Kim Henrick, and Haruki Nakamura. Announcing the worldwide Protein Data Bank. *Nat Struct Biol*, 10(12), Dec 2003.

[28] K. Palczewski, T. Kumasaka, T. Hori, C.A. Behnke, H. Motoshima, B.A. Fox, I. Le Trong, D.C. Teller, T. Okada, R.E. Stenkamp, M. Yamamoto, and M. Miyano. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science*, 289:739–745, Aug 2000.

[29] D.C. Teller, T. Okada, C.A. Behnke, K. Palczewski, and R.E. Stenkamp. Advances in determination of a high-resolution three-dimensional structure of rhodopsin, a model of G-protein-coupled receptors (GPCRs). *Biochemistry*, 40:7761–7772, Jul 2001.

[30] T. Okada, Y. Fujiyoshi, M. Silow, J. Navarro, E.M. Landau, and Y. Shichida. Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. *Proc. Natl. Acad. Sci. U.S.A.*, 99:5982–5987, Apr 2002.

[31] J. Li, P.C. Edwards, M. Burghammer, C. Villa, and G.F. Schertler. Structure of bovine rhodopsin in a trigonal crystal form. *J. Mol. Biol.*, 343:1409–1438, Nov 2004.

[32] T. Okada, M. Sugihara, A.N. Bondar, M. Elstner, P. Entel, and V. Buss. The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure. *J. Mol. Biol.*, 342:571–583, Sep 2004.

[33] H. Nakamichi and T. Okada. Local peptide movement in the photoreaction intermediate of rhodopsin. *Proc. Natl. Acad. Sci. U.S.A.*, 103:12729–12734, Aug 2006.

[34] H. Nakamichi and T. Okada. Crystallographic analysis of primary visual photochemistry. *Angew. Chem. Int. Ed. Engl.*, 45:4270–4273, Jun 2006.

[35] D. Salom, D.T. Lodowski, R.E. Stenkamp, I. Le Trong, M. Golczak, B. Jastrzebska, T. Harris, J.A. Ballesteros, and K. Palczewski. Crystal structure of a photoactivated deprotonated intermediate of rhodopsin. *Proc. Natl. Acad. Sci. U.S.A.*, 103:16123–16128, Oct 2006.

[36] J. Standfuss, G. Xie, P.C. Edwards, M. Burghammer, D.D. Oprian, and G.F. Schertler. Crystal structure of a thermally stable rhodopsin mutant. *J. Mol. Biol.*, 372:1179–1188, Oct 2007.

[37] H. Nakamichi, V. Buss, and T. Okada. Photoisomerization mechanism of rhodopsin and 9-cis-rhodopsin revealed by x-ray crystallography. *Biophys. J.*, 92:L106–108, Jun 2007.

[38] V. Cherezov, D.M. Rosenbaum, M.A. Hanson, S.G. Rasmussen, F.S. Thian, T.S. Kobilka, H.J. Choi, P. Kuhn, W.I. Weis, B.K. Kobilka, and R.C. Stevens. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science*, 318:1258–1265, Nov 2007.

[39] S.G. Rasmussen, H.J. Choi, D.M. Rosenbaum, T.S. Kobilka, F.S. Thian, P.C. Edwards, M. Burghammer, V.R. Ratnala, R. Sanishvili, R.F. Fischetti, G.F. Schertler, W.I. Weis, and B.K. Kobilka. Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature*, 450:383–387, Nov 2007.

[40] T. Shimamura, K. Hiraki, N. Takahashi, T. Hori, H. Ago, K. Masuda, K. Takio, M. Ishiguro, and M. Miyano. Crystal structure of squid rhodopsin with intracellularly extended cytoplasmic region. *J. Biol. Chem.*, 283:17753–17756, Jun 2008.

[41] M.A. Hanson, V. Cherezov, M.T. Griffith, C.B. Roth, V.P. Jaakola, E.Y. Chien, J. Velasquez, P. Kuhn, and R.C. Stevens. A specific cholesterol binding site is established by the 2.8 A structure of the human beta2-adrenergic receptor. *Structure*, 16:897–905, Jun 2008.

[42] J.H. Park, P. Scheerer, K.P. Hofmann, H.W. Choe, and O.P. Ernst. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature*, 454:183–187, Jul 2008.

[43] T. Warne, M.J. Serrano-Vega, J.G. Baker, R. Moukhametzianov, P.C. Edwards, R. Henderson, A.G. Leslie, C.G. Tate, and G.F. Schertler. Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature*, 454:486–491, Jul 2008.

[44] P. Scheerer, J. H. Park, P. W. Hildebrand, Y. J. Kim, N. Krauss, H. W. Choe, K. P. Hofmann, and O. P. Ernst. Crystal structure of opsin in its G-protein-interacting conformation. *Nature*, 455:497–502, Sep 2008.

[45] V.P. Jaakola, M.T. Griffith, M.A. Hanson, V. Cherezov, E.Y. Chien, J.R. Lane, A.P. Ijzerman, and R.C. Stevens. The 2.6 Angstrom Crystal Structure of a Human A2A Adenosine Receptor Bound to an Antagonist. *Science*, Oct 2008.

[46] F. Fanelli and P. G. De Benedetti. Inactive and active states and supramolecular organization of GPCRs: insights from computational modeling. *J. Comput. Aided Mol. Des.*, 20:449–461, 2006.

[47] M. Murakami and T. Kouyama. Crystal structure of squid rhodopsin. *Nature*, 453:363–367, May 2008.

[48] D. M. Rosenbaum, V. Cherezov, M. A. Hanson, S. G. Rasmussen, F. S. Thian, T. S. Kobilka, H. J. Choi, X. J. Yao, W. I. Weis, R. C. Stevens, and B. K. Kobilka. GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science*, 318:1266–1273, Nov 2007.

[49] B. B. Fredholm, A. P. IJzerman, K. A. Jacobson, K. N. Klotz, and J. Linden. International Union of Pharmacology. XXV. Nomenclature and classification of adenosine receptors. *Pharmacol. Rev.*, 53:527–552, Dec 2001.

[50] P. Fishman and S. Bar-Yehuda. Pharmacology and therapeutic applications of A3 receptor subtype. *Curr Top Med Chem*, 3:463–469, 2003.

[51] M. Parsons, L. Young, J. E. Lee, K. A. Jacobson, and B. T. Liang. Distinct cardioprotective effects of adenosine mediated by differential coupling of receptor subtypes to phospholipases C and D. *FASEB J.*, 14:1423–1431, Jul 2000.

[52] S. Merighi, A. Benini, P. Mirandola, S. Gessi, K. Varani, E. Leung, S. Maclennan, and P. A. Borea. A3 adenosine receptor activation inhibits cell proliferation via phosphatidylinositol 3-kinase/Akt-dependent inhibition of the extracellular signal-regulated kinase 1/2 phosphorylation in A375 human melanoma cells. *J. Biol. Chem.*, 280:19516–19526, May 2005.

[53] T. M. Palmer and G. L. Stiles. Identification of threonine residues controlling the agonist-dependent phosphorylation and desensitization of the rat A(3) adenosine receptor. *Mol. Pharmacol.*, 57:539–545, Mar 2000.

[54] S. Rorke and S. T. Holgate. Targeting adenosine receptors: novel therapeutic targets in asthma and chronic obstructive pulmonary disease. *Am J Respir Med*, 1:99–105, 2002.

[55] K. A. Jacobson. Adenosine A3 receptors: novel ligands and paradoxical effects. *Trends Pharmacol. Sci.*, 19:184–191, May 1998.

[56] T. W. Stone. Purines and neuroprotection. *Adv. Exp. Med. Biol.*, 513:249–280, 2002.

[57] J. Linden. Molecular approach to adenosine receptors: receptor-mediated mechanisms of tissue protection. *Annu. Rev. Pharmacol. Toxicol.*, 41:775–787, 2001.

[58] UniProt Consortium. The Universal Protein Resource (UniProt) 2009. *Nucleic Acids Res.*, Oct 2008.

[59] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne. The Protein Data Bank. *Nucleic Acids Res.*, 28:235–242, Jan 2000.

[60] A.R. Leach. *Molecular Modelling - Principles and Applications.* Pearson Education Limited, 2nd edition, 2001.

[61] H.-D. Holtje, W. Sippl, D. Rognan, and G. Folkers. *Molecular Modeling - Basic Principles and Applications.* Wiley-VCH, 3nd edition, 2008.

[62] G. Jones, P. Willett, R.C. Glen, A.R. Leach, and R. Taylor. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.*, 267:727–748, Apr 1997.

[63] MOE, Chemical Computing Group Inc, 1010 Sherbrooke St. West Suite 910 Montreal, Quebec H3A 2R7 Canada.

[64] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, and A. J. Olson. Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function. *J. Computational Chemistry*, 19:1639–1662, 1998.

[65] Glide, version 3.5, Schrödinger, LLC, New York, NY, 2005.

[66] O. Korb, T. Stutzle, and T.E. Exner. An Ant Colony Optimization Approach to Flexible Protein-Ligand Docking. *Swarm Intelligence*, 1(2):115–134, 2007.

[67] T. Okada, I. Le Trong, B.A. Fox, C.A. Behnke, R.E. Stenkamp, and K. Palczewski. X-Ray diffraction analysis of three-dimensional crystals of bovine rhodopsin obtained from mixed micelles. *J. Struct. Biol.*, 130:73–80, May 2000.

[68] B.K. Kobilka and X. Deupi. Conformational complexity of G-protein-coupled receptors. *Trends Pharmacol. Sci.*, 28:397–406, Aug 2007.

[69] F.G. Sajjadi and G.S. Firestein. cDNA cloning and sequence analysis of the human A3 adenosine receptor. *Biochim. Biophys. Acta*, 1179:105–107, Oct 1993.

[70] C.A. Salvatore, M.A. Jacobson, H.E. Taylor, J. Linden, and R.G. Johnson. Molecular cloning and characterization of the human A3 adenosine receptor. *Proc. Natl. Acad. Sci. U.S.A.*, 90:10365–10369, Nov 1993.

[71] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, 117:5179–5197, 1995.

[72] A. Martinelli and T. Tuccinardi. Molecular modeling of adenosine receptors: new results and trends. *Med Res Rev*, 28:247–277, Mar 2008.

[73] F. Libert, J. Van Sande, A. Lefort, A. Czernilofsky, J. E. Dumont, G. Vassart, H. A. Ensinger, and K. D. Mendla. Cloning and functional characterization of a human A1 adenosine receptor. *Biochem. Biophys. Res. Commun.*, 187:919–926, Sep 1992.

[74] K. D. Pierce, T. J. Furlong, L. A. Selbie, and J. Shine. Molecular cloning and expression of an adenosine A2b receptor from human brain. *Biochem. Biophys. Res. Commun.*, 187:86–93, Aug 1992.

[75] T. J. Furlong, K. D. Pierce, L. A. Selbie, and J. Shine. Molecular characterization of a human brain adenosine A2 receptor. *Brain Res. Mol. Brain Res.*, 15:62–66, Sep 1992.

[76] Z. G. Gao, S. K. Kim, A. S. Gross, A. Chen, J. B. Blaustein, and K. A. Jacobson. Identification of essential residues involved in the allosteric modulation of the human A(3) adenosine receptor. *Mol. Pharmacol.*, 63:1021–1031, May 2003.

[77] H. T. Duong, Z. G. Gao, and K. A. Jacobson. Nucleoside modification and concerted mutagenesis of the human A3 adenosine receptor to probe interactions between the 2-position of adenosine analogs and Gln167 in the second extracellular loop. *Nucleosides Nucleotides Nucleic Acids*, 24:1507–1517, 2005.

[78] A. Chen, Z. G. Gao, D. Barak, B. T. Liang, and K. A. Jacobson. Constitutive activation of A(3) adenosine receptors by site-directed mutagenesis. *Biochem. Biophys. Res. Commun.*, 284:596–601, Jun 2001.

[79] Z. G. Gao, A. Chen, D. Barak, S. K. Kim, C. E. MÃ¼ller, and K. A. Jacobson. Identification by site-directed mutagenesis of residues involved in ligand recognition and activation of the human A3 adenosine receptor. *J. Biol. Chem.*, 277:19056–19063, May 2002.

[80] Z. G. Gao, S. K. Kim, T. Biadatti, W. Chen, K. Lee, D. Barak, S. G. Kim, C. R. Johnson, and K. A. Jacobson. Structural determinants of A(3) adenosine receptor activation: nucleoside ligands at the agonist/antagonist boundary. *J. Med. Chem.*, 45:4471–4484, Sep 2002.

[81] K. A. Jacobson, Z. G. Gao, A. Chen, D. Barak, S. A. Kim, K. Lee, A. Link, P. V. Rompaey, S. van Calenbergh, and B. T. Liang. Neoceptor concept based on molecular complementarity in GPCRs: a mutant adenosine A(3) receptor with selectively enhanced affinity for amine-modified nucleosides. *J. Med. Chem.*, 44:4125–4136, Nov 2001.

[82] O. Lenzi, V. Colotta, D. Catarzi, F. Varano, G. Filacchioni, C. Martini, L. Trincavelli, O. Ciampi, K. Varani, F. Marighetti, E. Morizzo, and S. Moro. 4-amido-2-aryl-1,2,4-triazolo[4,3-a]quinoxalin-1-ones as new potent and selective human A3 adenosine receptor antagonists. synthesis, pharmacological evaluation, and ligand-receptor modeling studies. *J. Med. Chem.*, 49:3916–3925, Jun 2006.

[83] V. Colotta, D. Catarzi, F. Varano, F. Capelli, O. Lenzi, G. Filacchioni, C. Martini, L. Trincavelli, O. Ciampi, A.M. Pugliese, F. Pedata, A. Schiesaro, E. Morizzo, and S. Moro. New 2-arylpyrazolo[3,4-c]quinoline derivatives as potent and selective human A3 adenosine receptor antagonists. Synthesis, pharmacological evaluation, and ligand-receptor modeling studies. *J. Med. Chem.*, 50:4061–4074, Aug 2007.

[84] E. Morizzo, F. Capelli, O. Lenzi, D. Catarzi, F. Varano, G. Filacchioni, F. Vincenzi, K. Varani, P.A. Borea, V. Colotta, and S. Moro. Scouting human A3 adenosine receptor antagonist binding mode using a molecular simplification approach: from triazoloquinoxaline to a pyrimidine skeleton as a key study. *J. Med. Chem.*, 50:6596–6606, Dec 2007.

[85] C. Bolcato, C. Cusan, G. Pastorin, G. Spalluto, B. Cacciari, K.N. Klotz, E. Morizzo, and S. Moro. Pyrazolo-triazolo-pyrimidines as adenosine receptor antagonists: Effect of the N-5 bond type on the affinity and selectivity at the four adenosine receptor subtypes. *Purinergic Signal.*, 4:39–46, Mar 2008.

[86] V. Colotta, D. Catarzi, F. Varano, O. Lenzi, G. Filacchioni, C. Martini, L. Trincavelli, O. Ciampi, C. Traini, A.M. Pugliese, F. Pedata, E. Morizzo, and S. Moro. Synthesis, ligand-receptor modeling studies and pharmacological evaluation of novel 4-modified-2-aryl-1,2,4-triazolo[4,3-a]quinoxalin-1-one derivatives as potent and selective human A3 adenosine receptor antagonists. *Bioorg. Med. Chem.*, 16:6086–6102, Jun 2008.

[87] V. Colotta, O Lenzi, D. Catarzi, F. Varano, G. Filacchioni, C. Martini, L. Trincavelli, O. Ciampi, A.M. Pugliese, C. Traini, F. Pedata, E. Morizzo, and S. Moro. The Pyrido[2,3-e]-1,2,4-triazolo[4,3-a]pyrazin-1-one as a New Scaffold to Develop Potent and Selective Human A3 Adenosine Receptor Antagonists. Synthesis, Pharmacological Evaluation and Ligand-Receptor Modeling Studies. *submitted*.

[88] S. Moro, F. Deflorian, M. Bacilieri, and G. Spalluto. Ligand-based homology modeling as attractive tool to inspect GPCR structural plasticity. *Curr. Pharm. Des.*, 12:2175–2185, 2006.

[89] S. Moro, P. Braiuca, F. Deflorian, C. Ferrari, G. Pastorin, B. Cacciari, P.G. Baraldi, K. Varani, P.A. Borea, and G. Spalluto. Combined target-based and ligand-based drug

design approach as a tool to define a novel 3D-pharmacophore model of human A3 adenosine receptor antagonists: pyrazolo[4,3-e]1,2,4-triazolo[1,5-c]pyrimidine derivatives as a key study. *J. Med. Chem.*, 48:152–162, Jan 2005.

[90] MOPAC 7, J. J. P. Stewart, (1993) Fujitsu Limited, Tokyo, Japan.

[91] C.A. Baxter, C.W. Murray, D.E. Clark, D.R. Westhead, and M.D. Eldridge. Flexible docking using Tabu search and an empirical estimate of binding affinity. *Proteins*, 33:367–382, Nov 1998.

[92] Thomas A. Halgren. Merck molecular force field. i. basis, form, scope, parameterization, and performance of mmff94. *Journal of Computational Chemistry*, 17:490–519, 1996.

[93] V. Colotta, D. Catarzi, F. Varano, F.R. Calabri, O. Lenzi, G. Filacchioni, C. Martini, L. Trincavelli, F. Deflorian, and S. Moro. 1,2,4-triazolo[4,3-a]quinoxalin-1-one moiety as an attractive scaffold to develop new potent and selective human A3 adenosine receptor antagonists: synthesis, pharmacological, and ligand-receptor modeling studies. *J. Med. Chem.*, 47:3580–3590, Jul 2004.

[94] S. Moro, F. Deflorian, M. Bacilieri, and G. Spalluto. Novel strategies for the design of new potent and selective human A3 receptor antagonists: an update. *Curr. Med. Chem.*, 13:639–645, 2006.

[95] S. Moro, M. Bacilieri, F. Deflorian, and G. Spalluto. G protein-coupled receptors as challenging druggable targets: insights from *in silico* studies. *New Journal of Chemistry*, 30:301–308, 2006.

[96] Z.G. Gao, A. Chen, D. Barak, S.K. Kim, C.E. Muller, and K.A. Jacobson. Identification by site-directed mutagenesis of residues involved in ligand recognition and activation of the human A3 adenosine receptor. *J. Biol. Chem.*, 277:19056–19063, May 2002.

[97] D. Catarzi, V. Colotta, F. Varano, F.R. Calabri, O. Lenzi, G. Filacchioni, L. Trincavelli, C. Martini, A. Tralli, C. Montopoli, and S. Moro. 2-aryl-8-chloro-1,2,4-triazolo[1,5-a]quinoxalin-4-amines as highly potent A1 and A3 adenosine receptor antagonists. *Bioorg. Med. Chem.*, 13:705–715, Feb 2005.

[98] D. Catarzi, V. Colotta, F. Varano, O. Lenzi, G. Filacchioni, L. Trincavelli, C. Martini, C. Montopoli, and S. Moro. 1,2,4-Triazolo[1,5-a]quinoxaline as a versatile tool for the design of selective human A3 adenosine receptor antagonists: synthesis, biological evaluation, and molecular modeling studies of 2-(hetero)aryl- and 2-carboxy-substituted derivatives. *J. Med. Chem.*, 48:7932–7945, Dec 2005.

[99] S. Moro, F. Deflorian, G. Spalluto, G. Pastorin, B. Cacciari, S.K. Kim, and K.A. Jacobson. Demystifying the three dimensional structure of G protein-coupled receptors (GPCRs) with the aid of molecular modeling. *Chem. Commun. (Camb.)*, pages 2949–2956, Dec 2003.

[100] S. Moro, G. Spalluto, and K.A. Jacobson. Techniques: Recent developments in computer-aided engineering of GPCR ligands using the human adenosine A3 receptor as an example. *Trends Pharmacol. Sci.*, 26:44–51, Jan 2005.

[101] A. Tafi, C. Bernardini, M. Botta, F. Corelli, M. Andreini, A. Martinelli, G. Ortore, P.G. Baraldi, F. Fruttarolo, P.A. Borea, and T. Tuccinardi. Pharmacophore based receptor modeling: the case of adenosine A3 receptor antagonists. An approach to the optimization of protein models. *J. Med. Chem.*, 49:4085–4097, Jul 2006.

[102] S.K. Kim, Z.G. Gao, L.S. Jeong, and K.A. Jacobson. Docking studies of agonists and antagonists suggest an activation pathway of the A3 adenosine receptor. *J. Mol. Graph. Model.*, 25:562–577, Dec 2006.

[103] Q. Jiang, B.X. Lee, M. Glashofer, A.M. van Rhee, and K.A. Jacobson. Mutagenesis reveals structure-activity parallels between human A2A adenosine receptors and biogenic amine G protein-coupled receptors. *J. Med. Chem.*, 40:2588–2595, Aug 1997.

[104] Q. Jiang, A.M. Van Rhee, J. Kim, S. Yehle, J. Wess, and K.A. Jacobson. Hydrophilic side chains in the third and seventh transmembrane helical domains of human A2A adenosine receptors are required for ligand recognition. *Mol. Pharmacol.*, 50:512–521, Sep 1996.

[105] V. Colotta, D. Catarzi, F. Varano, G. Filacchioni, C. Martini, L. Trincavelli, and A. Lucacchini. Synthesis and structure-activity relationships of a new set of 1,2,4-triazolo[4,3-a]quinoxalin-1-one derivatives as adenosine receptor antagonists. *Bioorg. Med. Chem.*, 11:3541–3550, Aug 2003.

[106] V. Colotta, D. Catarzi, F. Varano, G. Filacchioni, C. Martini, L. Trincavelli, A. Lucacchini, and V. Colotta. Synthesis and structure-activity relationships of 4-cycloalkylamino-1, 2, 4-triazolo[4, 3-a]quinoxalin-1- one derivatives as A1 and A3 adenosine receptor antagonists. *Arch. Pharm. (Weinheim)*, 337:35–41, Jan 2004.

[107] P.G. Baraldi, B. Cacciari, R. Romagnoli, G. Spalluto, K.N. Klotz, E. Leung, K. Varani, S. Gessi, S. Merighi, and P.A. Borea. Pyrazolo[4,3-e]-1,2,4-triazolo[1,5-c]pyrimidine derivatives as highly potent and selective human A(3) adenosine receptor antagonists. *J. Med. Chem.*, 42:4473–4478, Nov 1999.

[108] P.G. Baraldi, B. Cacciari, R. Romagnoli, G. Spalluto, S. Moro, K.N. Klotz, E. Leung, K. Varani, S. Gessi, S. Merighi, and P.A. Borea. Pyrazolo[4,3-e]1,2,4-triazolo[1,5-c]pyrimidine derivatives as highly potent and selective human A(3) adenosine receptor antagonists: influence of the chain at the N(8) pyrazole nitrogen. *J. Med. Chem.*, 43:4768–4780, Dec 2000.

[109] P.G. Baraldi, B. Cacciari, R. Romagnoli, S. Moro, X.D. Ji, Jacobson K.A., , S. Gessi, P.A. Borea, and G. Spalluto. Fluorosulfonyl- and bis-(beta-chloroethyl)amino-phenylamino functionalized pyrazolo[4,3-e]1,2,4-triazolo[1,5-c]pyrimidine derivatives: irreversible antagonists at the human A(3) adenosine receptor and molecular modeling studies. *J. Med. Chem.*, 44:2735–2742, 2001.

[110] P.G. Baraldi, B. Cacciari, S. Moro, G. Spalluto, G. Pastorin, T. Da Ros, K.N. Klotz, K. Varani, S. Gessi, and P.A. Borea. Synthesis, biological activity, and molecular modeling investigation of new pyrazolo[4,3-e]-1,2,4-triazolo[1,5-c]pyrimidine derivatives as human A(3) adenosine receptor antagonists. *J. Med. Chem.*, 45:770–780, Feb 2002.

[111] A. Maconi, G. Pastorin, T. Da Ros, G. Spalluto, Z.G. Gao, K.A. Jacobson, P.G. Baraldi, B. Cacciari, K. Varani, S. Moro, and P.A. Borea. Synthesis, biological properties, and molecular modeling investigation of the first potent, selective, and water-soluble human A(3) adenosine receptor antagonist. *J. Med. Chem.*, 45:3579–3582, Aug 2002.

[112] G. Pastorin, T. Da Ros, C. Bolcato, C. Montopoli, S. Moro, B. Cacciari, P.G. Baraldi, K. Varani, P.A. Borea, and G. Spalluto. Synthesis and biological studies of a new series of 5-heteroarylcarbamoylaminopyrazolo[4,3-e]1,2,4-triazolo[1,5-c]pyrimidines as human A3 adenosine receptor antagonists. Influence of the heteroaryl substituent on binding affinity and molecular modeling investigations. *J. Med. Chem.*, 49:1720–1729, Mar 2006.

[113] V. Colotta, D. Catarzi, F. Varano, L. Cecchi, G. Filacchioni, C. Martini, L. Trincavelli, and A. Lucacchini. Synthesis and structure-activity relationships of a new set of 2-arylpyrazolo[3,4-c]quinoline derivatives as adenosine receptor antagonists. *J. Med. Chem.*, 43:3118–3124, Aug 2000.

[114] V. Colotta, D. Catarzi, F. Varano, G. Filacchioni, C. Martini, L. Trincavelli, and A. Lucacchini. Synthesis of 4-amino-6-(hetero)arylalkylamino-1,2,4-triazolo[4,3-a]quinoxalin-1-one derivatives as potent A(2A) adenosine receptor antagonists. *Bioorg. Med. Chem.*, 11:5509–5518, Dec 2003.

[115] William Humphrey, Andrew Dalke, and Klaus Schulten. VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics*, 14:33–38, 1996.

[116] R Development Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria, 2005. ISBN 3-900051-07-0.

[117] L. Kaufman and P. J. Rousseeuw. *Finding groups in data: an introduction to cluster analysis.* John Wiley and Sons, New York, 1990.

[118] J.H. Ward. Hierachical grouping to optimize an objective function. *J. Am. Statist. Assoc.*, 58:236–244, 1963.

[119] H. Frauenfelder, S. G. Sligar, and P. G. Wolynes. The energy landscapes and motions of proteins. *Science*, 254:1598–1603, Dec 1991.

[120] C. Kandt, W.L. Ash, and D.P. Tieleman. Setting up and running molecular dynamics simulations of membrane proteins. *Methods*, 41:475–488, Apr 2007.

[121] E. Lindahl, B. Hess, and D. Van der Spoel. Gromacs 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Mod.*, 7:306–317, 2001.

[122] H. J. C. Berendsen, D. van der Spoel, and R. van Drunen. Gromacs: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, 91:43–56, 1995.

[123] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Dinola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81:3684–3690, 1984.

[124] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. Pedersen. A smooth particle mesh ewald method. *J. Chem. Phys.*, 103:8577, 1995.

[125] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An n [center-dot] log(n) method for ewald sums in large systems. *The Journal of Chemical Physics*, 98:10089–10092, 1993.

[126] Berk Hess, Henk Bekker, Herman J. C. Berendsen, and Johannes G. E. M. Fraaije. Lincs: A linear constraint solver for molecular simulations. *J. Comp. Chem*, 18:18–1463, 1997.

[127] V. Colotta, D. Catarzi, F. Varano, L. Cecchi, G. Filacchioni, C. Martini, L. Trincavelli, and A. Lucacchini. 1,2,4-Triazolo[4,3-a]quinoxalin-1-one: a versatile tool for the synthesis of potent and selective adenosine receptor antagonists. *J. Med. Chem.*, 43:1158–1164, Mar 2000.