MODELING OF 5-HT_{2A} AND 5-HT_{2C} RECEPTORS AND OF THEIR COMPLEXES WITH ACTUAL AND POTENTIAL ANTIPSYCHOTIC DRUGS

PhD Thesis
Computer- Assisted Drug Design Laboratory
Research group in Biomedical Informatics
Pompeu Fabra University

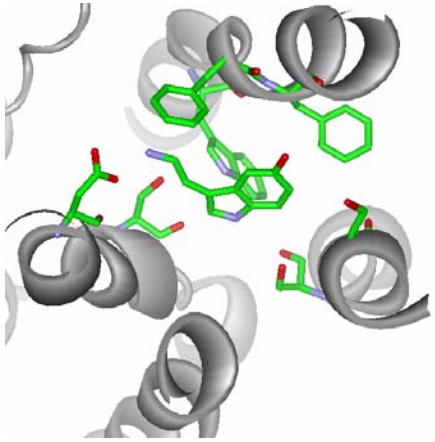
Cristina Dezi

c/ Pere IV 191, p, 1º
08018 Barcelona
Tel. 651539047
Email: cristina.dezi@gmail.com

Cristina Dezi

Computer- Assisted Drug Design Laboratory
Research group in Biomedical Informatics
Pompeu Fabra University

MODELING OF 5-HT_{2A} AND 5-HT_{2C} RECEPTORS AND OF THEIR COMPLEXES WITH ACTUAL AND POTENTIAL ANTIPSYCHOTIC DRUGS



Barcelona, 2007

PhD Thesis directed by:
Ferran Sanz i Carreras, Manuel Pastor Maeso

“Naturalmente li omini boni
desiderano di sapere”
Leonardo da Vinci

Dedicated to my family:
Thank you for your patience and love.

Dedicated to Patrizio Paoletti:
Without your teachings and advices I would not be what I am

Acknowledgements

At the end of this journey of study and research, I would like to thank all the persons that supported me during this period.

First of all I would like to thank my PhD tutors: Prof. Ferran Sanz i Carreras and Dr. Manuel Pastor Maeso, thank you for sharing your experience and for your guide, thank you for learning so much while working with you.

Then, thank you, thank you, thank you with all my love to all the persons that supported me with food, chocolate, hugs, smiles, laughs, and love: Genis Parra (for our moments), Josep Pareja (I love you Peppolino), Hugo & Raquel (for your advices and love), Xavi & Renata (for the dinners and the laughs), Jordi Rodrigo (for your support), Sergi Castellano (for the Passiò d’Esparraguera and the rest), Enrique Blanco (to be so spontaneous), JanJaap (for our talks and coffees), LoPep (for the scripts and Dr Slump), Robert (for the good moments), Eduardo (to be so wonderful), Beth (for your creativity), Ruth (for the theatre), Montse (for the beach), Carlos Sanz (for your help), Francisco (for Princesa 23, my big brother), Cristina Herraiz (for the coffee), Oscar Gonzalez (for the ham and the flip-flops), Alfons Gonzalez (without you my thesis would never happen!), Ramon Aragues (for our talks, the travels and the gadgets).

A special thank to the (ex) FUNDITEC staff: Gorka Izquierdo Zubiate (for your teachings and the voll-damm), Xavier Leal (for the laughs), Gemma (for being our “mamma”).

Another special thank to Javier Giron (for the flights and our good moments). Thank you to the PhDs at SISSA (Trieste), Andrea Della Chiesa, Michele Cirasuolo, Katrin Spiegel, Angelo Rosa: thank you for all the moments, and for still speaking to me even if I never phone… ☺. Thank you Davide, for our 9 years together and your help when needed, thank you Ylenia and Federica, my sisters.

Again, a very special thank to CIDEM staff that adopted me, making me feel at home and a real Catalan

between the Catalans: Mariona, Monica and the Europeans (Noemi, Mar, David, Marta, Marc), thank you Ivon of Fundaciò Bosch i Gimpera.

Thank to Gracia & Isabel, to come with me in my personal journey! Another special thank to my wonderful family: your support, patience and love are priceless! Thank you mum, dad, Cristiano & Paola & Francesco, Cinzia, thank you grandparents.

The most special thank to the person that was near me in the last one year and the half, Luca: thank you to believe in me, thank you for choosing me as travel mate, thank you for your love, thank you for your example, I love you now and I will love you forever.

Index

Abbreviations

5-HT_{2A}	Subtype 2A serotonin receptor
5-HT_{2B}	Subtype 2B serotonin receptor
5-HT_{2C}	Subtype 2C serotonin receptor
5-HT	Serotonin
APDs	Atypical Antipsychotic
bRho	Bovine Rhodopsin
BRO	Bacterio Rhodopsin
CD	Circular Dicroism
CNS	Central Nervous System
EM	Energy minimization
GDP	Guanosine diphosphate
GPCR	G protein-coupled receptors
GTP	Guanosine triphosphate
MD	Molecular dynamics
MIF	Molecular Interaction Field
PLS	Partial Least Square
QSAR	Quantitative Structure Activity Relationship
TA	Typical Antipsychotic
TMH	Transmembrane Helix

1. Introduction

Schizophrenia: ethiology, symptoms and therapy

Schizophrenia, in Greek “split mind”, is a mental disorder that affects about 1% of the population in industrialised countries, and the first symptoms typically occur between the age of 15 and 25.

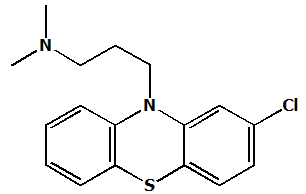
After the treatment and depending on the severity of the episodes, about 15% of patients return to a normal life, while 60% have intermittent episodes throughout their lives and another 25% never recover their ability to live as independent adults. Schizophrenia symptoms are divided into two broad categories: positive and negative symptoms (Wickelgren I.A. 1988).

Positive symptoms generally include distortion or exaggeration of normal sensation and perception and include delusions, hallucinations, paranoia, and incoherent speech. Instead, negative symptoms include the absence or diminution of normal cognitive, behavioural, motivational and affective functions, often characterised as apathy and lack of emotional expression. Other common deficits include reduced memory, attention and verbal fluency. In addition, schizophrenic patients also suffer of depression and anxiety.

Due to the complexity of the neural networks involved in cognition, perception and emotion, it is not surprising that the neurobiological basis of this pathology is still not clearly understood. Experts correlate positive symptoms with the dopamine neurotransmission in the limbic system and negative symptoms with hypodopaminergic activity in cortical regions involved in communication and cognition.

The first drugs for schizophrenia were discovered more than 60 years ago with Chlorpromazine (figure 1.1) that was the first drug to be used for psychotic symptoms in the late 1940s.

Figure 1.1: Chlorpromazine

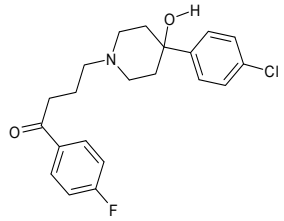


2-chloro-10-[3(-dimethylamino) propyl] phenothiazine monohydrochloride "Chlorpromazine"

Since then, as Chlorpromazine is effective against dopamine D_2 receptors, antipsychotic drug discovery was focussed on agents blocking D_2 receptors (Carlsson A. 1988). This first family of drugs, of which Haloperidol (Figure 1.2) is one of the most used, was called typical antipsychotics (TA). TA usually fail to manage negative symptoms, are ineffective in about one third of patients with schizophrenia and cause extrapyramidal side effects (EPS).

The best known EPS is tardive dyskinesia (involuntary, irregular muscle movements, usually in the face). Other common EPS include akathisia (restlessness), dystonia (muscular spasms of neck - torticollis, eyes - oculogyric crisis, tongue, or jaw; more frequent in children), drug-induced parkinsonism (muscle stiffness, shuffling gait, drooling, tremor; more frequent in adults and the elderly).

Figure 1.2: Haloperidol



These heavy extrapyramidal side effects were overcame with the discovery of clozapine, the first compound in use as atypical antipsychotics drug (APD). Its discovery was a combination of paradoxes, luck and unfortunate incidents (Hippius H. 1989; Hippius H. 1999). The main characteristics of Clozapine and the other atypical agents are the absence of EPS, the efficacy in patients for whom the typical neuroleptics had failed, and on negative symptoms and on symptoms related to mood and cognition. After this first accidental discovery, trying to improve the profile of TA by overcoming their main limitations, in the last ten years, the discovery of new antipsychotic agents has shifted from selective dopamine antagonist to compounds that have a broader receptor affinity profile: a new generation of antipsychotics arose, and they were called atypical antipsychotics.

Atypical antipsychotics (atypical APDs) also block serotonin receptors and are characterized by improved clinical efficacy and fewer side effects (Campbell M. 1999).

Since the discovery of Clozapine, the list of atypical APDs increased rapidly, including risperidone, olanzapine, quetiapine and ziprasidone, amisulpiride, sertindole and zotepine, to remind the most used both in North America and in Europe.

The action of APDs in schizophrenia is probably due to their blockage or inverse agonism activity on various 5-HT receptors, especially 5-HT_{2A} and 5-HT_{1A} receptors. The ratio of D_2 / 5-HT_{2A} blockage seems to be determinant for the atipicality, even if other receptors could be also involved (Reynolds 2004). Anyway, blockade of 5-HT_{2A} receptors coupled to weaker antagonism of dopamine (DA) D_2 receptors has been found to be the only classical pharmacologic feature those atypical APDs share in common, as highlighted in the original work by Melzter and co-workers (Meltzer H.Y. 1989; Woodward ND 2006)

This research group addressed the inhibition ratio of the 5-HT_{2A} and D_2 (the so called “Meltzer index”) as an

important factor to discriminate between typical and atypical antipsychotics (Meltzer H.Y. 1989; Roth B.L. 1995). In fact, studying a series of 37 compounds, they found that compounds with a Meltzer index < 1.09 showed behaviour of typical antipsychotics, while compounds with a Meltzer index > 1.12 showed behaviour of atypical antipsychotics.

The “Meltzer index” continues to be generative of novel APDs with efficacy for psychosis, cognitive dysfunction, and low extrapyramidal side effects (EPS) and continues to be used to discriminate and classify new compounds as “typical” or “atypical” agents.

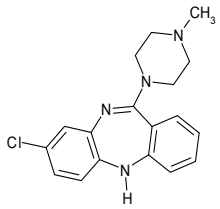
In addition to the implication of the serotonin receptor 5-HT_{2A} in schizophrenia, also 5-HT_{1A} receptor agonism has been suggested to be able to contribute to an atypical APD profile. Thus, both 5-HT_{2A} antagonism and 5-HT_{1A} agonism may be important for atypical APD action. In addition, some of these atypical APDs also have affinities for 5-HT_{2C}, 5-HT_6 or 5-HT_7 receptors that are in the same range as the affinity for the 5-HT_{2A} receptor. The involvement of the 5-HT_2 receptors in the pharmacological profile of atypical antipsychotics is supported by many biological, pharmacological and clinical studies (Green M.F. 1997; Aghajanian G.K. 2000).

Some studies point out that the effects of atypical antipsychotics that have been attributed to the blockade of the 5-HT_{2A} receptor may instead be due to the blockade of the 5-HT_{2C} that is another important feature for discriminating classical from atypical antipsychotics (Herrick-Davis K. 2000): Clozapine (Figure 1.3), Olanzapine, Seroquel, and other atypical antipsychotics have indeed greater affinity for 5-HT_{2C} (and in addition for 5-HT_{2A}) than for D_2 receptors (Cussac D. 2000).

In summary, there are several receptors of importance to the action of antipsychotic drugs, as reviewed in 2004 (Reynolds 2004), but the alternatives to D_2 receptor blockade as the essential antipsychotic mechanism are still under development even if there is general agreement on the

essential role of serotonin receptor 5-HT_{2A}. More controversial is the role of 5-HT_{2C}: even if it is not widely accepted as one of the receptors implied in the ethiology of schizophrenia, it should be taken into account, as potentially implied in this disorder.

Figure 1.3: Clozapine



Considering all these pharmacological data regarding the existing drugs for schizophrenia, the laboratory of Pharmaceutical Chemistry, (Faculty of Pharmacy, University of Santiago de Compostela), in collaboration with the Department of Pharmacology (Faculty of Pharmacy, University of Santiago de Compostela) synthesized and tested pharmacologically a series of new compounds with interesting chemical and pharmacological characteristics (Brea J. 2002; Brea J. 2003).

The series comprises 76 conformationally constrained butyrophenones, with a large molecular diversity in their alkanone and / or amino moieties, as shown in the Chart I and Chart II. The combination of the alkanone and amine fragments in Chart I and Chart II generates the 76 compounds, which present diversity in their affinity and specificity for 5-HT_{2A} and 5-HT_{2C} receptors.

Chart I – Cycloalcanone scaffolds

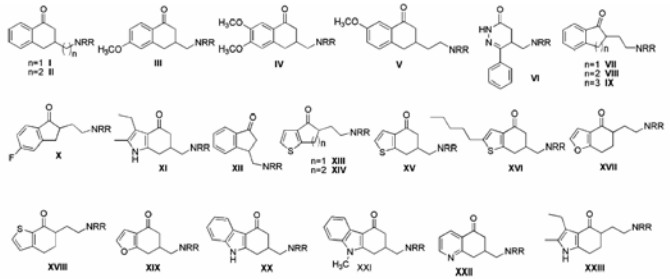
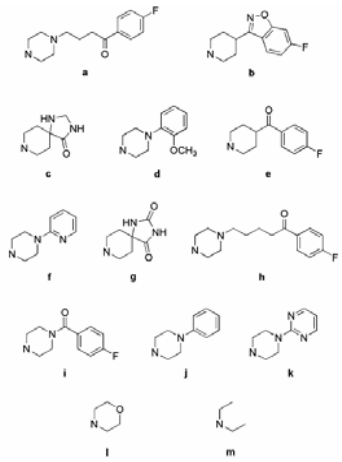


Chart II – NRR Substituents



This large series of compounds have been tested as antagonists of the 5-HT_{2A} and 5-HT_{2C} serotonin receptors and have shown to be useful tools for defining the pathophysiological role of each 5-HT_2 subtype and for identifying structural features relevant to receptor recognition and subtype discrimination (Brea J. 2006).

Some compounds were found highly active (pKi > 8.76), and the series showed interesting pharmacological properties (Brea J. 2002; Brea J. 2003; Alvarado, Coelho et al. 2005; Brea, Castro et al. 2006):

a. they inhibited both 5-HT_{2A} and 5-HT_{2C}, covering a wide range of potency and selectivity

b. the series comprise both antagonists and inverse agonists

c. the compounds also inhibited D_2 receptors

d. the series comprises some compounds with a Meltzer index characteristic of atypical antipsychotics and others with a Meltzer index of classical antipsychotics

e. some compounds showed a monophasic curve while others showed a biphasic curve when considering the inhibition of the second messenger response

Within this scenario, it is important to understand the molecular mechanisms that govern ligand receptor interaction, in order to discover more effective and selective drugs towards schizophrenia.

Considering the importance of serotonin receptors 5-HT_{2A} and 5-HT_{2C} in this pathology and the amount of pharmacological data about known atypical APDs drugs and new ligands (such as the butyrophenone series), it was decided to approach the problem of drug discovery for

schizophrenia with computational techniques, that could help to rationalize the experimental data and to identify the essential chemical features governing potency and selectivity toward one receptor subtype (5-HT_{2A}) versus the other (5-HT_{2C}).

Therefore the present thesis was focussed on the modelling of 5-HT_{2A} and 5-HT_{2C} receptors and of their complexes with actual and potential antipsychotic drugs.

The research, carried out with computational techniques at the state of the art, aimed at study the structural requirement for ligands potent and selective towards one subtype versus the other and at design of ligands with new desired properties.

2. Background

2.1. The 5-HT_{2A} and 5-HT_{2C} receptors

The target receptors, serotonin 5-HT_{2A} and 5-HT_{2C} receptors belong to the GPCR superfamily, which is a very large and complex family of receptors activated by a variety of stimuli that span from light to proteins, through hormones and amines as it will be explained in the following section.

5-HT_{2A} and 5-HT_{2C} receptors are localized both in the periphery and in the Central Nervous System (CNS), and they mediate many diverse physiological functions. In fact, 5-HT_{2A} receptors play a role in appetite control, thermoregulation and sleep. They are also involved in cardiovascular function and muscle contraction and seem to play an important role in different neuropsychiatric disorders.

Various antipsychotic agents and antidepressants bind with relatively high affinity at 5-HT_{2A} receptors. 5-HT_{2A} antagonists also possess anxiolitic properties and 5-HT_{2A} receptors may be involved in the actions of the classical hallucinogens: indolealkylamine such as 5-methoxy-N,N-dimethyltryptamine; 5-Ome DMT and ergot related compounds (such as LSD),DOB, DOI (Vernier P. 1995).

Phenylalkilamine hallucinogens also bind at 5-HT_{2B} and 5-HT_{2C} receptors, and there is a significant correlation between human potency and receptor affinity for different agents (Nelson D.L. 1994).

Recent studies suggest that 5-HT_{2A} receptors may play a more prominent role than 5-HT_{2C} in the behavioural actions of hallucinogens (Ismaiel A.M. 1993; Schreiber R. 1994; Fiorella D. 1995).

2.1.1.GPCR superfamily

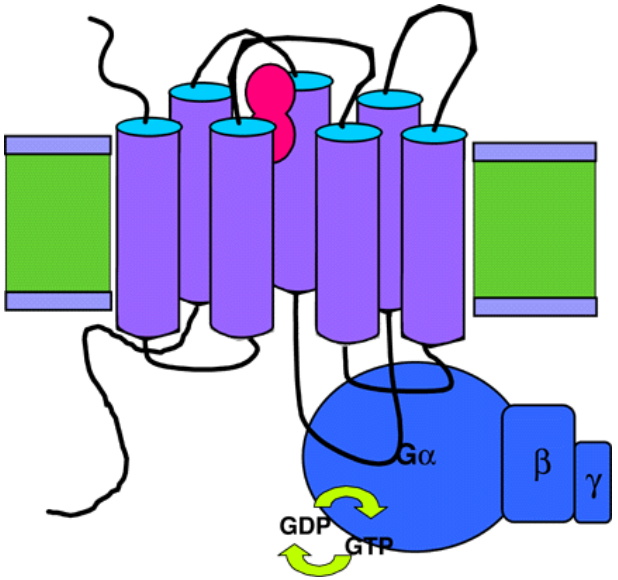
G-protein-coupled receptors (GPCRs) constitute a large and diverse family of proteins whose primary function is to transduce extracellular stimuli into intracellular signal, through the activation of their intracellular domains with heterotrimeric G-proteins.

They are the largest and most diverse protein family in mammalian genomes (constituting the 1% of the whole mammalian genomes, 950 genes in human genome (Takeda S. 2002) and are present ubiquitous in the genome of a variety of organisms, such as bacteria, yeast, plants, nematodes and other invertebrate groups.

Upon the activating stimuli (such as light, neurotransmitters, odorants, biogenic amines, lipids, proteins, aminoacids, hormones, nucleotides or chemokines), the GPCRs undergo a conformational change to an activated state in which they can interact with G proteins.

G proteins are heterotrimeric proteins consisting of α, β and γ subunits (Flower D.R. 1999). Their name originates in their ability to bind guanosine diphosphate (GDP) or guanosine triphosphate (GTP). After the activation of the GPCR, the conformation in the associated G protein α-subunit changes and leads to release of GDP followed by binding of GTP.

Figure 2.1.1.4. Schematic representation of a GPCR.



This ternary GPCR-G(αβγ)-GTP complex is highly unstable and dissociates into free GPCR, Gα-GTP and Gβγ. Both Gα-GTP and Gβγ can stimulate specific effectors, by activating or inhibiting the production of a variety of second messengers.

The Gα-subunit has catalytic GTPase activity that promotes the return to the inactivated state of the complex GPCR-G-protein.

GPCR receptors subfamilies

Several classification systems have been used to sort out this superfamily. Some systems group the receptors by

how their ligand binds, and others have used both physiological and structural features.

One of the most frequently used systems divides mammalian GPCRs into five distinct subfamilies based on sequence similarities: A-E (Attwood T.K. 1994; Kolakowski L.F. 1994; Bockaert J. 1999; Pierce K.L. 2002).

The letter code stands for:

- Class A Rhodopsin like
- Class B secretin like
- Class C metabotropic glutamate / pheromone
- Class D fungal pheromone
- Class E cAMP receptors

There is also another class, for the frizzled / smoothened receptors.

This A-E system is designed to cover all GPCRs, in both vertebrates and invertebrates, and in fact some families do not exist in humans.

More recently, GPCRs have been classified into five distinct groups based on phylogenetic analyses of sequences from the human genome (Fredriksson R. 2003). This classification system has been named GRAFS, which is an acronym for the five different groups: Glutamate (G), Rhodopsin (R), Adhesion (A), Frizzled/Taste2 (F), and Secretin (S).

Of all these families, the largest one is the Rhodopsin receptor family, which is formed by the highest number of receptors and corresponds to what has previously been called the Class A in the A-F classification system.

The receptors belonging to the Rhodopsin family have several structural characteristics in common, such as the NsxxNPxxY motif in TMHVII or the DRY motif (also defined as D(E)-R-Y(F), if the most frequent variations are taken into account) at the border between TMHIII and IL2. The ligand for

most of the Rhodopsin receptors binds within a cavity between the TM regions (Baldwin J.M. 1994).

This group includes: the prostaglandin receptor cluster, the amine receptor cluster (including serotonin receptors), opsin receptors clusters, melatonin receptors, melancortin receptors, hypocretin receptors, neuropeptide receptors, takichinin receptors, colecystokinin receptors, neuropeptide Y receptors, the endothelin-related receptors, gastrin-releasing peptide receptor, neuromedin receptors, tyrotropine releasing hormone receptors, gonadotropin-releasing hormone receptors and oxytocin receptor.

GPCR receptors: structural features common to all the subfamilies. The lessons learnt from the crystal structure of bovine Rhodopsin.

First information about the structure of GPCRs was obtained thanks to electron cryo-microscopy map of Bacteriorhodopsin (Henderson R. 1990; Hiber M.F. 1991).

Bacteriorhodopsin is a membrane protein of the microrganism *Halobacterium halobium* that also consists of seven membrane-spanning helical structures, even if it is not coupled to G proteins and there is very low sequence similarity between Bacteriorhodopsin and GPCRs (the sequence identity percentage is about 20%).

More experimental information came in 1993, when Baldwin derived by analysis of ~200 GPCR sequences a model establishing the potential arrangement of the seven TMHs that was compatible with a low-resolution 2D projection density map of bovine Rhodopsin obtained by cryo-microscopy (Baldwin J.M. 1993; Schertler G.F.X. 1993). This model suggests that the relative organization of the helixes of GPCRs is somewhat different from that of

Bacteriorhodopsin but confirms nevertheless a large degree of overall similarity. Further advances were achieved with the determination of a new low resolution density map of

frog Rhodopsin that had a significant resolution perpendicular to the membrane plane (Unger V.M. 1997).

Baldwin consequently presented an updated model based on an analysis of ~500 sequences (Baldwin J.M. 1997) that included these new data.

In 2000 Palczewski et al (Palczewski K. 2000) succedeed in solving a x-ray structure of bovine Rhodopsin (Figure 2.1.1.5) at a resolution of 2.8 Å, the first experimentally determined high resolution three-dimensional structure of a GPCR. This structure was subsequently refined and solved to a resolution of 2.6 Å (Teller D.C. 2001) and more recently to 2.2 Å (Okada, Sugihara et al. 2004).

Figure 2.1.1.5 – Bovine Rhodopsin crystal structure at 2.8 Å resolution (pdb code: 1F88)

Rhodopsin activation stimulus is the light, and its ligand (11-cis-retinal) is covalently bound to the receptor (the apoprotein opsin). The 11-cis-retinal isomerizes upon absorption of a photon to all-trans-retinal, and all the isomerisation occur within the binding pocket. The resolution of a crystal structure of a protein belonging to the GPCR superfamily, constituted a great advance as allowed changing and improving some wrong assumption about the structure and the organization of the helixes (Palczewski K. 2000; Teller D.C. 2001).

The seven TM helixes have an anti-clockwise arrangement when viewed from the extracellular side and they have different lengths due to several kinks and bends as well as varying tilts with respect to a hypothetical membrane surface.

The kinks and the deviations from the straightness are due to aminoacids such as prolines and glycines that prevent the formation of intra-helical hydrogen bonds because of the lack of a polar side chain. Most of the residues that cause the bend of the helixes are highly conserved residues within the GPCR superfamily as shown by Baldwin J.M. in 1997; therefore they are supposed to play an important structural role common to all receptors, maybe being involved in the activation mechanisms.

Helix I contains a bend due to P53, whereas helix II has a bend due to a G-G motif (G89-G90). Helix III is the most tilted and longest helix. It makes multiple inter-helical contacts, with helix II with the outer leaflet of the bilayer, with helixes VI and VII within the inner leaflet, and with helix IV and V near the intracellular surface.

Helix III contains two small bends, the first one at the residues G120-G121, and the second one at S127. Helix IV is the least tilted and shortest helix. It contains a bend at the extracellular end due to P170-P171. Helix V contains two small bends, one at F203, the other one at H211. Helix VI is the most bent helix. The residue P267 causes this bend, which is

one of the most conserved amino acids in the Rhodopsin-like family. Helix VII contains a significant distortion around K296, the retinal attachment side. P303 produces a significant bend.

Additionally to the transmembrane domains, there is another short helix, Helix VIII, in the cytoplasmic surface. This helix is a part of the fourth cytoplasmic loop. It starts at K311, and end with C322 and C323 that attach palmitoyl moieties within the membrane.

The N-terminus of the receptor contains five strands, two of which form a typical β-sheet fold. Most surprising is the conformation of the extracellular loop E2. A part of his loop folds deeply into the centre of Rhodopsin, thus restricting the size of the retinal binding site. This fold is constrained by a disulfide bridge between the E2 and helix III (Palczewski K. 2000; Teller D.C. 2001), which seems to be conserved through the GPCR superfamily.

The inactive conformation of bovine Rhodopsin as found in the crystal structure structure is stabilized by a number of interhelical hydrophobic interactions and hydrogen bonds; most of them formed by highly conserved residues among Rhodopsin-like GPCR receptors.

An overview of the most conserved residues and of the hydrogen bond network that they form in the crystal structure of Rhodopsin is given in the Figure 2.1.1.6. For clarity, in the next sections residues will be indicated using the numbering scheme developed by Ballesteros and Weinstein (Ballesteros J. 1994).

According to this numbering scheme, residues in the GPCR family are defined by two numbers: one indicating the transmembrane helix where they are located and one for the position of these residues with respect to the most conserved residue in the same helix, which is given the number 50. For example, the most conserved residue in helix III is the R3.50, and the conserved aspartic acid in helix III (common to receptors activated by biogenic amines) is the D3.32.

Figure 2.1.1.6. Schematic representation of bovine Rhodopsin (1F88). The colouring method allows to appreciate the seven TM helixes (TMH). The residues responsible for the bending in each helix are represented as stick. Loops are not shown for clarity.

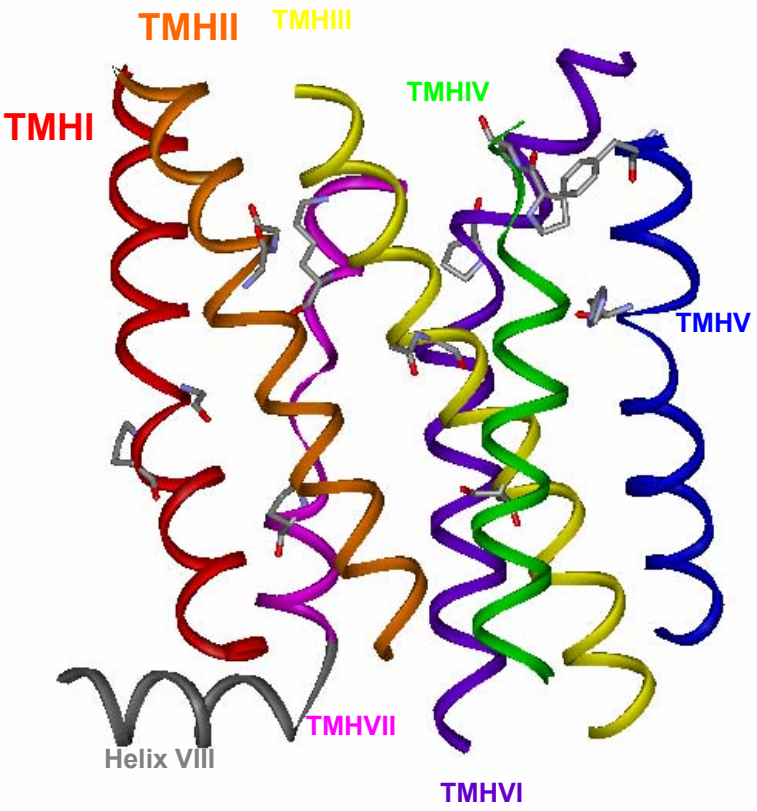
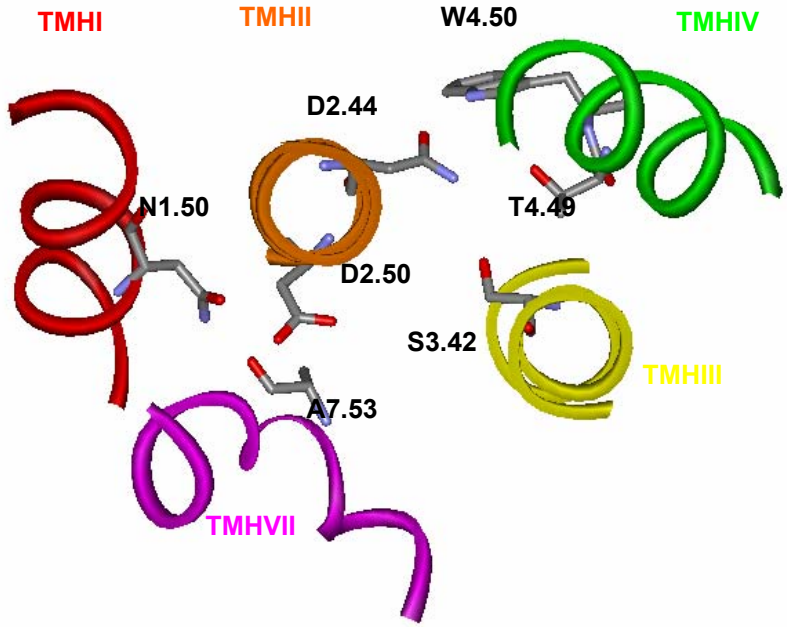


Figure 2.1.1.7. The most relevant hydrogen bond interactions are shown, as described in the paper (Palczewski K. 2000)



The main interactions found are:

- D2.50 of TMHII is directly hydrogen bonded with the conserved N1.50 of TMHI that additionally interacts with the carbonyl of oxygen of A7.46 of TMHVII.

- D2.50 is interacting with N7.49 via a water molecule, thus connecting the three highly conserved residues N1.50-D2.50-N7.49 via a hydrogen bond chain.

- N2.44 in TMHII hydrogen bonds the OH group of S3.42 of TMHIII and the OH group of T4.49 and the ring nitrogen of W4.50 of TMHIV.

- E3.49 of TMHIII forms a salt bridge with its neighbour R3.53 on TMHIII, E6.35 of TMHVI as well as a hydrogen bond with the OH group of T6.39 of TMHVI.

- E3.49 and R3.50 belong to the DRY sequence that is highly conserved among the Rhodopsin-like GPCRs.

- H5.47 of TMHV and E3.37 in TMHIII form another strong hydrogen bond.

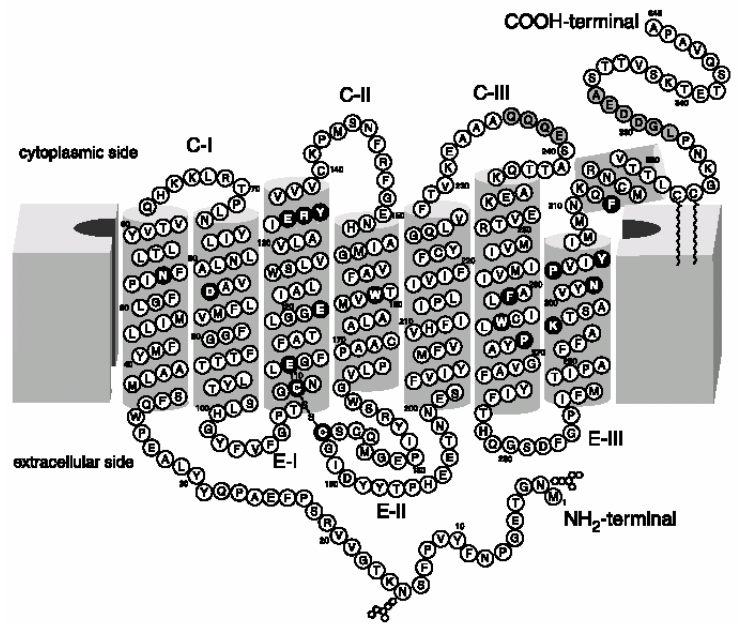
- K7.50 in TMHVII and its counter ion E3.28 in TMHIII

Particular important seems to be the role played by water molecules in stabilizing some of these hydrogen bonds (Okada T. 2002).

Three water molecules are clustered between TMHs II, III, VI and VII, mediating the interaction of N302 that is a part of the highly conserved residue NpxxY motif in TM VII with residues of other helixes (Figure 2.1.1.8 key residues in the bovine Rhodopsin structure). Two water molecules also mediate the hydrogen bond network around the Schiff base.

One of these water molecules is close to E3.28, and the second water molecule is interacting with the side chains of E181 and S186 that belong to the second extracellular loop (EL2). Another water is located between the helixes VI and VII, close to C6.51 and P6.55. The last observed water mediates interactions between helixes I and II at the cytoplasmic surface.

Figure 2.1.1.8. Key residues in the bovine Rhodopsin structure (and in most GPCRs), and their location in the TMH are presented in the following scheme. The scheme is taken by the original paper published after the release of 1F88 (Palczewski K. 2000).



The complex hydrogen bond network described above is supposed to stabilize the inactive form of GPCR receptors, as the bovine Rhodopsin structure was determined and solved in its inactive form. In order to activate, the structure is supposed to break some of the most conserved interhelical interaction, and in this process essential could be the role played by some residues, like prolines and glycines, known to give flexibility to the protein structures.

However, even if the experimental methods are improving faster and faster allowing better insights into the helical movements that occur during the activation process, few detailed information is available about the conformational changes. New insights could be provided by computational methods based on experimental data, as reviewed by various authors in the last years (Ballesteros J. 2001; Sylte 2001; Shapiro D.A. 2002; Bissantz 2003; Weinstein H 2006)

Activation: what is known, what is supposed, the disulfide bond between TMHIII and EL2

Since, so far, the only known structure of a GPCR correspond to the high resolution x-ray structure of an inactive state of GPCR (bovine Rhodopsin), it is not possible to directly compare an inactive with an active conformation to learn which conformational changes take place during activation.

However, in the last years, biophysical studies carried out mainly on Rhodopsin and Class A receptors provided some information about these rearrangements

In the inactive state of Rhodopsin, the ligand 11-cis-retinal acts is assumed to stabilize the receptor in the inactive conformation, much like the “inverse agonists”. Upon absorption of a photon, the retinal isomerizes to the all-trans-conformation. In this conformation, the retinal stabilizes or promotes conformational rearrangements of the protein leading to the active conformation, that contains all-trans-retinal as a covalently bound agonist and that is capable of interacting with the G protein transducin Gt.

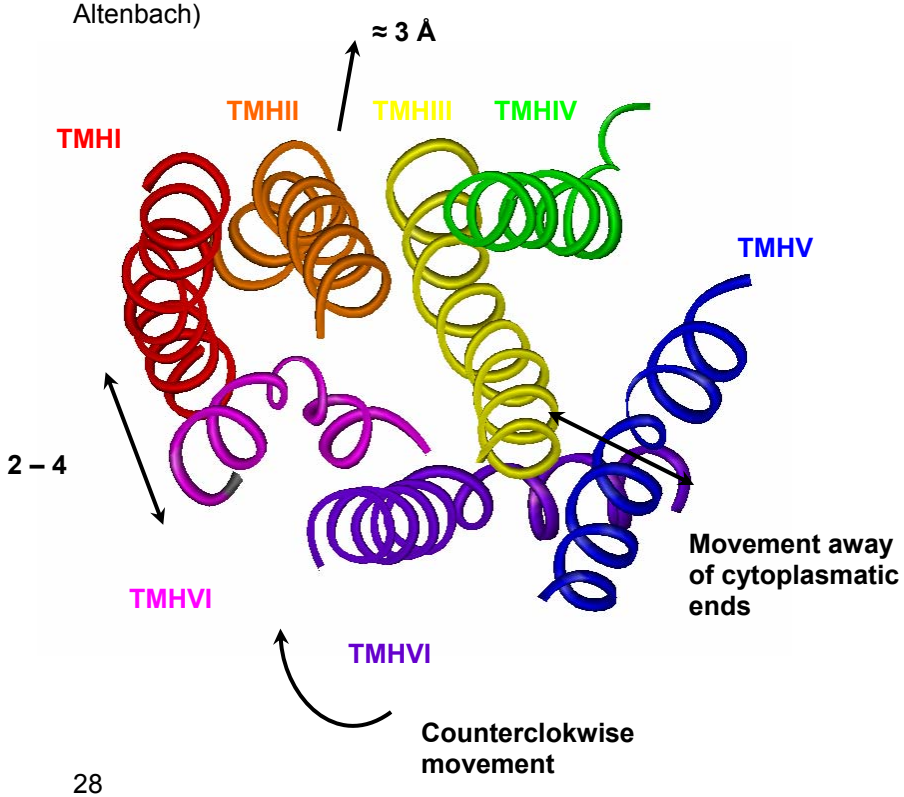
Using different biophysical experimental techniques (such as UV absorption spectroscopy, site direct spin-labelling, fluorescence and circular dicroism (CD) spectroscopy) it was possible to have insights about the helical movements involved in Rhodopsin and other Class A receptors activation.

The use of these different techniques leads to the suggestion that the formation of an active intermediate

involves mainly rigid body motion of TMHIII and TMHVI, with the motion of TMHVI being significantly lager than that of TMHIII that was interpreted to be smaller than 5Å. Small changes were also observed in TMHII, although these were smaller if compared to those in TMHs VI and III (Farahbakhsh Z.T. 1995; Altenbach C. 1996; Altenbach C. 1999).

Other biophysical studies lead to the suggestion of a rigid-body motion of TMHVI in a counterclockwise direction when viewed from the extracellular side and an outward movement of the cytoplasmatic end of TMHVI away from TMHIII (Farrens D.L. 1996).

Figure 2.1.1.9. The seven transmembrane bundle and the schematic representation of the movements that could possibly occur during the activation process (according to Farrens and Altenbach)



Further studies carried out by Altenbach and co-workers, found a movement of the cytoplasmic portion of TMHVII away from TMHI, by 2-4 Å (Altenbach C. 2001) and an outward displacement of TMHII relative to helix VIII by ≈ 3 Å (Altenbach C. 2001).

Using fluorescence and circular dicroism (CD) spectroscopy to study the structural properties of helix VIII, Krishna and co-workers (Krishna A.G. 2002) suggested that helix VIII has different conformation in the dark and in the light activated state of Rhodopsin.

They suggested that helix VIII remains in helical conformation in the inactive state, and that upon illumination it adopts a looplike structure (Krishna A.G. 2002).

Several biophysical studies that analyzed the conformational changes during receptor activation of GPCRs with diffusible ligands were carried out using the adrenergic β_2 receptor, the muscarinic M3 receptor, the parathyroid hormone receptor (PTHR) (Gether U. 1995; Gether U. 1997; Sheikh S.P. 1999; Ghanouni G. 2001; Jensen A.D. 2001; Ward S.D.C. 2002).

The data obtained from such studies are in agreement with the results of the studies with Rhodopsin suggesting that at least all Class A GPCRs share a similar activation mechanism. In summary, the current activation model includes an agonist-induced separation of the cytoplasmic ends of the TMHs III and VI mainly achieved by a movement of the cytoplasmic part of the TMHVI away from the receptor, but also by a smaller movement of TMHIII. A counterclockwise rotation (when viewed from the extracellular surface) of TMHVI around its axis might also occur as well as movements of other helixes.

It is assumed that the inactivated state is stabilized by interactions that have to be broken during receptor activation. Evidence therefore is given by the observation of mutations that lead to mutant receptors with significant agonist-independent constitutive activity (Cohen G.B. 1993; Scheer A.

1997; Rasmussen S.G.G. 1999; Alewijnse A.E. 2000), suggesting that these mutations break interactions stabilizing the inactive state and thus promote the transition to an active conformation (Table I). Table I summarizes the residues, for different GPCRs, which are found to stabilize the inactive conformation of the receptor, and, if mutated, could increase the constitutive activation of the receptor (reviewed in (Seifert R. 2002).

While the above-cited interactions stabilize the inactive state and are disrupted during receptor activation, it is also expected that a new set of interactions stabilizes the active state. The hypothesis has emerged that an interaction between the conserved D2.50 in the TMHII and the conserved N7.49 of the NPxxY motif in TMHVII are important for receptor activation.

In the gonadotropin-releasing hormone receptor (GnRHR) the two residues are switched, D2.50 is replaced by N2.50 and N7.49 by D7.49. Replacing N2.50 in GnRHR by an aspartate (N2.50D) disrupted the function of the receptor, which could, however, be restored by the double-revertant mutant N2.50D/D7.49N (Zhou W. 1994).

A similar result was obtained for the 5-HT_{2A} receptor (Sealfon S.C. 1995). Whereas the D2.50N mutation eliminated coupling of the receptor, the double mutant D2.50N/N7.49D was again functional, indicating that such favourable network of hydrogen bonding is required for proper agonist-induced conformational changes in the receptor (Sealfon S.C. 1995). Analogous results were obtained for the tyrotropine-releasing hormone receptor (Perlman J.H. 1997) and the tachykinin NK2 receptor (Donnelly D. 1999).

Table I. Summary of the mutations that were found to play a relevant role in increasing the constitutive activity of some GPCR.

Residue	GPCR	Constitutive activity
E3.49Q	Opsin	+
E3.49D	Opsin	-
D3.49A (Rasmussen S.G.G. 1999)	β2	+
R3.50 (Ballesteros J. 1998), (Alewijnse A.E. 2000)	H2	+
D3.32, E6.30, K7.36 (Flower D.R. 1999), (Greasley P.J. 2002), (Shapiro D.A. 2002), (Porter J.E. 1996)	α1B 5-HT_{2A}	+
C6.34, V6.36, V6.40 (Shapiro D.A. 2002); E6.30Asn, E6.30Gln, E6.30L, D3.49Ala (Visiers 2001; Visiers 2001); E6.30Ala, V6.40A, E6.30R (Shapiro D.A. 2002)	5-HT_{2A}	+
N6.29A, K6.32A, R3.50E and the double R3.50 / E6.30 mutant (Shapiro D.A. 2002)	5-HT_{2A}	-
K7.43, E3.28, M6.40 (Robinson P.R. 1992), (Han M. 1998)	Rhodopsin	+
D3.32, Tyr7.43 (Befort K. 1999)	δ-opiod	+
N3.35, Y7.43 (Groblewski T. 1997).	angiotensin AT-1	+

One of the questions that still remain to be answered is how binding of an agonist or isomerisation of retinal can trigger the described conformational switch that lead to an active conformation. Shi et al (Shi L. 2002) recently proposed that the highly conserved C6.47 modulates the configuration of the highly conserved aromatic cluster in TMHVI (W6.48, F6.51 and F6.52) and the TMHVI Pro-kink (P6.50) through specific interactions in its different rotamer configurations.

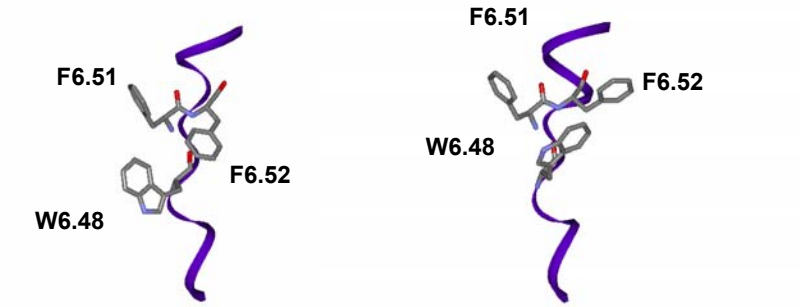
Using biased Monte-Carlo technique of conformational memories, they showed that in the β_2-adrenergic receptor the rotamer changes among C/S/T6.47, W6.48, and F6.52 are highly correlated. A change of W6.48 from *g+*, its conformation in the Rhodopsin inactive state structure, to *t* must be accompanied by a corresponding change of F6.52 to *t* to avoid steric clashed. Vice versa, a change of F6.52 from *t* to *g+* must be accompanied by a corresponding change of W6.48 from *g+* to *t*. Moreover, the rotamer conformation of C6.47 was found to be correlated with the rotamer of W6.48. C6.47 in *g+* was found to be associated with W6.48 in *t*, and C6.47 in *t* with W6.48 in *g+*. They thus propose that C6.47*g+*/W6.48*t*/F6.52*t* represents the active state and that C6.47*t*/W6.48*g+*/F6.52*g+* represents the inactive state. This rotamer switch is suggested to impact the α-helix backbone and thus the movement of TMHVI about the Pro-kink during receptor activation through rotamer-dependent hydrogen bonds of C6.47. They supported their hypothesis with mutants such as C6.47Thr.

For GPCR such as the β_2-adrenergic receptor that are activated by diffusible ligands, a ligand that stabilizes W6.48 in the *g+* conformation should behave as an inverse agonist. Agonists, on the other side, should induce the *t* conformation of W6.48. For the biogenic amine receptors, this role is probably taken over by the aromatic ring of the agonist.

Ebersole et al (Ebersole BJ 2003) proposed for the 5-HT_{2A} receptor that residues in TMHIII and TM V serve to adjust the position of the aromatic indole ring of the studied ligands in the binding pocket. This indole ring then interacts with the

aromatic cluster inducing the conformational switch (Ebersole BJ 2003). For the natural full agonist serotonin, the aromatic ring is optimally positioned, thus inducing full response (Ebersole BJ 2003). If for other agonists the aromatic ring is not positioned in the same way, they do not induce the full effect (Ebersole BJ 2003). Information about all these features regarding the GPCR superfamily and the major differences between the GPCR sub-families were recently summarised by Kristiansen (Kristiansen 2004). In the Figure 2.1.1.10 above, the aromatic cluster for serotonin 5-HT_{2A} model with the two different conformation of the aromatic cluster (W6.48 and F6.51) is shown.

Figure 2.1.1.10. Two different conformation of the aromatic cluster (W6.48 and F6.51) for 5-HT_{2A} receptor.



2.1.2.Serotonin receptors

Serotonin receptors belong to the Class A of GPCR receptors superfamily (Rhodopsin like receptors), and to the subclass of receptors activated by biogenic amines (Kroeze W.K. 2002). Seven 5-HT receptor families with defined functions, 5-HT_1 through 5-HT_7, currently are recognized (Roth B.L. 1984;

Hoyer D. 1994) (Table II), and have a predicted membrane topology composed of an extracellular N terminal segment linked to an intracellular C terminus by seven transmembrane spanning segments. The 5-HT_3 receptor is the only member of the serotonin family that is not coupled to G-proteins, but it is a ligand-gated ion channel that gates Na+ and K+ and has a predicted membrane topology similar to that of the nicotinic cholinergic receptor.

The 5-HT receptor subtypes are expressed in distinct but often overlapping patterns (Bloom F.E. 1995) and are coupled to different transmembrane signalling mechanism (Table II). 5-HT has been implicated in a variety of physiological and behavioural functions, such as affection, aggression, appetite, cognition, emesis, endocrine function, motor function, neurotrophism, perception, sensory function, sex, sleep, and vascular function (Bloom F.E. 1995). The variety of physiological processes in which serotonin is involved seems to be related with the molecular diversity and differential cellular distribution of the many 5-HT receptor subtypes that are expressed in brain and other tissues.

In fact, serotonin is found in high concentrations in enterochromaffin cells throughout the gastrointestinal tract, in platelets, and in specific regions of the CNS, but the exact sites and modes of action of 5-HT still remain not precisely defined. In non neuronal tissues, 5-HT has many roles, including smooth muscle growth and contraction and platelet aggregation. The G-protein coupled serotonin receptors are predicted to be typical heptahelical GPCRs, and although the structures of none of the 5-HT receptors have yet been solved, it is likely that the basic features of these structures will be similar to those of bovine Rhodopsin (Palczewski K. 2000). The sequences of the mammalian G-protein coupled 5-HT receptors are very highly conserved. For example, among the 12 human G-protein coupled 5-HT receptors, which vary in total length from 358 to 482 residues, there are 33 residues

that are 100% conserved, and an additional 27 residues that are at least 80% conserved (Table III).

Table II. Location of 5-HT subtype receptors, second messengers to which they are coupled and the phisiological response that each subtype mediates.

Receptor	Location	G protein	Second messenger	Effect
5-HT1	Raphe nuclei Axon terminals Substantia nigra Basal ganglia	Gi / G0	Adenylyl ciclase	Inhibition of adenylil cyclase
5-HT2	Prefrontal cortex, claustrum, platelets Gastrointestinal tract	Gq / G11	Phospholipase C (PLC) Phospholipase A2 (PLA2) Diacylglicerol	Activation protein C Release Ca2+
5-HT3 ligand gated ion channel	Gastrointestinal tract Nucleus tractus solitarii Area postrema	No coupled to G proteins	Na+ / K+	More concentration of Na+ and K+
5-HT4	CNS Heart Gastrointestinal tract	Gi / G0 / Gs	Adenylyl cyclase	Activation of adenylil cyclase
5-HT5	CNS	Gi / G0	Adenylyl cyclase Phospholipase C	Inhibition of adenylil cyclase
5-HT6	Basal ganglia	Gs	Adenylyl cyclase	Activation of adenylil cyclase
5-HT7	CNS, Blood vessels	Gs	Adenylyl cyclase	Activation of adenylil cyclase

Table III. Structural common features of serotonin receptors and sequence conservation

Receptor segment	Conserved residues (> 80%) whithin serotonin receptor subfamilies
N terminus	Poor conservation
TMH I	P1.30, P1.31, R1.50
IL1	L1.63
TMHII	P2.38 or P2.39, S2.45, D2.50, V2.57, P2.59, L2.46
EL1	Motif PXP
TMHIII	C3.25, D3.32, S3.39, I3.40, I3.46, D3.49, R3.50, Y3.51, I3.54
IL2	Poor conservation
TMHIV	S4.38 or T4.38, M4.58, W4.50
EL2	Poor conservation, except for a Cys residue near the extracellular end of TMHIII
TMHV	P5.37 or G5.37, F5.47, P5.50, Y5.58, K5.66 or Q5.66 or R5.66
IL3	Porly conserved in sequence and length (from 23 to 106 residues)
TMHVI	Charged residues in 6.29, 6.31, 6.32, 6.35, E6.30, F6.44, W6.48, P6.50, F6.51, F6.52
EL3	Poor conserved in sequence and length
TMHVII	W7.40, G7.42, Y7.43, S7.46, R7.49, P7.50, Y7.53
C-terminus	Poor conserved in sequence and length

Most of these residues are found in the regions of the receptors most likely to be highly involved in various aspects of receptor function, especially ligand binding and coupling to G-proteins.

Residues that are highly conserved among all Class A receptors could play important roles in either the maintenance of the proper structure or folding of all Class A receptors, or in signalling mechanism. However, a few residues are conserved only within subsets of Class A receptors (i.e. D3.32 in mammalian biogenic amine receptors and F6.51 and F6.52 in mammalian serotonin receptors).

In the following table, the most relevant mutations in 5-HT receptors and their effect on ligand binding are presented

Table IV Mutagenesis at 5-HT serotonin receptor: effects on ligand binding

	Mutation	Effect on ligand binding
5-HT$_{1A}$	D2.50 (Chanda P.K. 1993)	Abolished
	D2.50R (Ho B.Y. 1992)	Reduction on 5-HT, no effect on pindolol
	D3.32 (Ho B.Y. 1992), (Wang C.D. 1993; Kristiansen 2000), (Manivet P. 2002), (Mialet 2000; Mialet 2000), (Boess 1998)	Abolish 5-HT binding
	S5.43A (Ho B.Y. 1992)	Decrease effect for 5-HT, no effect for pindolol
	W6.48	Reduced antagonist binding, no effect on 5-HT
5-HT$_{1B}$	F4.61A (Granas 2001)	Modest effect in antagonist binding
	S5.43A (Granas 1998)	Effect on5-carboxamidotryptamine (5-CT), no effect for 5-HT
	S6.55A (Granas 1998; Granas 1998)	Increased affinity for 8-OH-DPAT
	T7.39N (Oksenberg D. Marsters S.A. 1992; Parker E.M. 1993)	Change tha pharmacology of the human receptor to be more likely to the rat 5-HT1B
	D7.36 (Granas 1999)	Reduced binding of 5-HT

Table IV Mutagenesis at 5-HT serotonin receptor: effects on ligand binding (continue from the previous page).

	Mutation	Effect on ligand binding
5-HT$_{2A}$	T1.39I, T1.46I (Roth B.L. 1993), W1.34A (Roth B.L. 1997; Roth B.L. 1997)	No effect on 5-HT, α-Me-5-HT or bufotenine, decrease affinity for Ketanserin and DOM
	D2.50N (Wang C.D. 1993)	No effect on LSD and spiperone, reduction for 5-HT, Ketanserin and mianserin
	F2.55L	No effect on 5-HT
	F2.55S (Roth B.L. 1993; Choundary M.S. 1995)	Reduced affinity for mesulergine
	M2.62L, or T2.64A	Reduced Haloperidol binding, no effect for spiperone, Ketanserin and mesulergine
	D3.32 (Ho B.Y. 1992), (Wang C.D. 1993; Kristiansen 2000), (Manivet P. 2002), (Mialet 2000; Mialet 2000), (Boess 1998)	Abolish 5-HT binding
	S3.36A (Almaula 1996; Manivet P. 2002)	Reduce 5-HT binding and N,N-dimethyl-5-HT, no effect on LSD
	S3.36C (Almaula 1996)	Decrease affinity for 5-HT, no change for LSD or N,N-dimethyl-5-HT
	S4.57A (Johnson 1997)	No effect
	S5.43A (Johnson 1997; Shapiro D.A. 2000; Shapiro D.A. 2002)	Reduction for 5-HT, less effect on 5-methxytryptamine and N-1-isopropyl-5-methoxytryptamine, LY86057, LY53857, DOI, no effect on Ketanserin, ritanserin, spiperone, great effect on α-methyl-5-HT, N-ω-methyl-5-HT and 5-methyl-DMT
	S5.45A (reverse effect in human 5-HT2C where this residue is Ala, with the mutation A5.46A) (Almaula 1996; Almaula 1996)	Decrease LSD binding, increase mesulergine binding, no effect for (±)-1-(2,5-dimethoxy-4-iodophenyl)-2-aminopropane (DOI), psilocin, bufotenine, and Ketanserin
	F5.47A (Shapiro D.A. 2000)	Increased the affinity of DOI, decreased the affinity of α-methyl-5-HT, N-ω-methyl-5-HT, Ketanserin, ritanserin and spiperone, and had no effect on the binding of 5-HT and 5-methyl-DMT
	F5.48A (Shapiro D.A. 2000)	Decreased the affinity of 5-HT, α-methyl-5-HT, N-ω-methyl-5-HT and DOI
	Alanine scan from S6.28 to V6.40 (Shapiro D.A. 2002)	No effect on DOI and 5-HT binding
	E6.30R, R3.50E the double R3.50/E6.30 mutant	Similar binding affinities for Ketanserin, the E6.30 mutant has higher affinity for 5-HT
	F6.52L (Choundary M.S. 1993)	No effect on Ketanserin, abolish DOI and mesulergine

	F6.51L (Choundary M.S. 1993)	Marked effect on Ketanserin, no effect on 5-HT
	W6.48A (Roth B.L. 1997; Roth B.L. 1997)	Marked effect on agonist and antagonist
	Y7.43 (Roth B.L. 1997; Roth B.L. 1997)	Significant effects on 5-HT, Ketanserin and DOM
	W7.40 (Roth B.L. 1997; Roth B.L. 1997)	Abolish all ligand binding
	F7.56 (Roth B.L. 1997; Roth B.L. 1997)	No effect

Table IV Mutagenesis at 5-HT serotonin receptor: effects on ligand binding (continue from the previous page).

	Mutation	**Effect on ligand binding**
5-HT$_{2B}$	D2.50N, N749D (Manivet P. 2002) (the effect is reversed by the double mutation D2.50N / N7.49D)	Reduced 5-HT
	D3.32 (Ho B.Y. 1992), (Wang C.D. 1993; Kristiansen 2000), (Manivet P. 2002), (Mialet 2000; Mialet 2000), (Boess 1998)	Abolish 5-HT binding
	A4.57S (Manivet P. 2002)	No effect
	S5.43A (Manivet P. 2002)	No effect on 5-HT
	T5.49A / A5.52T (Manivet P. 2002)	Decrease on 5-HT
	N6.55S (Boess 1998)	Decrease on 5-HT
	D7.36 (Manivet P. 2002)	Decrease on 5-HT

Table IV Mutagenesis at 5-HT serotonin receptor: effects on ligand binding (continue from the previous page).

	Mutation	**Effect on ligand binding**
5-HT$_4$	T4.54V (Mialet 2000; Mialet 2000)	No effect
	D3.32 (Ho B.Y. 1992), (Wang C.D. 1993; Kristiansen 2000), (Manivet P. 2002), (Mialet 2000; Mialet 2000), (Boess 1998)	Abolish 5-HT binding
	P4.53S (Mialet 2000; Mialet 2000)	Abolish binding of antagonists
	S5.43A (Mialet 2000; Mialet 2000)	Decrease binding for GR113808 and 5-HT
	A6.54L / N6.55S (Boess 1998)	Increased affinity for sumatriptan
	F6.52A (Boess 1998)	Decreased affinity for methiotepin
	Y7.43 (Mialet 2000; Mialet 2000; Rosendorff A. 2000)	Effect on agonist and antagonist binding
5-HT$_6$	W3.28F (Boess 1998)	Reduction
	D3.32 (Ho B.Y. 1992), (Wang C.D. 1993; Kristiansen 2000), (Manivet P. 2002), (Mialet 2000; Mialet 2000), (Boess 1998)	Abolish 5-HT binding
	T5.43A (Boess 1997)	Reduced affinity for LSD, 5-HT, ergotamine and lisuride, no effect on metergoline, methysergide, mesulergine, methiothepine, Clozapine, ritanserin, amitriptyline, and mianserin

The 5-HT_2: 5-HT_{2A} and 5-HT_{2C} receptors

5-HT_2 receptors were one of the first populations to be identified. As early as 1954, it was evident that subtypes of the 5-HT_2 receptors existed. Since then, different 5-HT_2 subtypes were identified (Gaddum J.H 1954; Gaddum J.H 1955; Gaddum J.H 1957; Peroutka S.J. 1979; McKenna D.J. 1990; Kursar J.D. 1994), until the modern classification into 5-HT_{2A}, 5-HT_{2B} and 5-HT_{2C} was achieved.

Numerous 5-HT_2 ligands, agonists and antagonists, had been developed and the 5-HT_2 was being extensively studied, and after the discovery of different subfamilies, the research focused on the discovery of agents able to discriminate between the 5-HT_2 receptors subtypes.

The following sections will be focussed in the description of the pharmacological and physiological characteristics of the serotoninergic receptors 5-HT_{2A} and 5-HT_{2C}, which are the subject of the research performed during this study. Also, a description of the main agonists, antagonists and inverse agonists for each one of these two subtypes will be provided. Therefore, a short paragraph will introduce and remind the pharmacological concepts of agonism, antagonism, and inverse agonism.

The allosteric ternary complex model and Agonism, Inverse agonism and antagonism

For GPCRs, there is a well-established model that assumes the active form of the receptor as a ternary complex involving the agonist, the receptor and the G-protein: the so-called ternary complex model. Several constitutively activating mutations (i.e. mutations that lead to basal receptor activation in the absence of agonist) have been identified for various GPCRs (reviewed in (Seifert R. 2002), as cited in the section 2.1.1.

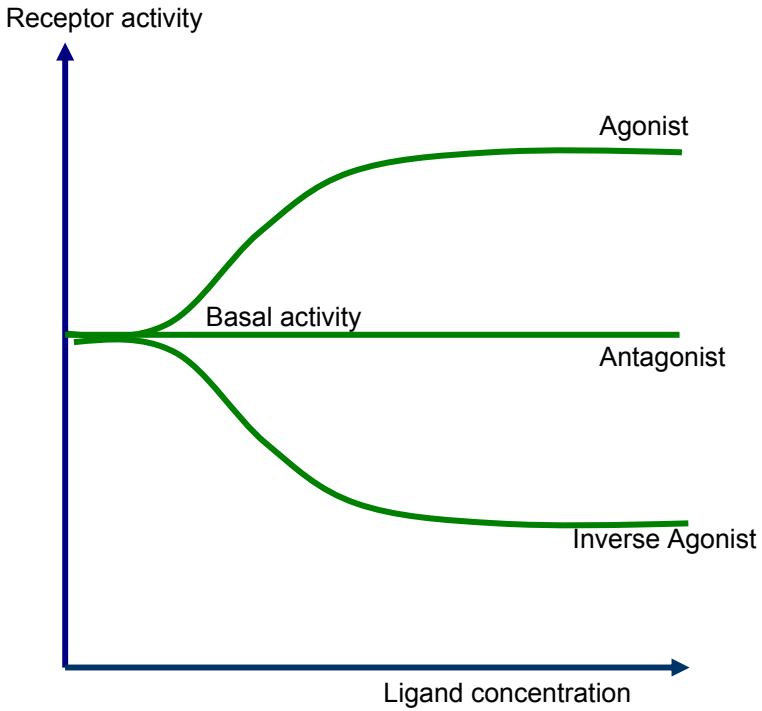
Studies on the properties of constitutively active mutant GPCRs contributed to a deeper understanding of the underlying activation mechanisms. For example mutation of the Cys6.34 in the C-terminus was found to increase constitutive activity of the 5-HT_{2A} receptor (Shapiro D.A. 2002).

The allosteric ternary complex model (Samama P. 1993) was developed to interpret the molecular properties of the mutant receptor. According to this model, receptors are assumed to exist in equilibrium between 2 states, R and R*, and there is a structural constraint in R so that only R* can effectively interact with G-proteins (Samama P. 1993). The binding of a ligand could have different effects on this equilibrium: the equilibrium could be shifted towards R, R* or could be not modified.

Therefore the ligands that interact with a specific receptor could be classified into three classes, depending on their effect: agonists, inverse agonist and antagonists (Figure 2.1.2.11).

The binding of an agonist shifts this equilibrium toward R* and lead to the formation of a high affinity agonist-receptor (R*) – G protein ternary complex. Among the agonists, further classification could be done between “full agonists” and “partial agonists”, depending on the percentage of receptors activated. Antagonists maintain the existent equilibrium between R and R*, while avoiding the activation by the agonist. Instead inverse agonist reduces the basal activity of the receptor population, shifting the equilibrium toward R.

Figure 2.1.2.11. Agonist, antagonist and inverse agonist. Effect on the basal activity of receptor population



2.1.3.Pharmacological relevance

Serotonin subtype 2 receptors play an important role in the regulation of different physiological processes, such as: mental state, food intake, sexual behaviour, sleep and wake cycle, thermoregulation, cardiovascular functions, migraine, hallucinogens effects, and aggressive behaviour. For their

predominant role in the regulation of the mental state, they are a common pharmacological target in the therapy of mental diseases. In fact, several 5-HT_{2A} antagonists are currently in clinical trials as potential antipsychotic agents, addressing diseases such as schizophrenia, mania or mood disorders ("Introduction", "Schizophrenia: ethiology and symptoms").

The importance of serotonin receptor for the therapy of schizophrenia is well known since several years (Kapur and Remington 1996), but the exact role of serotonin and other receptor systems is still partially unknown. In fact, for example, many pharmacological functions once attributed to 5-HT_{2A} receptors, on the basis of their being produced by DOB or DOI and / or their being antagonized by Ketanserin and related agents, and may actually involve a 5-HT_{2C} (or 5-HT_{2B}) mechanism.

Another example is the hypothermic activity of a series of phenylisopropylamines: once attributed with their 5-HT_{2A} affinity, was then shown to be correlated also to 5-HT_{2C} affinity (Glennon R.A. 1990). Several atypical antipsychotic agents bind at 5-HT_{2A} and 5HT_{2C} receptors; however, there is still confusion about the correlation between their atypical nature and binding (Roth B.L. 1994).

In fact, recent molecular research has indicated also 5-HT_{1A} and 5-HT_7 (Ikeda, Iwata et al. 2006; Wood, Scott et al. 2006) as possible new targets for antipsychotic atypical agents. The more recent hypothesis support the idea that inhibition of different serotonin and dopamine receptors in the correct degree could produce the ideal "cocktail" for the therapy of schizophrenia and other psychosis.

2.2. Computer-assisted drug discovery methods

The discovery and commercialization of new drugs is known to be a time consuming and cost intensive process (Ooms F 2000) and Computer Aided Drug Discovery (CADD)

is one of the tools that can be used to facilitate and speed up such process.

Strategies for CADD depend on the extent of information available regarding the target and the ligands and have been traditionally divided into two groups.

The first group (*direct approaches*) is focused on understanding the interactions in receptor-ligand complexes and is usually applied when the structure of the biological target is known, while the second group (*indirect approaches*) is centred on the structure-activity problem in the absence of detailed structural information about the receptor.

Both approaches were recently reviewed (Agrafiotis DK 2002; Congreve M 2005; Orry JWA 2006), in some cases with a special focus on GPCRs ligands discovery (Ashton M 2002; Lowrie JF 2004).

2.2.1. Direct and indirect modelling in drug discovery methods

Indirect approaches

The indirect approaches in drug design are used whenever the structure of the pharmaceutical target is unknown. In this case, as there is uncertainty regarding the active site, the discovery of new compounds with the desired characteristics for that target or the explication of some pharmaceutical properties has to be determined by studying series of known ligands for the same target.

This approach comprises different techniques, of which the most used are QSAR, molecular similarity / diversity technique and combinatorial chemistry

Quantitative Structure – Activity Relationship (QSAR)

The first quantitative structure-activity relationship (QSAR) techniques were developed nearly 40 years ago. They

were based on the idea that the activity of a ligand was a function of its structure, depending on its electronic, hydrophobic and steric properties. This methodology (Kubinyi H 1976) since then was applied.

Any QSAR method can be generally defined as the application of mathematical and statistical methods to the problem of finding empirical relationships (QSAR models) of the form:

$$P_i = F(D_1, D_2, D_3 \ldots D_n)$$

where P_i are biological activities (or the properties of interest) of molecules 1, 2, 3…., n, and D_1, D_2, D_3, …, D_n are the structural descriptors (calculated or experimentally measured) for each molecule and *F* is an empirical mathematical function that should be applied to the descriptors to obtain the property values for all molecules.

The goal of QSAR modelling is to establish a trend in the descriptor values that correlates with the biological activity. All QSAR approaches imply, directly or indirectly, a simple similarity principle: compounds with similar structures are expected to have similar biological activities.

Since Hansch initial work (Hansch C 1977), many different approaches to QSAR have been developed. One of the most used approaches is based on descriptors that consider the three-dimensional representation of the molecular structures, which is usually known as 3D-QSAR, or CoMFA method (Comparative Molecular Field Analysis (Cramer RD 1998). CoMFA allows the evaluation of the influence of different structural features on the behaviour of a molecule in the active site. CoMFA (Cramer RD 1998) has combined the power of 3D molecular modelling and partial-least square (PLS) regression analysis and has been successfully applied in medicinal chemistry and toxicity analysis (Kubinyi 1993).

In 3D-QSAR, a grid is placed around the collection of superimposed molecules. For each molecule in the ensemble,

interaction energies are calculated using a probe placed at each grid point in turn. This interaction energy is composed by a Van der Waals term and an electrostatic term. PLS regression analysis is then used to handle the hundreds of variables that are calculated.

The analysis can be done graphically and the grid points surrounding the collection of active molecules can be coloured according to the coefficients: in this way cluster of points can be used to identify regions of common steric or electrostatic energy and thus provide clues about the most important regions in determining activity.

The main limitation of both CoMFA and the GRID/GOLPE method is the requirement of 3D alignment of the ligands according to a pharmacophore model (Cramer RD 1998) or based on ligand docking to a receptor-binding site.

A similar approach, the GRID / GOLPE method (Baroni M. 1993) is based on the calculation of Molecular Interaction Fields (MIFs) using the molecular mechanic program GRID (Goodford PJ 1985), and allows statistical pretreatment of the MIFs.

Other methodologies that rely on the 3D-QSAR approach and that are devoted to overcome some limitations of the method include the Smart Region Definition (SRD) (Pastor M 1997). SRD constituted an improvement of the predictive ability and interpretability of the 3D QSAR. The method relies on the idea that changes in the structure of the compounds will induce changes in the field variable, and these changes will be reflected in a group of field variables in neighbour regions. This methodology is able to extract groups of neighbour variables that represent the same information, and therefore leads to more stable and easier to interpret models.

The GRid – INdependent descriptors (GRIND) (Pastor M 2000) methodology constitutes an original approach to overcome the problem of alignment of compounds that is the first step before performing the 3D-QSAR analyses. GRIND

comprises a novel class of molecular descriptors that are obtained from MIFs by simplifying and transforming the fields into alignment – independent variables using a particular type of autocorrelation function. GRIND have been used in many different cases for building 3D-QSAR models of comparable quality to those obtained with classical 3D-QSAR methods.

Direct approaches

The direct approaches in drug discovery, also known as Structure Based Drug Design (SBDD), is applied whenever the 3D structural information for the biological target is known, as reminded before (“2.2 Drug discovery methods”). This methodology combines information from several fields, such as X-ray crystallography and/or NMR, molecular modelling, synthetic organic chemistry, QSAR, and biological assays. The experimental methods that are able to solve the three-dimensional structure of the target have rapidly evolved in the last years, and the number of entries in protein structure databases like the Protein Crystallographic Database (PDB) (Sussman, Lin et al. 1998) has greatly increased. Despite the amount of structural information available, still it is not an easy task to understand the exact interaction mode for ligands. Biological target are often crystallized with the natural agonist or without ligand, and few structural information is available regarding the binding of other ligands such as antagonists or inverse agonists. In addition, the X-ray structure could present biases due to the crystallization methodology used, and it represents only one of all the possible low energy conformations for that protein. Nevertheless, the knowledge of the 3D structure of the biological target can provide useful information to guide the discovery of potential new drugs. In addition, molecular modelling techniques could speed up the drug discovery process by integrating data from different experimental sources and predicting protein structures and binding modes for known and new ligands. Methodologies that

rely on the 3D structure of the target include 3D databases / 3D searching and the modelling of 3D protein structures.

Several databases are available, such as the Cambridge Structural Database (CSD) (Allen 2002) or the Brookhaven Protein Databank that collects the crystal coordinates of proteins and other large macromolecules (37.874 in July 2006). The knowledge of the three-dimensional orientation of the key regions in the active site is crucial for drug discovery and is often the first step of the process.

Nevertheless receptor active sites are complex, and the residues side chains in their geometrical features and often flexible conformations. Therefore, even knowing the orientation of the residues in the active site, drug discovery of new compounds is not a trivial task.

Anyway, combining this information with indirect computational methods, to define and develop a pharmacophore model for a specific biological target for example, allows to speed up the process.

2.2.2.Homology modelling of protein target

When the 3D structure of the biological target has not been solved by experimental techniques, computational molecular modelling may be used to obtain a suitable approximation of the protein structure.

As well as other computational methods, also protein-modelling methods have greatly evolved in the past few years, leading to models with higher predictive ability. Depending on the type of protein to be modelled (membrane or soluble protein) and on the amount of data available (number of crystallized proteins of the same family and sequence identity between such proteins and the target), different methods could be applied to predict the 3D structure of a protein: topology modelling, homology / comparative modelling, and ab initio modelling methods.

In contrast to the first methods cited, de novo or ab initio methods predict the structure from sequence alone without relying on a 3D template. This method still remains the less reliable to predict the 3D structure of a protein with an acceptable accuracy, and it will not be described here, while the others will be described shortly in the next section.

Topology modelling, homology modelling and comparative modelling

The expression “topology modelling” is usually applied to build membrane proteins and especially GPCRs, before the discovery of Rhodopsin crystal structure (Palczewski K. 2000). Before then, there was no template crystallized for GPCR modelling and therefore it was necessary to guess the transmembrane regions and the orientation of secondary structure elements on the basis of the few experimental info available.

In 1994, Ballesteros and Weinstein proposed a specific protocol for topology building of GPCRs (Ballesteros J. 1994), using the electronic density map of bovine Rhodopsin as a template to organize the orientation of the transmembrane helixes. In all the other cases, for membrane or soluble proteins with homology at least higher than 30% between the target and the template and with availability of numerous experimental data (for example crystallographic structure of more than one template), other modelling techniques are usually preferred such as the “homology modelling” or “comparative protein structure modelling”.

Homology modelling or comparative modelling rely on the similarity between the sequence and at least one known structure (Marti-Renom M.A. 2000), and still remains the most used and reliable method to predict the 3D structure of a protein with an accuracy that can be comparable to a low-resolution, experimentally determined structure.

This method has been assessed through the CASP (Critical Assessment of Structure Prediction) that monitored in the past ten years the state of the art in modelling protein structure from sequence (Moult J 2005).

Homology or comparative modelling is carried out in four sequential steps:

1. Identification of known structures (templates) related to the sequence to be modelled (target)
2. Alignment of the target sequence with the templates and fold assignment
3. Building of the model
4. Assessment of the model

1. Identification of known structures (templates) related to the sequence to be modelled (target)

The identification of templates related to the target sequence has to consider that comparative modelling is only applicable when sequence identity between the target sequence and the templates protein structure is at least a 30%, as below this cut-off the accuracy of models decreases sharply, mainly as a result of a rapid increase of the errors in the alignment step (Chothia C 1986; Sander C 1991).

The identification of suitable templates is usually carried out by means of different alignment methods, such as the pairwise alignment methods BLAST (Altschul and Koonin 1998) and FASTA (Pearson and Lipman 1988), profile-sequence alignment methods, such as PSI-BLAST (Altschul, Madden et al. 1997), profile-profile alignment methods, such as Hidden Markov Models (Eddy R.S. 1998), and sequence-structure threading methods (Jones, Taylor et al. 1992) that can sometimes reveal more distant relationship than purely sequence-based methods.

All these methods give as output an alignment between the target sequence and the templates: the similarity between the target and the template and the reliability and accuracy of the alignment are expressed by a score that allows an easy comparison between alternative alignments for the same sequences.

Once suitable templates have been identified, the alignment obtained could be refined by other alignment programs, in order to assign the folding of the target sequence.

2. Alignment of the target sequence with the templates and folding assignment

The alignment obtained in the previous step could be refined using other programs, like CLUSTALW (Thompson J.D. 1994) or CLUSTALX (Thompson, Gibson et al. 1997), that take into consideration the secondary structure elements and the folding of the templates.

Therefore CLUSTALX program allows aligning the target and the templates using two different criteria: the sequence similarity between the proteins and the secondary structure organization, in order to avoid the gap opening within secondary structure elements.

Alignment errors are often the first cause of important distortion in the structure of the resulting models, therefore great attention should be put on this phase of the modelling process.

Secondary structure prediction methods could also be used to identify sequence patterns repeated within the family, and therefore the end or start residues for secondary structure elements or aminoacids important for the function of the protein. Among the most widely used programs for secondary structure prediction, there are PHD (Rost B 1996), PSI-PRED (McGuffin, Bryson et al. 2000; Bryson, McGuffin et al. 2005) and JPRED (Cuff JA 1999).

PHD uses the sequence information from the protein family and the profile derived from multiple alignment to perform the prediction through a neuronal network system. JPRED is a consensus method based on a multiple sequence alignment which measures the conservation of the aminoacids in the given target sequence when compared with the other sequences in the alignment (Cuff, Clamp et al. 1998)

If the target sequence belongs to transmembrane protein family, transmembrane region prediction methods are also recommended, such as TMAP (Persson and Argos 1997), Swiss-Prot (http://www.expasy.ch) and TMPRED (http://www.ch.embnet.org), and in this case, also other criteria could be helpful in manual refinement, like hydrophobic and hydrophilic profile, and position of arginines and lysines.

3. Building of the model

Once the structural templates have been selected and their sequences have been aligned with the target sequence, the next step is building of the model by transferring the coordinates of the template to the target sequence.

Comparative protein modelling produces an all-atom model of the target sequence by modelling the backbone of the protein, the loops, and the side-chains. In the original comparative approach, a model is constructed from a few template core regions, and from loops and side chains obtained from either aligned or unrelated structures (Browne W.J. 1969; Greer J. 1981; Blundell T.L. 1987).

More recent comparative modelling methods use distance geometry or optimization techniques to satisfy spatial restraints obtained from sequence-template alignment (Havel T.F. 1991; Sali A. 1993; Kolinski A. 2001). In general, when the percentage of sequence identity is high enough and many crystallized templates are available, it is possible to transfer the coordinates from the templates to the target.

Anyway, such ideal conditions are rare, and often there is no template available for some part of the target, therefore other methods or less related sequences should be used. Usually the loops and the sidechain conformation present the highest level of uncertainty, and their modelling requires special attention.

Loop modelling

Target sequences often have regions that are structurally different from the corresponding regions in the templates, and this occur mostly in the surface loops, thus no structural information about these segments can be extracted from the template structures. Loops often play an important role in defining the functional specificity of a given protein, forming the binding site, or in ligand recognition as it is the case of the second extracellular loop in GPCR Class A and class B receptors (Massotte D 2005).

Therefore accuracy of loop modelling can be an important factor determining the usefulness of comparative models in applications such as ligand docking. Loops are generally too short to provide sufficient information about their local folds and even identical decapeptides in different proteins do not always have the same conformation.

Some additional restraints are provided by the core anchor regions that span the loop and by the structure of the rest of a protein. Many loop modelling methods have been described but modelling correctly loops longer than 8-10 residues is still challenging (Moult J. 1986; Bruccoleri R.E. 1987; Shenkin P.S. 1987).

Loop modelling methods could be divided roughly in two main classes: (i) database search approaches that scan a database of all known protein structures to find segments fitting the anchor core regions; (ii) conformational search approaches that rely on optimizing a scoring function.

Among the available loops database we will mention the ArchDB, that structurally classifies the loops of known protein structures (Espadaler J. 2004; Fernandez-Fuentes N 2005).

Regarding conformational methods, there are different methods available based on objective functions and optimisation algorithms. The search algorithms could include molecular dynamics simulations, genetic algorithms, Monte Carlo and simulated annealing (Skolnick, Kolinski et al. 1988).

Sidechain modelling

Two simplifications are frequently applied in the modelling of sidechain conformations. First, aminoacid residue replacements often leave the backbone structure almost unchanged (Browne W.J. 1969), fixing the backbone during the search for the best sidechain conformations. Second, most sidechains in high-resolution crystallographic structures can be represented by a limited number of conformers that comply with stereochemical and energetic constraints (Janin J. 1978).

These two simplifications allowed the development of different libraries of sidechains rotamers, which are used in some comparative modelling approaches.

As it could be easily deduced by the previous discussion, building reliable predicted models is not trivial, and manual refinement is often needed: as a final step of the modelling process, it is useful to perform manual reorientation of important residue side chains followed by energy minimization and molecular dynamic simulations in order to relax the structure.

Such steps are time consuming and sometimes a source of human errors, but at the same time are essential before proceeding with further studies, like docking simulation, that are highly dependent on the starting side chain conformation. This issue will be fully treated in the following sections.

The most widely accepted approach in order to avoid human errors is the use of fully automated methods for building and validation of the models. Automatic servers are available for the modelling of target sequences, but often the results are not very accurate and allow only the prediction of the general fold.

Among the automatic methods, one of the most used computer programs is MODELLER: a suite of programs for comparative structure modelling (Sali A. 1993; Fiser A. 2000). MODELLER input in the simplest case comprises the alignment of the target sequence to be modelled with those of the template structures, the atomic coordinates of the templates, and a short script file. MODELLER fast and automatically calculates a model containing all non-hydrogen atoms.

This program performs comparative protein structure modelling by satisfaction of spatial restraints that include:

1. Homology derived restraints on the distances and dihedral angles in the target sequence

2. Stereochemical restraints, such as bond length and bond angles preferences, obtained from the CHARMM-22 molecular mechanics force field

3. Statistical preferences for dihedral angles and non-bonded distances, obtained from a representative set of known protein structures

4. User defined restraints, such as those from NMR spectroscopy, rules of secondary structure packing, cross-linking experiments, fluorescence spectroscopy, image reconstruction from electron microscopy, site-directed mutagenesis, and intuition.

The spatial restraints are expressed as probability density functions, and are combined into an objective function that is optimized by conjugate gradients and molecular dynamics with simulated annealing.

4. Assessment of the model

Once the models are built, and before using them to drug discovery or other purpose (studies of stability, protein-protein interactions, enzymes catalysis mechanism etc etc), it is essential to validate the 3D structure in terms of two parameters: the accuracy of the models and the geometric / physico - chemical values.

Regarding the accuracy, it is important to bear in mind that the quality that can be expected is related to the percentage of sequence identity between the target and the templates. Usually, models are divided in high (50% sequence identity), medium (30% - 50%) and low accuracy (below 30%) comparative models.

Alignment errors increase rapidly in low accuracy models and become the most significant source of errors.

In the model assessment phase it is also important to consider the physico-chemical and geometrical parameters such as the hydrophobic and hydrophilic profile, distances, angles and dihedrals distribution (that should follow the Ramachandran plot), as well as the orientation of important residue that must agree with experimental data reported in the literature (for example conserved residues, mutagenesis experiments etc)

There are various programs available for assessing the geometric quality of the models, and among the most used validation suites, there are PROSA (Sippl 1993) and PROCHECK (Laskowski R.A. 1993). PROSA detects misfolded proteins using a knowledge based mean field to analyze the energy distribution of the protein, while PROCHEK suite of programs provides a detailed check on the geometry

and stereochemistry of a protein structure (such as covalent geometry, planarity, dihedral angles, chirality, non bonded interactions etc etc)

Geometry optimization and dynamic behaviour of the models: energy minimization and molecular dynamics simulations

Geometrical optimization methods relax the structures and improve the side chains packing allowing to obtain the conformation corresponding to the nearest local minimum in the energy landscape. Molecular dynamic simulations allow to explore conformational states other than the local minimum found with the first minimization step and also to study the interactions stabilizing different ligand-protein complexes within the time scale affordable in such methodology.

The easiest geometrical optimizations (in term of simplicity of the model and time of the calculation) make use of molecular mechanics, which treats a molecule as a collection of atoms whose interactions could be described by Newtonian mechanics. Atoms are essentially a collection of billiard balls, with classical mechanics determining their positions and velocities at any moment in time.

The basic assumption underlying molecular mechanics is that classical physical concepts can be used to represent the forces between atoms. In other words, one can approximate the potential energy surface by the summation of a set of equations representing pairwise and multibody interactions. Equilibrium parameters have been developed for most of the biological interactions, including distances, angles, dihedrals and improper dihedrals and distortions from these equilibrium parameters are penalized.

Regarding the interactions between bonded atoms, usually three laws are used to penalise the distortion of bond distances, bond angles and dihedrals from their equilibrium value. Pairwise interactions are often represented by a harmonic potential like the potential represented by Hooke’s

law (derived for a spring) for bonded atoms, restoring the bond distance to and equilibrium distance b_0:

$[1/2\ K_b(b-b_0)^2]$

Similarly, distortion from an equilibrium valence angle (θ_0) describing the angle between three bonded atoms sharing a common atom is also penalized:

$[1/2\ K_\theta\ (\theta-\theta_0)^2]$

A third class of interaction dependent on the dihedral angle φ between four bonded atoms is the torsional potential, used for account for orbital delocalization and to compensate for other deficiencies in the force field:

$K_\varphi\ [1 + \cos(\varphi - \delta)]$

A harmonic term is often introduced for dihedral angles ξ that are relatively fixed, such as those in aromatic rings:

$[1/2\ K_\xi\ (\xi-\xi_0)^2]$

Regarding the interactions between no bonded atoms, the Van Der Waals and Coulomb contributions are taken into consideration, which are the simplest approaches to account for steric and electrostatic term to the general potential surface V.

Van der Waals potential for two nonbonded atoms *i* and *j*:

$[C^{12}(i,j) / r_{ij}^{12} - C^6(i,j) / r_{ij}^6]$

Coulomb's law for the two atoms *i* and *j*:

$[q_i\ q_j / (4\pi\varepsilon_0\varepsilon_r r_{ij})]$

The potential surface V, on which the force field is based, includes all the terms presented above and describes the energy of a molecule by a combination of all these forces:

$$V = \sum ½\ K_b(b-b_0)^2 + \sum ½\ K_\theta(\theta-\theta_0)^2 + \sum ½\ K_\varphi\ [1 + \cos(\varphi - \delta)] + \sum ½\ K_\xi(\xi-\xi_0)^2 + \sum ½\ [C^{12}(i,j)/r_{ij}^{12} - C^6(i,j)/r_{ij}^6] + \sum q_i q_j/(4\pi\varepsilon_0\varepsilon_r r_{ij})$$

One of the energy terms most difficult to represent is the electrostatic term. In most force fields, the electronic distribution surrounding each atom is treated as a monopole with a simple coulombic term for the interaction. The effect of the surrounding medium is generally treated with a continuum model by use of a dielectric constant.

Electrostatic interactions range from those operating only at very short distances that are non-specific (dispersive interaction, r^6 dependency) to those operating at very long distances with a high degree of specificity (charge-charge interactions, r^1 dependency).

The energy of interaction between two charges q_1 and q_2 is given by Coulomb's law:

$E = q_1 q_2 / 4\pi\varepsilon r^{12}$

Where r^{12} is the distance separating charges and ε is the dielectric constant of the medium.

To evaluate atom-atom interaction using the Coulomb's law, charges are considered as a point, a monopole (net atomic charge).

Another difficult aspect of molecular mechanics is the dielectric problem and the solvation of the biological system and different approximations were developed to face this limitation. Simple or more sophisticated approaches could be applied depending on the purpose of the research and an

exhaustive treatment of this issue is beyond the scope of the dissertation.

Energy Minimization and Molecular Dynamics Simulations

Energy Minimization

Energy minimization algorithms have been characterised with regard to their convergence properties, but, in general, they only locate the closest local minima to the starting geometry of the system. A stochastic approach to the starting geometries can be combined with minimization to find a subset of minima in the hope that the global minimum is contained within the subset and can readily be identified by its potential value.

Among the most used algorithms for energy minimization of a system, there are the steepest descent and the conjugate gradient methods.

The steepest descent is most used in the first steps of minimization, and it was shown to be highly effective when applied to rough conformations, that are supposed to be quite far from a local minimum. If we imagine the energy surface as a hilly landscape, then the easiest and more reliable way to find a valley is to follow the landscape downhill. The gradient method has no information about the curvature of the energy surface, therefore minimization by this method slows down as the gradient decreases. In subsequent steps of energetic refinement, conjugate gradient is preferred as it uses information from previous steps to modify the move in the next step.

The method usually has much better convergence properties then steepest descent, and it is more sensitive for the final steps of energy minimization.

Molecular dynamics simulations

Molecular dynamics is a process based on the simulation of molecular motion by solving Newton's equations of motion for each atom of the system and incrementing the position and velocity of each atom by use of a small time increment.

From physics:

$$F = ma = -\delta V / \delta r = m\delta^2 r / \delta t^2$$

Where *F* is the force on the atom, *m* is the mass of the atom, *a* is the acceleration, *V* is the potential energy function, and *r* represents the Cartesian coordinates of the atom. Using the first derivative of the analytical expression for the force field allows the calculation of the system energy as a function of the position of the atoms.

In this simulation we use numerical integration; that is, we choose a small time step (smaller than the period of fastest local motion in the system) such that our simulation moves atoms in sufficiently small increments, so that the position of surrounding atoms does not change significantly per incremental move. In general, this means that the time increment is on the order of 10^{-15} s (1 femtosecond). This reflects the need to adequately represent atomic vibrations that have a time scale of 10^{-15} s to 10^{-11} s. The total energy of the system E_{tot} is represented as the sum of kinetic energy E_{kin} and potential energy V_{pot}:

$$E_{tot}(t) = E_{kin}(t) + V_{pot}(t)$$

Where the potential energy is a function of the coordinates, $V_i = f(r_i)$ for atoms *i* to *N* and r_i represents cartesian coordinates of atom $_i$; and the kinetic energy depends on the motion of the atoms:

$$E_{kin}(t) = \sum \tfrac{1}{2} M_i V_i^2(t)$$

Where M_i is the mass of the atom *i*, and V_i is the velocity of atom *i*.

The energy is constantly redistributed because of the movements of the atoms, resulting in changes in their positions on the potential surface and in their velocities. At each iteration ($t \rightarrow t + \Delta t$), an atom *i* moves to a new position [$r_i(t) \rightarrow r_i(t + \Delta t)$], and it experiences a new set of forces.

For simulations one must be able to control the temperature of the simulation. The temperature of a system is a function of the kinetic energy, E_{kin} (t):

$$T(t) = E_{kin}(t) / 3/2\ Nk$$

Where *k* is Boltzmann's constant.

Depending on the simulation, either the pressure or the volume must be maintained constant. Constant volume is the easiest to perform because the boundary of the system are maintained with all molecules confined within those boundaries and the pressure allowed to change during the simulation.

Because molecular mechanics is empirical, parameters are derived by iterative evaluation of computational results, such as molecular geometry (bond lengths, bond angles, dihedrals) and heats of formation, compared with experimental values. In the case of bond lengths, bond angles and VDW parameters, crystallography has provided most of the essential experimental database, while quantum mechanics has provided excellent charge approximation.

Charge estimation for being used in molecular mechanics can be derived by application of one of the many different quantum chemical approaches, either ab initio or semiempirically. Quantum mechanical methods are available for calculating the electron probability distributions for all the electrons in a molecule and then partitioning those distributions to yield representations for the net atomic charges of atoms in

the molecule, either as atom centred charges or as more complex distributed multipole methods.

After these geometric optimizations, the protein structure can be used in further studies, such as docking simulations.

2.2.3.Docking studies

Ligands provoke a cascade of chemical events in the intracellular medium by means of molecular recognition mechanism with the receptor (the protein target). The rational discovery of new drugs implies the understanding of molecular recognition in terms of both structure and energetics (Boehm H.J. 1996). Docking, as a computational tool, allows the investigation of the binding between macromolecular targets and potential ligands to form noncovalent protein-ligand complexes and in general could be considered as an energy-optimization problem (Trotov M. 2001) with two components: the search and the score (Kuntz I.D. 1994).

The "search" is performed by exploring the configuration space accessible for the interaction between the two molecules, with the goal of finding the orientation and conformation of the interacting molecules corresponding to the global minimum of the free energy of binding. Reported techniques for automated docking fall into two broad categories: matching methods and docking simulation methods. Matching methods create a model of the active site, typically including sites of hydrogen bonding and sites that are sterically accessible, and then attempt to dock a given ligand structure into the model as a rigid body by matching its geometry to that of the active site (for example DOCK (Kuntz I.D. 1982; Shoichet B.K. 1993)).

In the second class of docking techniques the ligand is oriented randomly outside the protein, and explores translations, orientations, and conformations until the ideal site is found (for example AUTODOCK (Morris G.M. 1998) and

GOLD (Jones G. 1997; Verdonk M.I. 2003). Other programs, like QXP (quick explore) (McMartin C. 1997) use search algorithms derived from Monte Carlo perturbation combined with energy minimization, and allow flexibility both on the ligand and on the protein.

The “score” refers to the fact that any docking procedure must evaluate and rank the configuration generated by the search process, in order to predict the most favourable and realistic conformation for a specific docking problem. The above-mentioned programs use different scoring functions, with variable results depending on the complexity of the problem addressed. Bissantz et al. in 2000 (Bissantz C. 2000) evaluated different combination of docking / scoring functions and found that even if scoring seems to predominate docking accuracy, predicting which scoring function would perform the best is still a very difficult task. More recently, docking and scoring in virtual screening for drug discovery were reviewed (Kitchen, Decornez et al. 2004).

2.2.4.Challenges in GPCR modelling

The importance of GPCR as pharmaceutical target was outlined in the “Introduction” of the present thesis, and the most important milestones of GPCR research were recently and exhaustively reviewed (Jacoby E 2006; Klabunde and Jager 2006).

In such review and in other GPCRs research articles computational methods and 3D structure prediction appear to be fundamental tools to understand biological mechanism such as ligand-receptor recognition, activation mechanism and signalling (Weinstein H 2006). In such articles, while is prominent the knowledge of 3D receptor structure is also evident that the modelling of receptors belonging to the GPCR superfamily, still constitutes a challenging research problem, due to different limitating issues:

Lack of three-dimensional structures of GPCRs of other subfamilies than bovine Rhodopsin. The three-dimensional structure of GPCRs is extremely difficult to obtain with X-ray, as it is not trivial to crystallize enough pure amount of protein for crystallography. NMR technique could also be used but the main limitation is to obtain solution at high concentration of GPCR, as most GPCR are not soluble in the commonly used solvents for NMR. So far, only the three-dimensional structure of bovine Rhodopsin is available by X-ray crystallography at an acceptable resolution.

Low homology among GPCRs. The bovine Rhodopsin has a low sequence identity with other GPCRs, even when considering other Class A GPCRs. In general, sequence identity is not above the 30%, which makes difficult to build of high quality models.

Modelling of the loops. In general, when approaching protein modelling, the most variable regions between the target and the template are in the surface loops. This is the case also for GPCRs, whose loops are poorly conserved in sequence and length. In absence of templates, the flexibility and the degree of freedom of loops larger than 10 residues makes them difficult to model.

Modelling of explicit solvent. Nowadays, the number of GPCR computational models embedded in explicit solvent is increasing due to the higher potency of computers and the more efficiency in data storage. Therefore it is frequent to find in the literature the description of GPCR in the hydrophobic environment that mimic the lipidic membrane (Schlegel, Sippl et al. 2005), using dipalmotoyl phosphatidylcholine (DPPC), palmitoyloleoyl phosphatidylcholine (POPC) (Urizar, Claeysen et al. 2005) or dymiristoyl phosphatidylcholine (DMPC). Despite the evolution of computer programs, the lipidic environment is still difficult to take into account explicitly.

Different simplified systems were developed and protocols for large simulations with the explicit membrane model are becoming more and more popular, but the question whether it is worth to adopt this option still remains valid and depends on the research problem addressed. In some cases, such efforts are not essential to answer the scientific question proposed.

Other important limitations are related with the modelling of membrane protein in general, and the absence of experimental data, which can be used to build databases and to train programs especially for membrane proteins. For example the automatic modelling programs and the databases of available three-dimensional structures are based on biological data of soluble proteins. Therefore, they do not take into account the lipidic environment and the related physical-chemical differences in the geometrical parameters between membrane proteins and enzymes (as example of soluble proteins).

2.2.5. Models of GPCR and serotonin receptors

Together with the development and evolution of pharmacological tools and the consequent discovery of different subfamilies of GPCR, several attempts to predict the topology and 3D structure of receptors belonging to this wide family were made.

In most cases the predicted structures, since the early rough models to the more recent sophisticated models that include explicit solvent, were able to reproduce experimental data and to give significant support to drug discovery for pharmaceutical companies such as Novartis, Biofocus and Novo Nordisk (Jacoby E 2006). Such models and the results achieved constitute important milestones to be considered when approaching the modelling of GPCR. The most relevant models of serotonin receptor until the present are summarised in the following table V.

Table V. Main serotonin receptors models published.

Year	Receptor model	Reference
1991	5-HT_2	(Westkaemper R 1991; Westkaemper 1991)
1992	5-HT_2	(Trumpp-Kallmeyer S 1992)
	5-HT_{1A}	(Weinsten H 1992)
1993	5-HT_2	(Westkaemper R 1993; Zhang D. 1993)
1995	5-HT_{2A}	(Holtje HD 1995; Sealfon S.C. 1995)
1996	5-HT	(Kristiansen K. 1996)
	5-HT_{2A}	(Almaula 1996)
1999	5-HT_{1A}	(Homan EJ 1999)
2000	5-HT_{2A}	(Shapiro D.A. 2000)
2001	5-HT_{1A}, 5-HT_{2A}	(Bronowska A 2001; López-Rodríguez ML 2001)
	5-HT_4	(López-Rodríguez ML 2001; López-Rodríguez ML 2001)
2002	Aminergic receptors	(Shi L. 2002)
	5-HT_{1A}	(López-Rodríguez ML 2002)
	5-HT_{2A}	(Westkaemper R 2002)
	5-HT_{2A}, 5-HT_{2B}	(Brea J. 2002)
	5-HT_{2B}	(Manivet P. 2002)
	5-HT_4	(López-Rodríguez ML 2002; López-Rodríguez ML 2002)
2003	5-HT_{1A}	(López-Rodríguez ML 2003)
	5-HT_{2A}, 5-HT_{2B}, 5-HT_{2C}	(Rashid M 2003)
	5-HT_{2A}	(Mehler EL 2003)
	5-HT_4	(López-Rodríguez ML 2003)
	5-HT_6	(Hirst WD 2003)
	5-HT_7	(López-Rodríguez ML 2003)
2004	5-HT_4	(Rivail L 2004)
	5-HT_7	(López-Rodríguez ML 2004)
2005	5-HT_{1A}, 5-HT_6	(López-Rodríguez ML 2005; López-Rodríguez ML 2005)
2006	5-HT_{1A}	(Nowak M 2006)
	5-HT_{2A}, 5-HT_{2C}	(Brea J 2006)
	5-HT_{2A}	(Weinstein H 2006)

3. Hypothesis and objectives of the thesis

Computational techniques have been proved to be useful tools for the study of the interactions of molecules at atomic and molecular level. In fact, as explained in previous chapters, bioinformatics and molecular modelling (using both direct and indirect approaches) have been applied successfully to achieve a better insight on the mechanism involved in the drug action either at biochemical and molecular level as well as for the discovery of new drugs.

Among the most used applications of computational techniques in drug discovery, we can mention the obtention of suitable approximated 3D structures for relevant proteins (even in absence of experimental data), the development of pharmacophores for specific targets and the investigation of the binding between macromolecular targets and potential ligands.

Within this scenario, this thesis aims to **apply state-of-the-art computational techniques for addressing challenging issues regarding serotonin receptors and the design of new ligands with potential interest as antipsychotic drugs.**

In particular, the thesis aims to achieve the following concrete objectives:

1. Model the three-dimensional structure of the 5-HT2A and 5-HT2C receptors, with the highest level of accuracy and robustness. Since no experimentally obtained 3D structures for serotonin receptors were available, the first step was obtaining suitable 3D structures for all these targets, allowing to apply direct drug discovery methods.

2. Build complexes of these receptor models with the natural agonist serotonin (5HT) and with other well known ligands, studying the most important interactions at the receptor binding site. The docking of these ligands is a valuable tool for validating the models and the computational protocol. Special attention was paid to reproduce previously reported interactions between serotonin, 5-HT2A and 5-HT2C receptors.

3. Build complexes of these receptor models with a series of butyrophenones, studying the most important interactions at the receptor binding site. Among all the butyrophenones ligands, some ligands were selected as specially interesting due to their pharmacological properties, such as selectivity toward one subtype versus the other (5-HT2A versus 5-HT2C) or monophasic or biphasic inhibition of the second messenger production.

4. Study of the dynamic evolution of the ligand receptor complexes. Molecular dynamics simulations were carried out for each one of the ligands in complex with 5-HT2A and 5-HT2C receptors. Special attention was paid to the side chain conformations of some residues known to play a key role in the activation process. Molecular dynamics simulation tested the stability of the complexes, and the use of different starting points allowed a comparison of the interaction differences between them as well as between the 5-HT2A and the 5-HT2C receptors in general.

5. 3D QSAR study on the series of butyrophenones. This analysis combines direct and indirect approaches, improving the identification of the

most important interaction and the prediction of the binding affinity for new compounds.

4. Results and discussion

4.1. Modelling of 5-HT_{2A} and 5-HT_{2C} receptors

As explained before ("2.2.1 Direct and indirect approaches in drug discovery methods" and "2.2.2 Homology modelling of a protein target"), different methodologies are available at the "state of the art" level to predict the three-dimensional structure of a target protein.

In particular, in the section "2.2.2 Homology modelling of a protein target", the "topology modelling" and the "homology modelling" techniques are described with a special emphasis on the limitations of the older one (the topology modelling approach) and the progresses of the newer one (the homology modelling) that allowed to overcome them.

Currently, only the homology modelling technique with the highest level of automatization is in use, as the computer science has evolved rapidly and is now able to satisfy the necessary technical requirements. Nevertheless in the present thesis, both approaches are described, following the chronology of the research experiments, which will allow the reader to understand the progress made during the development of the 5-HT_{2A} and 5-HT_{2C} models that were carefully built, revisited, re-built with the introduction of the newest experimental data, validate and finally used to fulfil the scientific objectives of the thesis.

Therefore, the topological approach is described first, with a focus on the limitations encountered, and then, the homology modelling approach is presented, with a special emphasis on the methodological improvements introduced and the advantages of the new models.

Topology modelling of 5-HT_{2A} and 5-HT_{2C} receptors

As quoted from different sources several times through the text, the discovery of bovine rhodopsin three-dimensional structure (1F88) by means of X-ray crystallography constituted a revolution for GPCR modellers.

Therefore, the models built by topology were based on the 1F88 structure, which was the first GPCR crystal structure released in the PDB database, and the 3D general folding of the receptor was reproduced assuming that 5-HT2A and 5-HT2C should share the topology organization of Class-A GPCRs.

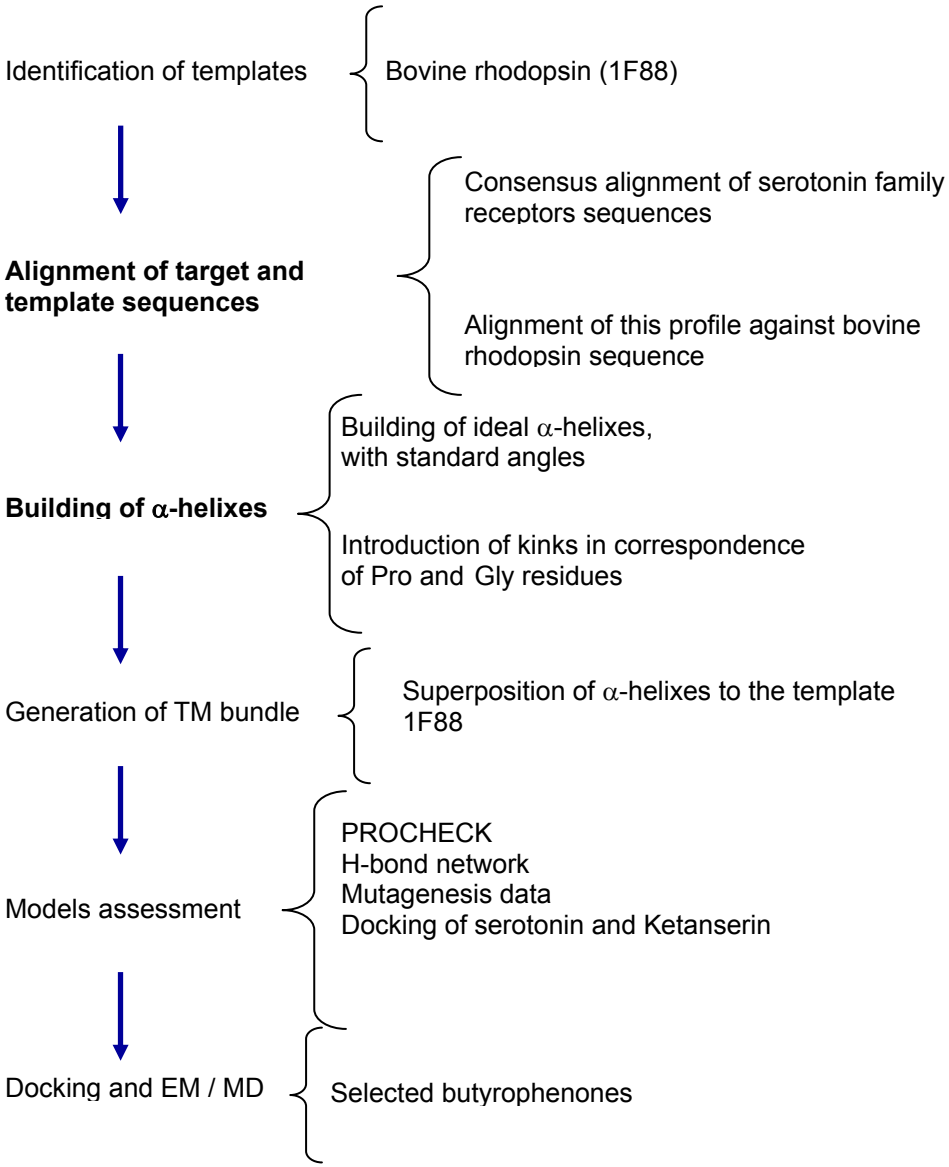
The experimental data supporting this hypothesis were the conservation of the residues within the GPCR family and the intra- and inter- helical hydrogen bond network described also in "2.1.1 GPCR superfamily", "*GPCR receptors: structural features common to all the subfamilies. The lesson learnt from the crystal structure of bovine rhodopsin*".

In brief, the topology approach was applied following six steps:

1. Identification of structural templates
2. Alignment of target and template sequences
3. Building of ideal α-helixes
4. Generation of the transmembrane bundle
5. Models assessment

After the quality check, the models were found suitable for docking, energy minimization and molecular dynamic simulations, in order to understand ligand – receptor interactions leading to the discovery of new drug candidates against schizophrenia.

Modelling by topology



1. Identification of a template

At the moment of building the models of 5-HT_{2A} and 5-HT_{2C} using modelling by topology, there was only one crystal structure of bovine rhodopsin available (1F88): the crystallization of bovine rhodopsin (Rh) constituted a revolution in the GPCR modelling, and it became soon the most widely accepted template by the majority of GPCRs modellers.

It allowed, for the first time, to verify the hypothesis about the orientation of the transmembrane helixes as could be deduced by the rhodopsin map obtained by electronic microscopy.

2. Alignment of target and template sequences:

Once the template was chosen, the second step was the analysis of the serotonin sequences available. The target sequences included were the human 5-HT_{2A} and human 5-HT_{2C} sequences, as the human receptors were the objectives of this thesis (see “Introduction” page 3).

At the moment of building the models, there were 58 serotonin receptors cloned; the serotonin receptor 5-HT_3 was not considered, as this receptor is a ligand gated ion channel (Table VI), not coupled to G-proteins to transduce the signal. All the sequences were retrieved from the Swiss-Prot database (Boeckmann B. 2003) and aligned with the ClustalX software (Thompson J.D. 1994), using the PAM250 matrix.

This multiple sequence alignment (called a “profile” of serotonin receptors) was then aligned against the sequence of bovine rhodopsin (Rh, OPSD_BOVIN). Secondary structure information derived from the crystal structure was given as input to the program in order to avoid the introduction of gaps in coincidence with alpha helixes.

Table VI. Sequence of serotonin receptor cloned at the time of development of the models

Receptor	Cloned sequences
5-HT	5-HT_BOMMO (Q17239), 5-HT_HELVI (Q25190), 5-HT_LYMST (Q25414)
5-HT$_1$	5-HT1_DROME (P20905), 5H1A_FUGRU (O42385), 5H1A_HUMAN (P08908), 5H1A_MOUSE (Q64264), 5H1A_RAT (P19327), 5H1B_CAVPO (O08892), 5H1B_CRIGR (P46636), 5H1B_DIDMA (P35404), 5H1B_FUGRU (O42384), 5H1B_HUMAN (P28222), 5H1B_MOUSE (P28334), 5H1B_RABIT (P49144), 5H1B_RAT (P28564), 5H1B_SPAEH (P56496), 5H1D_CANFA (P11614), 5H1D_CAVPO (Q60484), 5H1D_FUGRU (P79748), 5H1D_HUMAN (P28221), 5H1D_MOUSE (Q61224), 5H1D_RABIT (P49145), 5H1D_RAT (P28565), 5H1E_HUMAN (P28566), 5H1F_RAT (P30940), 5H1F_HUMAN (P30939), 5H1F_MOUSE (Q02284), 5H1F_CAVPO (O08890)
5-HT$_2$	5H2A_CRIGR (P18599), 5H2A_HUMAN (P28223), 5H2A_MACMU (P50128), 5H2A_MOUSE (P35363), 5H2A_PIG (P50129), 5H2A_RAT (P14842), 5H2B_HUMAN (P41595), 5H2B_MOUSE (Q02152), 5H2B_RAT (P30994),5H2C_HUMAN (P28335) 5H2C_MOUSE (P34968), 5H2C_RAT (P08909)
5-HT$_4$	5H4_HUMAN (Q13639), 5H4_CAVPO (O70528), 5H4_MOUSE (P97288), 5H4_RAT (Q62758)
5-HT$_5$	5H5A_HUMAN (P47898), 5H5A_MOUSE (P30966), 5H5A_RAT (P35364), 5H5B_MOUSE (P31387), 5H5B_RAT (P35365)
5-HT$_6$	5H6_HUMAN (P50406), 5H6_MOUSE (Q9R1C8), 5H6_RAT (P31388)
5-HT$_7$	5H7_CAVPO (P50407), 5H7_HUMAN (P34969), 5H7_MOUSE (P32304), 5H7_RAT (P32305), 5H7_XENLA (Q91559)

Legend: BOMMO silk moth, CANFA dog, CAVPO guinea pig, CRIGR chinese hamster, DIDMA opossum , DROME fruit fly, FUGRU japanese pufferfish, HELVIO wlet moth, HUMAN human, LYMST great pond snail, MACMU *rhesus macaque*, MOUSE mouse, RABIT rabbit, RAT rat, SPAEHE hrenberg's mole rat, XENLA african clawed frog

This first alignment was then refined manually, taking into consideration the following topics that helped the definition of transmembrane segments:

Alignment of highly conserved residues: the most conserved residues within the GPCR family were listed by Baldwin (Baldwin J.M. 1997) (Figure 4.1.12) and are though to play an important structural role within the family. Manual refinement was applied in order to make sure that these residues were aligned in all the sequences considered (bovine rhodopsin and serotonin receptors, Figure 4.1.13)

Figure 4.1.12. Conserved residues in class A GPCRs.

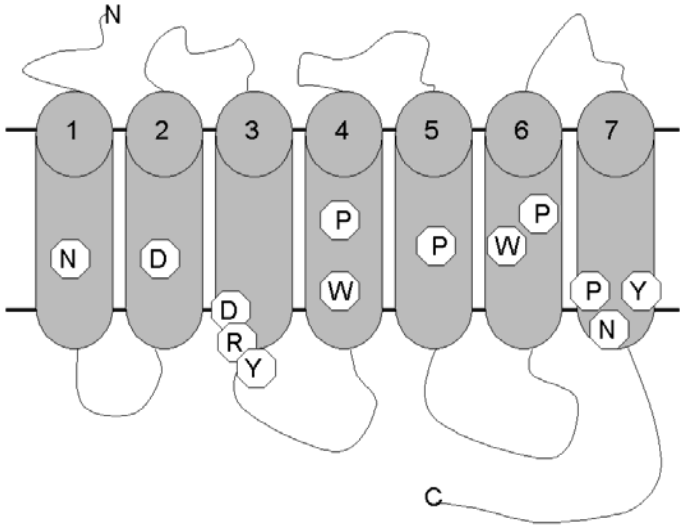


Figure 4.1.13. Alignment of human serotonin receptors sequences to the bovine rhodopsin sequence.

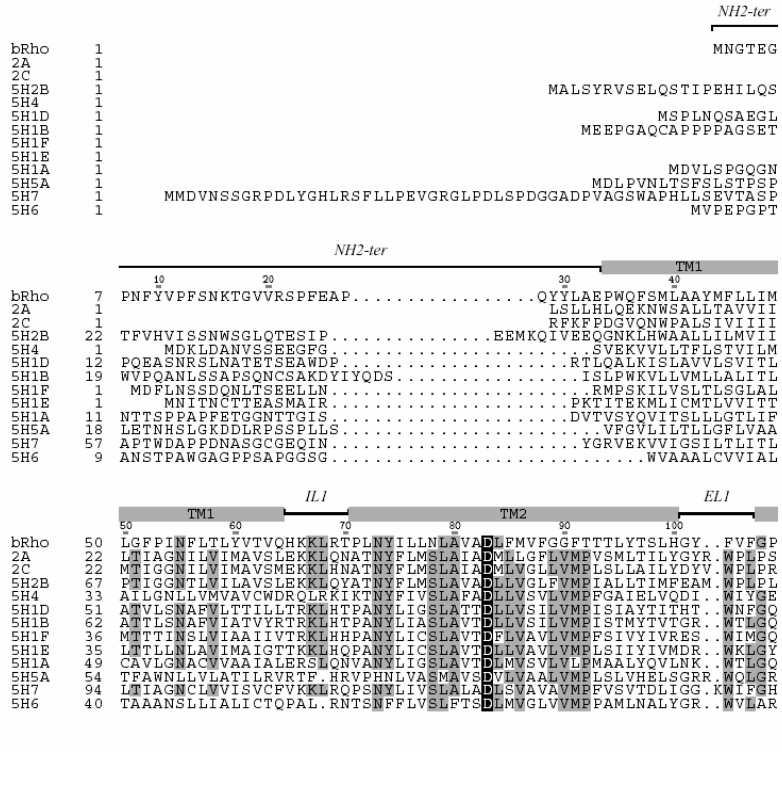


Figure 4.1.13. Alignment of human serotonin receptors sequences to the bovine rhodopsin sequence (continue from the previous page).

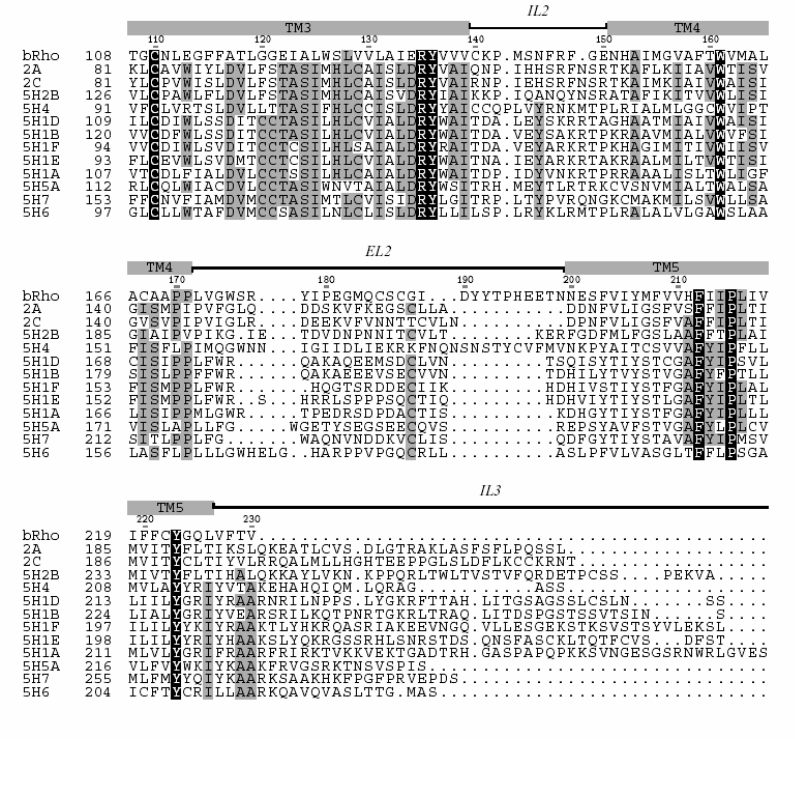
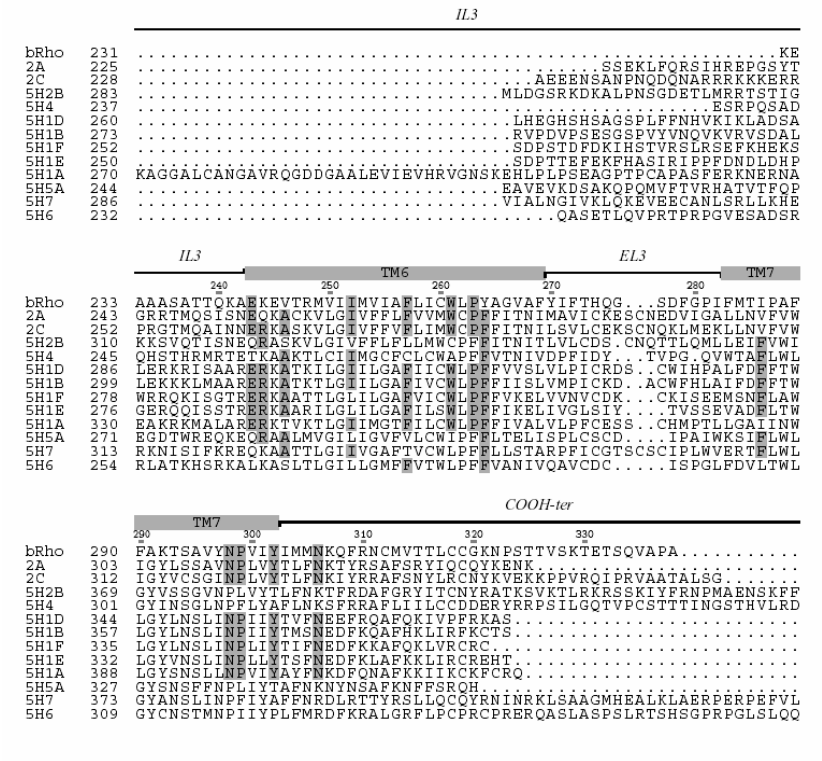


Figure 4.1.13. Alignment of human serotonin receptors sequences to the bovine rhodopsin sequence (continue from the previous page).



Prediction of transmembrane segments. The following software was used to predict the extension of transmembrane helixes: TMAP (Persson B. 1997), Swiss-Prot (web page: http://www.expasy.ch) and TMPRED (web page: www.ch.embnet.org). In order to obtain a more robust and reliable prediction, the consensus found between the three programs was used to refine the alignment.

Arg / Lys criterium and definition of helix length. The presence and position of Arg and Lys residues is a signal of the membrane boundaries (Zhang and Weinstein 1994), as their positive charge are thought to neutralize the negative charges of the phosphates groups of lipids.

In addition, alternative lengths for the helixes were considered by taking into account the experimental length of the Rh helixes (as explained in the previous section "*Alignment of target and template sequences*"), the sequence conservation and the secondary structure prediction by JPRED (Cuff, Clamp et al. 1998).

3. Building of ideal alpha helixes

Using the information gained in the previous steps by combining experimental data and predictions results, the seven alpha helixes for both 5-HT_{2A} and 5-HT_{2C} receptor were built using the Biopolymer module of InsightII. Transmembrane helixes were built using ideal geometrical parameters for alpha helixes, and kinks were introduced according to the Rh crystal structure in correlation of conserved prolines and glycines (Figure 4.1.12).

Before proceeding to build the transmembrane (TM) bundle, each one of the helixes was minimized separately with the Sander classic module of AMBER 6 (Case D.A. 1999) suite of programs, in order to optimize the side chains conformation.

4. Generation of the Transmembrane (TM) bundle

The 3D folding for 5-HT_{2A} and 5-HT_{2C} receptors was assumed to be similar to bovine rhodopsin therefore the TM bundle was built in order to reproduce the arrangement found in the crystal structure of the template (1F88).

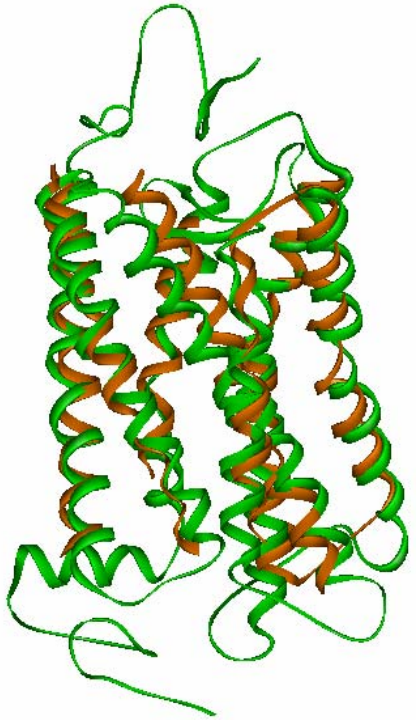
The alpha helixes of both receptors were therefore superimposed to the template (Rh). The superimposition was performed by superimposing the highly conserved residues. In fact, due the low homology (less than the 30%) between the serotonin and rhodopsin sequences, it was not possible to identify structurally conserved regions, therefore individual residues and the protein region comprise between them was used to perform the superposition.

In other words, the conserved residues identified by Baldwin were considered the boundaries of the “structurally conserved regions”.

The helixes were superimposed to Rh structure, with the Biopolymer module of InsightII (Biosym Technologies Inc., San Diego, CA, USA), taking care of minimizing the RMS between the helixes and the template.

Once the bundle was built (figure 4.1.14), the dihedrals of some conserved residues had to be manually modified to satisfy the structural restraints identified in the template. This was the case of some residues of helix 5 (5.43 and 5.45) not properly oriented towards the active site (in the inner part of the receptor).
As aminoacids 5.43 and 5.45 were shown to be essential for the binding of serotonin and other ligands (Johnson 1994; Choundary M.S. 1995), helix 5 was manually rotated in agreement with experimental data, to point residues side chain towards the binding site.

Figure 4.1.14. Bovine rhodopsin template (in green) with the 5-HT_{2A} seven transmembrane helixes (5-HT_{2A} bundle, represented in orange)



5. Assessment of the model

Before proceeding with further studies, the quality of the resulting 5-HT_{2A} and 5-HT_{2C} models was assessed by checking their geometrical and physico-chemical parameters, the reliability of the active site and the coherence with the mutagenesis data published. PROCHECK (Laskowski R.A.

1993) was used to assess the quality of bond lengths, dihedral values and angle distribution; the resulting Ramachandran plots were in agreement with good quality models (about 90% of dihedrals fell within allowed regions).

The conservation of hydrogen bond network between the conserved residues was also taken into account as well as the hydrophobic and hydrophilic distribution by using GRID probes. Molecular interaction fields were calculated with hydrophilic (WATER), hydrophobic (DRY) and phosphate group (PHO) probes. Hydrophilic and hydrophobic probes revealed the correct orientation of alpha helixes, while the phosphate probe disclosed the helixes boundaries.

Finally, the dockings of the natural ligand (serotonin) and of a known inverse agonist (ketanserin) were used to assess the usefulness of the binding site in linking structural features to pharmacological properties.

Docking, energy minimization and molecular dynamic simulations

The docking protocol span from a rigid ligand docking (only the ligand is flexible) to semi-flexible docking (that consider flexibility in the ligand and, partially, in the active site). This procedure was meant to convert the rigid and tight active site of the starting conformation, in a wider and more flexible structure able to allocate ligands larger and more bulky than serotonin.

Thus serotonin was retrieved by the Cambridge Crystallographic Data Centre (CCDC), and docked into 5-HT2A and 5-HT2C receptors active sites by means of three different docking programs, used in the following order: Autodock, QXP and GOLD (Verdonk M.I. 2003). The best position found in each one of the program was used as a starting point for the next step. Therefore the best position found with autodock, was used a reference for the docking with qxp.

The docked position reproduced the most important interactions between serotonin and the receptors, such as: the ionic interaction between the positive nitrogen and the negative D3.32 side chain and the hydrogen bond between Ser3.34 and the hydroxyl group of 5-HT. Nevertheless, these experiments were less successful in identifying other important interactions, for example the hydrogen bonds with S5.43 and S5.45.

In each one of the steps, the best docking position was chosen considering the score of the program, the population of the cluster (i.e. number of conformations belonging to the same cluster, as such repetition is considered a signal of a local minimum), and total energy of the ligand-receptor complex.

The best optimized ligand – receptor complex, outcome of the three-steps docking protocol, was then submitted to energy minimization and molecular dynamics simulation with Sander Classic (Case D.A. 1999).

Energy minimization and molecular dynamics simulations

Energy minimization and 1 ns molecular dynamic simulations of serotonin in complex with 5-HT_{2A} and 5-HT_{2C} were carried out. This first round of energetic simulations provided interesting results for serotonin in term of stability of the complexes, as shown in the following graphs for the total energy and the temperature (for the complex serotonin – 5-HT_{2A} model).

Figure 4.1.15. Serotonin in complex with 5-HT_{2A}, total energy during 1 ns molecular dynamic simulation.

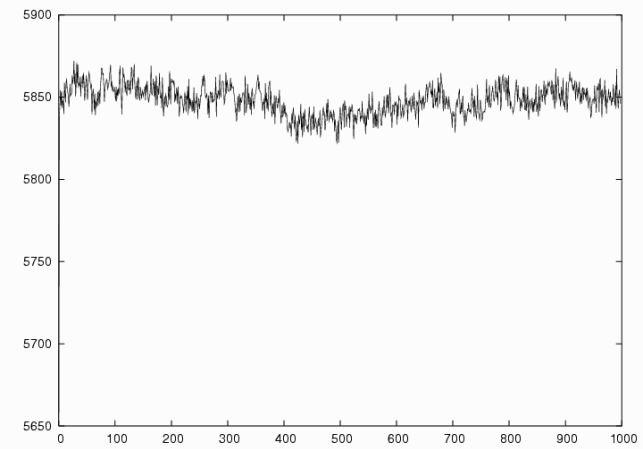
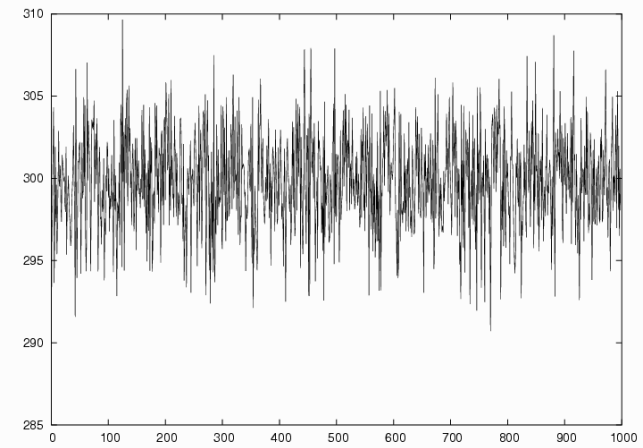
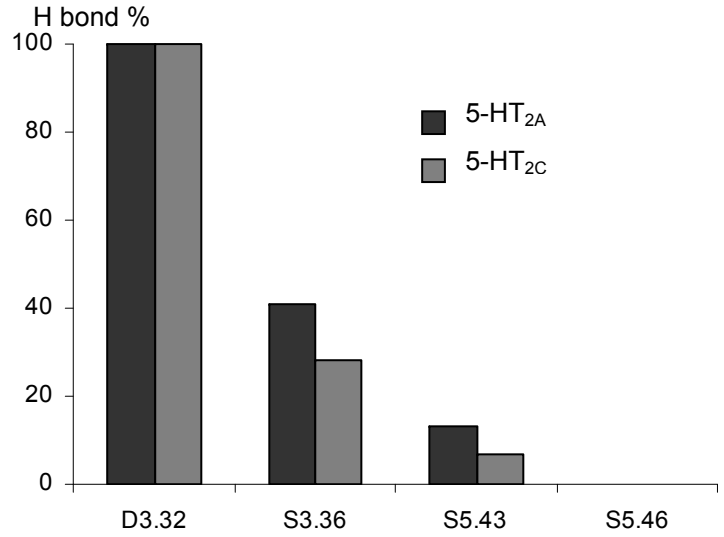


Figure 4.1.16. Serotonin in complex with 5-HT_{2A}, temperature during 1 ns molecular dynamic simulation.



This energetic stability should be related with the hydrogen bond interactions established during the MD simulation; unfortunately, polar and hydrophobic interactions were poorly represented and were unable to reproduce the mutagenesis data (Figure 4.1.17).

Figure 4.1.17. Hydrogen bond statistics for both 5-HT_{2A} and 5-HT_{2C} in complex with serotonin, for the interactions between 5-HT and relevant residues.



According to mutagenesis, we would expect Ser3.36 and Ser5.43 to interact strongly with serotonin, as the mutation of these residues had been found to cause important changes in the binding constant of 5-HT (chapter 2.1.2, Table IV).

Similar results were also found for ketanserin and QF0601b, as explained below.

Docking of ketanserin and QF0601b, energy minimization and molecular dynamic simulations

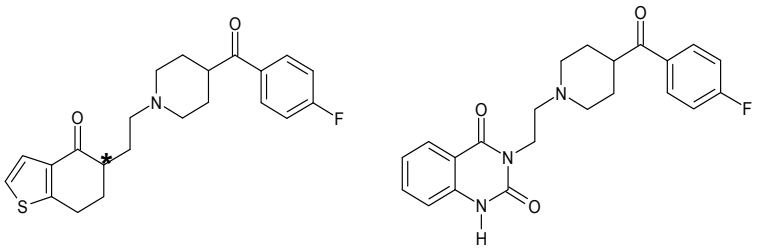
Ketanserin and QF0601b were chosen for docking into 5-HT_{2A} and 5-HT_{2C} receptors for different reasons. Ketanserin was chosen because it is one of the most studied ligand for serotonin receptors, even if the pharmacological mechanism of action is still not completely known (in fact, ketanserin is considered an antagonist or an inverse agonists), therefore a great amount of binding and mutagenesis data are available.

In addition, ketanserin is structurally related with butyrophenones, therefore it is the ideal candidate to test the active site of the models and to be used in comparison with butyrophenones.

Regarding the butyrophenones, QF0601b was selected because of its selectivity towards the 5-HT_{2A} receptor versus the 5-HT_{2C} (pKi for human 5-HT_{2A} = 8.84±0.17 and for 5-HT_{2C} = 6.48±0.70)., therefore it could help in understanding the structural requirement for selectivity for one subtype receptor with respect to others.

The 2D structures for both ligands are shown below (Figure 4.1.18). The 2D representation allows to appreciate the structural similarities in the extended conformation. QF0601b has chiral centre (represented by the symbol *), but chirality of the ligand, in this stage of the project was not taken into consideration as its pKi was measured for the racemic mixture.

Figure 4.1.18: 2D structures of QF0601b and ketanserin.



Qf0601b **ketanserin**

Both ligands were built in their extended conformations and were docked in the 5-HT_{2A} and 5-HT_{2C} receptors using the same protocol described before for the serotonin docking, using the aspartic D3.32 as the anchor point and the centre of the active site.

The best docking positions for ketanserin and QF0601b in 5-HT_{2A} and 5-HT_{2C} were submitted to energy minimization and to molecular dynamic simulation. Similarly to serotonin docking, also for ketanserin and QF0601b complexes the profiles for the total energy (figure 4.1.19 and 4.1.20) and the temperature of the systems demonstrated to be fully equilibrated in the time scale of the simulation (1 ns), even if the most important interactions were not properly represented in the hydrogen bond statistics (Figure 4.1.24).

Figure 4.1.19: QF0601b in complex with receptor 5-HT_{2A}, graph representing the total energy during 1ns simulation.

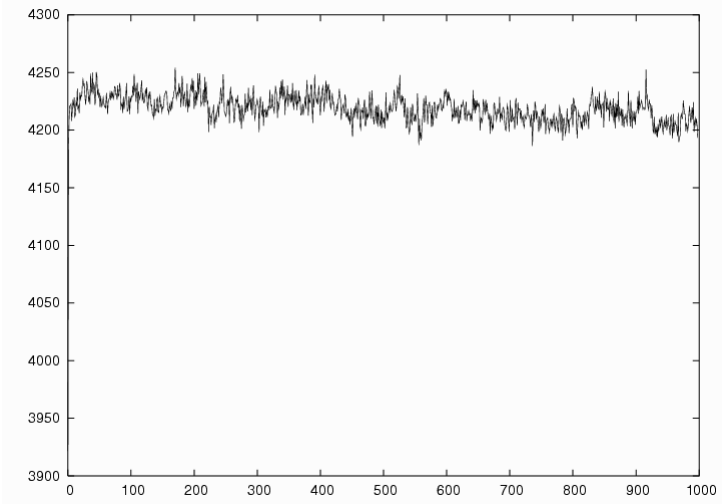
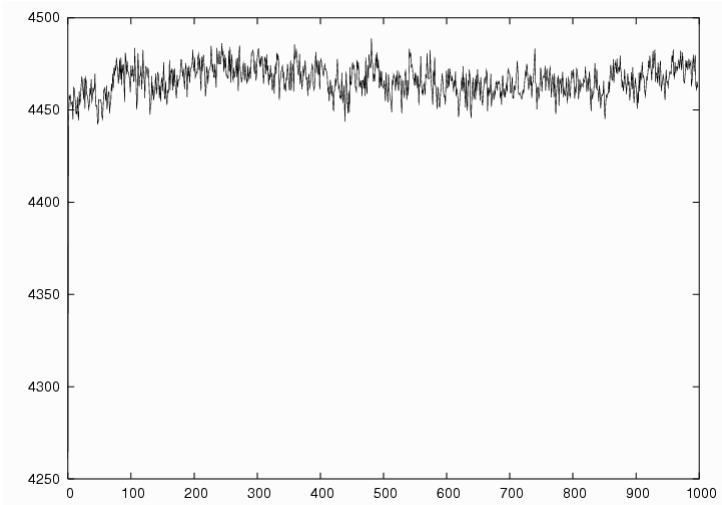


Figure 4.1.20: QF0601b in complex with receptor 5-HT_{2C}, graph representing the total energy during 1ns simulation.



Ketanserin and QF0601b binding sites were found between helixes TMHIII, IV, V and VI (figure 4.1.21, 4.1.22 and 4.1.23).

As shown in Figure 4.1.24, the ligands were kept in the binding site via the ionic interaction with D3.32, between the charged piperidinic nitrogen and the aspartic side chain, and via a hydrogen bond with S3.36. Hydrophobic interactions with aromatic residues found to play an important role in ketanserin binding by mutagenesis experiments, such as F6.52L (Choundary M.S. 1993), F6.51L (Choundary M.S. 1993), W6.48A (Roth B.L. 1997; Roth B.L. 1997), Y7.43 (Roth B.L. 1997; Roth B.L. 1997) and W7.40 (Roth B.L. 1997; Roth B.L. 1997), were not reproduced during these MD simulations.

Figure 4.1.21: the figure shows the preferred docking mode for ketanserin (green) and QF0601b (gray). ketanserin and QF0601b are located in a cavity between the helixes TMHIII, TMHV and TMHVI. The anchor point is the ionic interaction between the charged nitrogen and the D3.32. (view from the extracellular side)

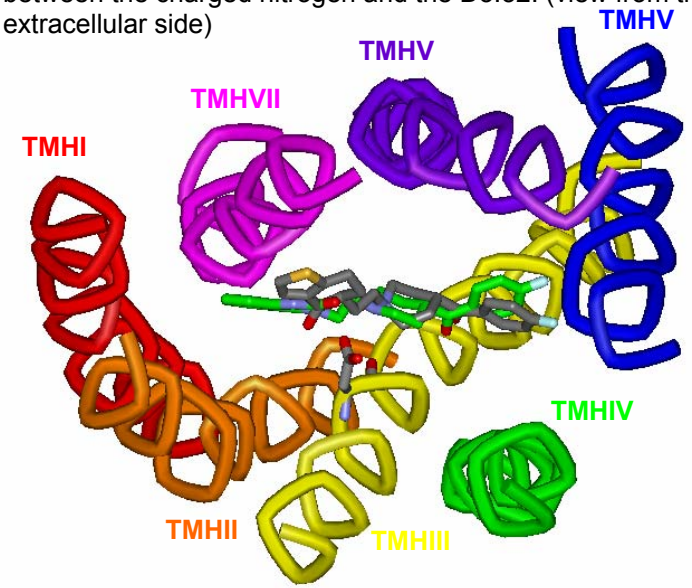


Figure 4.1.22: docking of ketanserin (gray) and QF0601b (green), seen from a different perspective than the previous figure. The ligands bind the receptor in an orientation almost parallel to the axis of the TM bundle. Some residues of TMHIII, IV and V are not shown for clarity.

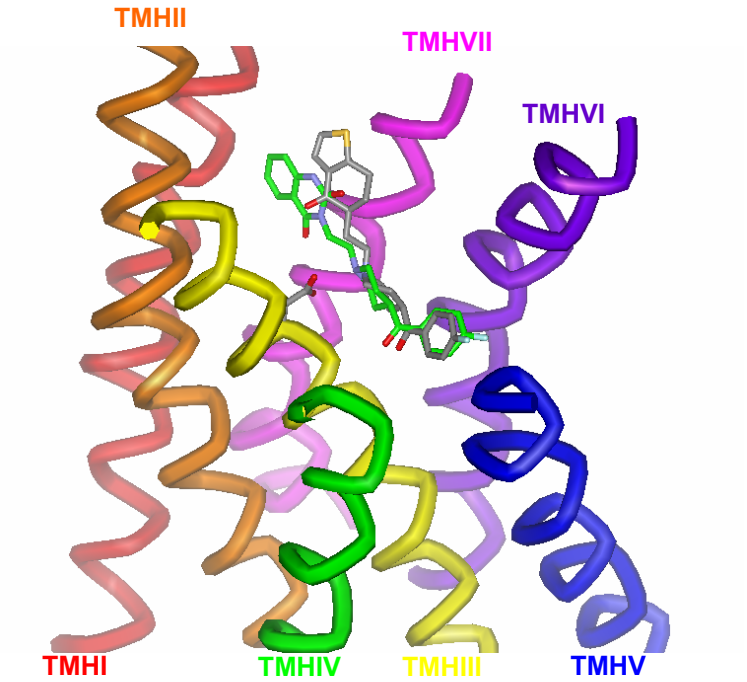


Figure 4.1.23: docking mode of ketanserin (green) and QF0601b (gray): the CPK representation of ligands shows the volume occupied in the binding site. Even if the anchor point (D3.32) is located approximately in the middle of TMHIII, the ligands protrude towards the extracellular loops (not built in these models).

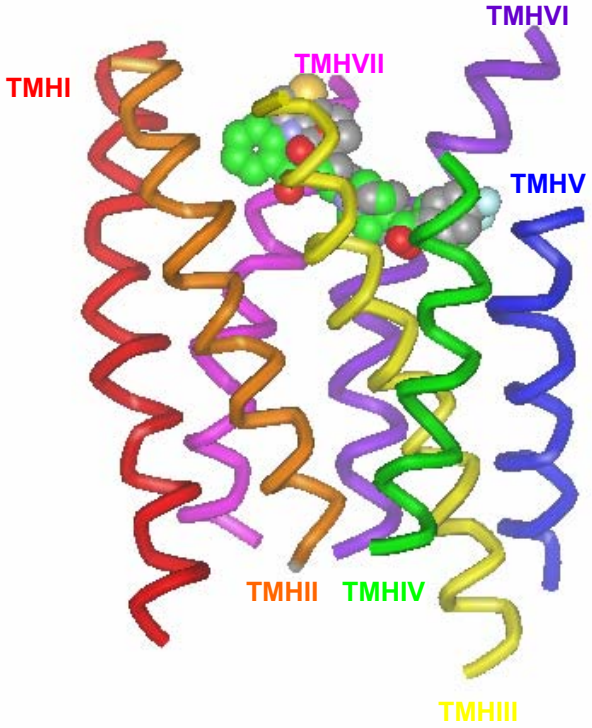
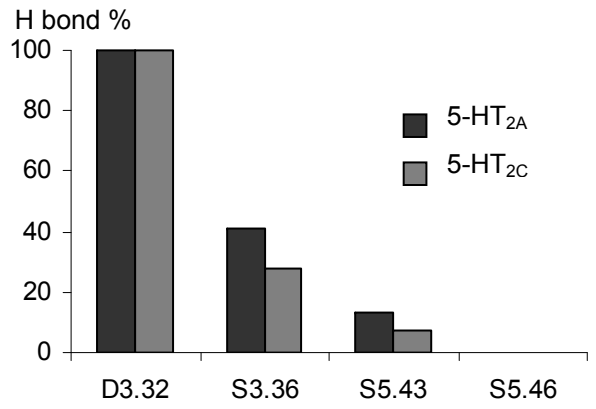


Figure 4.1.24. Hydrogen bond statistics with relevant residues during 1ns molecular dynamic simulation



Therefore, despite the stability shown during the MD simulations, the behaviour of the complexes shown in the previous graphs could not be related with mutagenesis experiments, as the interactions with the most important residues such as the serines S5.43 and S5.45 in TMHV for serotonin and the aromatic residues W6.48, F6.51, F7.40 and Y7.43 in TMHVI, were completely missed.

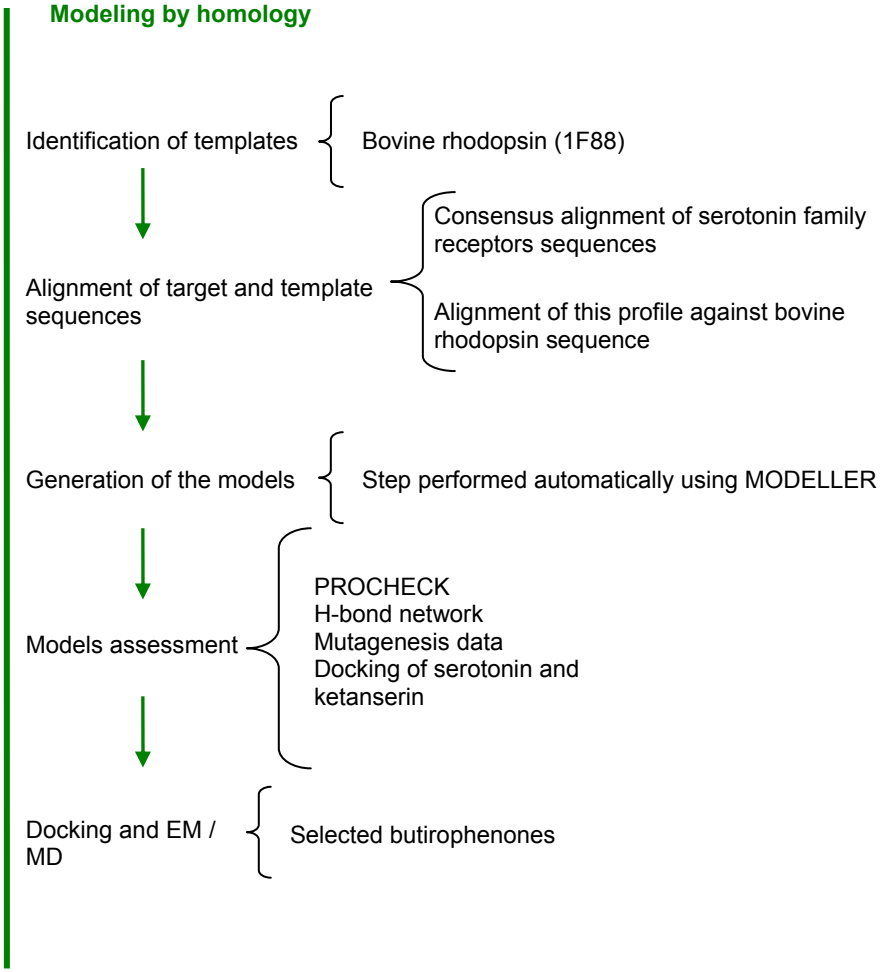
These limitations stimulated the development of improved protein models, which led to the introduction of new structural data and to more robust models.

Improving the 5-HT$_{2A}$ and 5-HT$_{2C}$ receptor models: homology modelling approach

A new homology modelling approach was developed in order to overcome the limitations of the first models. The protocol comprises different steps, some of which are in common with the topology modelling approach (see the previous chapter for details). Nevertheless, despite the similarities, some original contributions were introduced and led to new models better fitting experimental data.

The major novelties of the new homology modelling approach with respect to the previous one were: (i) the automatization of the building process by means of MODELLER, (ii) the use of alternative templates (see "Identification of structural templates" section for details) by browsing known databases of membrane proteins, and (iii) the building of loops, by means of the ArchDB loop database (Espadaler, Fernandez-Fuentes et al. 2004) and an energy minimization protocol.

An overview of the homology modelling protocol is presented below, with a mention of the key steps applied in building 5-HT$_{2A}$ and 5-HT$_{2C}$ models. All the single steps of the protocol are described in detail in the following section, in order to ease the comparison between this new protocol and the previous one based on the topology approach.



Identification of structural templates

One of the original issues of the homology modelling approach developed in our group resided in the use of other templates in addition to bovine rhodopsin crystal structure.

Since the generation of the first serotonin models, crystallization of GPCRs had progressed and at the moment of building the new models, different crystallized structures of bovine rhodopsin (Rh) were available. The analysis of these Rh structures led to the choice of 1F88, as other Rhs deposited in the PDB did not present any advantage in terms of resolution and integrity of the structure.

Considering that structures of other membrane proteins would be beneficial to the quality of 5-HT_{2A} and 5-HT_{2C} models, further 3D templates were searched. In fact, the use of only one template, with the limitations mentioned above (see the section “Building of 5-HT_{2A} and 5-HT_{2C} receptors: topological approach”, and the subsection “Identification of a template” for details) would generate models too similar to the template structure, therefore minimizing or eliminating the differences between serotonin receptors and rhodopsin. Thus, significant efforts were put in identifying suitable alternative templates among transmembrane proteins, following the idea that proteins embedded in a similar hydrophobic environment and with a known organization in seven transmembrane helixes would share some structural similarity. Searching for additional structural templates is quite common in homology modelling approach for soluble proteins, even if is not popular among GPCR modellers, probably due to the absence of crystallized GPCR proteins other than Rh.

The lack of GPCRs crystal structures poses sever limitations to the homology modelling, which could be overcome through other structural templates for the transmembrane part of the model and for the loops, retrieved from the “Membrane Proteins of known 3D-structure” database

(White S.H. 1999) and from the ArchDB loops database (Espadaler, Fernandez-Fuentes et al. 2004) .

In order to avoid as much as possible the introduction of bias due to a different folding, among all the membrane protein stored in the first database, we decided to incorporate into the models only proteins sharing the GPCR folding (seven transmembrane helixes) and displaying a minimum homology with subtype 2 serotonin receptors. According to these criteria only bacteriorhodopsin (Br) satisfied both requirements. It should be mentioned that many authors, especially after the release of the crystal structure of Rh, consider that Br is not a good template for modelling GPCR, since Br is not a GPCR itself and because there are major structural differences between GPCR and Br (Ballesteros and Palczewski 2001). Indeed, Rh and Br only share a ~20% of sequence homology and both share a ~20% of sequence homology with human 5-HT_{2A} and 5-HT_{2C}.

Considering only the homology requirement, Br has almost the same homology with the subtypes 2 of serotonin receptors than the Rh has, but when comparing their spatial organization, the superimposition of their crystal structures (1F88 and 1C3W) shows that the seven helixes bundle has a different orientation with respect to the plane of the lipid bilayer.

For this reason we decided to limit the use of the Br crystal structure to individual helixes and helixes pairs, considering that this would constitute an improvement with respect to procedures based on the use of a single template, without introducing biases due to the general folding. In order to extract this information, Br helixes were separated into pairs and these were superimposed to Rh using program STAMP (Russell and Barton 1992).

In addition and for the same reasons, ideal alpha helixes (built de novo using the sequence of human 5-HT_{2A} and human 5-HT_{2C} with ideal φ and ψ angles) were also

superimposed to the Rh structure and added to the templates library.

Regarding the modelling of the loops, suitable templates were searched in the ArchDB loops database (Espadaler, Fernandez-Fuentes et al. 2004). GPCR receptors present six loops connecting the secondary structure elements spanning the membrane: three intracellular (IL) and three extracellular (EL). They are of different length (Figure 4.1.25) and complexity. Their sequence was carefully analyzed and the loop database was browsed for templates.

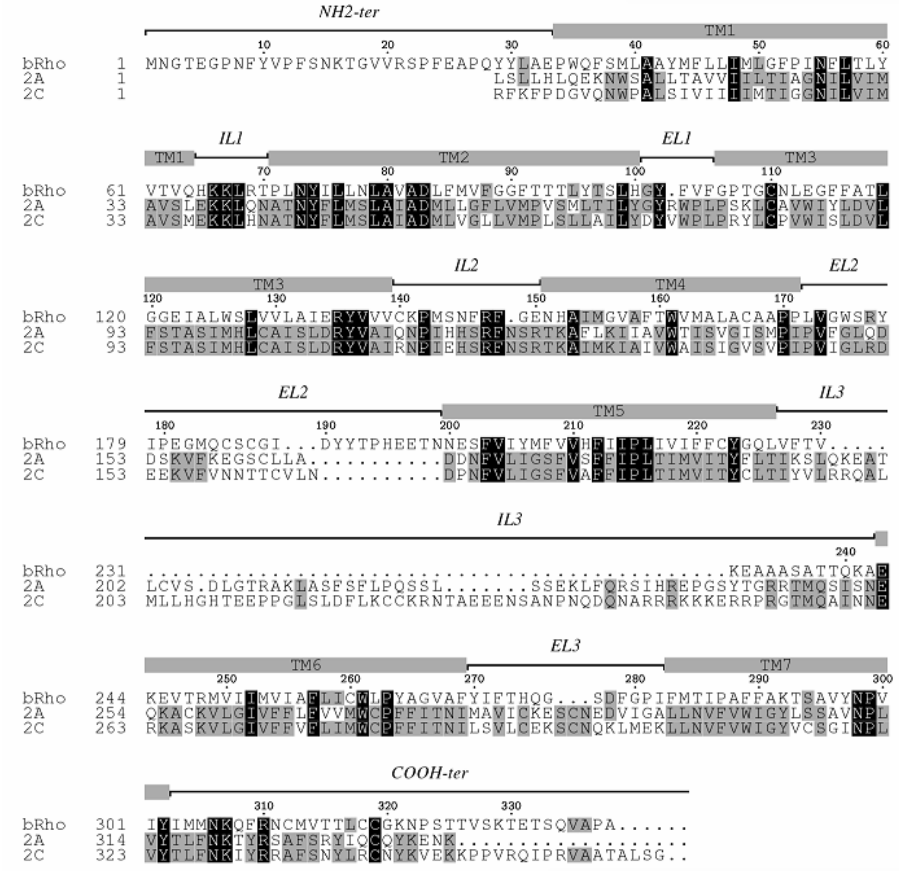
Template candidates should satisfied three characteristics:

1. be connected to two alpha helixes

2. have similar length (that is the same number of residues, with a tolerance of 2 residues) to the correspondent loop in bovine rhodopsin

3. have similar distance (between alpha helixes extremes) which braces the loop those found in rhodopsin (with a tolerance of 5 Å).

Figure 4.1.25. Figure of the alignment used as input in MODELLER. For clarity, only the bovine rhodopsin (bRho) and target sequences (2A and 2C) are shown. Conserved residues between these sequences are highlighted in black. Extensions of transmembrane regions and loops are also represented.



Suitable templates were found for almost all the loops, with the following exception: the IL3 loop (intracellular loop 3, between TMHV and TMHVI) and EL2 (extracellular loop 2, connecting the TMHIV and TMHV). IL3 interacts with the G protein and constitutes a special case, as it is the longest (almost 60 residues) and the most flexible loop. The structure of IL3 was not solved in the crystal structure of Rh deposited in the PDB and no valuable template was found. Fortunately, since IL3 is not though to play any role in the ligand binding, this difficulty was overcome by not incorporating this loop into the model.

The EL2 represents a different case; in fact its proximity to the binding site suggest that it can be important for the affinity and selectivity of some large ligands, even if biogenic amine natural ligands are known to interact with the receptor in a binding site that is located deeply in the helix bundle. As for IL3, no suitable templates were found, its structure was modelled starting from the equivalent loop present in the Rh crystal structure, despite the low sequence homology and the differences in length between the target receptors and the template (26 residues in Rh, 18 residues in 5-HT_{2A} and 19 residues in 5-HT_{2C}) (Figure 4.1.25).

However, two important structural constraints provide clues for reproducing the conformation of the template Rh, simplifying the modelling. First, the conserved disulfide bond between the EL2 and the TMHIII, formed by two cysteines highly common among GPCR receptors: the C163 and the C3.25. Second, the secondary structure prediction analysis suggested that a large part of the loop adopts a β-sheet conformation, which is the same conformation adopted by this flexible region in Rh.

Therefore, both constraints were incorporated, and the conformational space of the EL2 was consequently restricted that, apart from the solution proposed, no other conformation seemed possible. For IL1 and EL1, suitable templates were found and their structures were used in the modelling.

Alignment of target and template sequences

The template and target sequences were aligned according to the protocol explained previously (see the section: "Building of 5-HT_{2A} and 5-HT_{2C} receptors: topological approach", and the subsection: "Alignment of sequences of all serotonin receptors" for details). The final alignment used is shown in Figure 4.1.25.

Structural alignment

Structural alignment was the last step of refinement before modelling the target sequences automatically with MODELLER (Sali and Blundell 1993). Br helixes, ideal alpha helixes and loops were superimposed to bovine rhodopsin and such template alignment was introduced in the profile of serotonin receptors *versus* bovine rhodopsin sequence. This structural alignment was performed with STAMP, and the resulting alignment was considerably richer than the first alignment used in the topology modelling approach.

Generation of the models

3D models were then built automatically using MODELLER suite of program. The program is able to automatically derive restraints from the structures aligned with the target sequence when these are given as inputs, and to incorporate spatial restraints given by the user, as it was the case of the mentioned disulfide bond between C163 and C3.25. In this 3D modelling process one of the crucial step is to provide an appropriate sequence and structural alignment, since even very little differences in the gap opening positions might produce large changes in the final 3D structures.

Therefore, the 3D elements were superimposed carefully to the crystal structure, as previously explained, and the alignment was checked in order to avoid gaps, as far as

possible. The program MODELLER generated 50 candidate models: 25 for each receptor (5-HT_{2A} and 5-HT_{2C}). The main structural differences observed between these candidates were in the loops conformations (Figure 4.1.26). The disulfide bond was properly modelled in all the candidate structures. From these candidates, some structures were chosen for model assessment and for the following docking studies.

The criteria used in the choice were the MODELLER objective function, which is a measure of the robustness and reliability of the models and the visual inspection. The best structures according to both criteria were submitted to further investigation in order to assess their quality and then they were used in the docking simulations.

Figure 4.1.26. 5-HT_{2A} and 5-HT_{2C} models generated by MODELLER. The analysis of the picture shows that loops are the most variable regions.

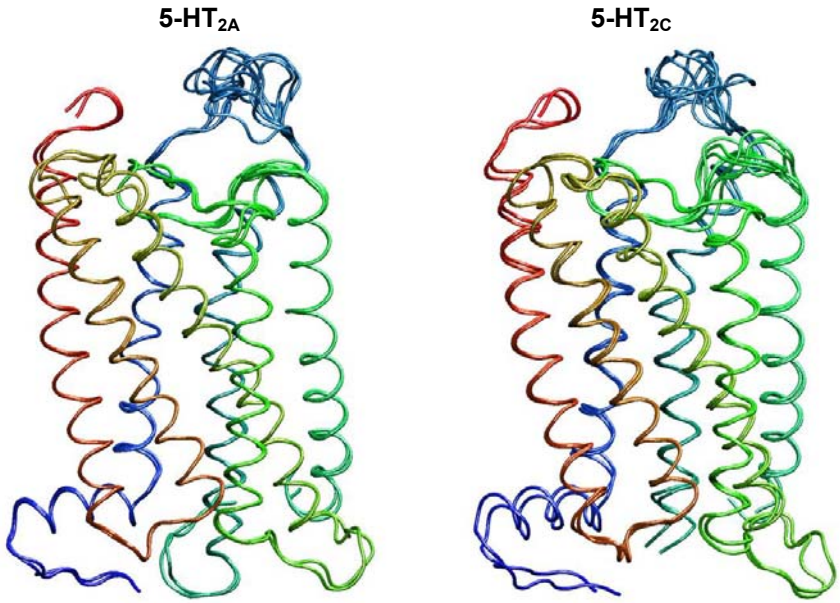
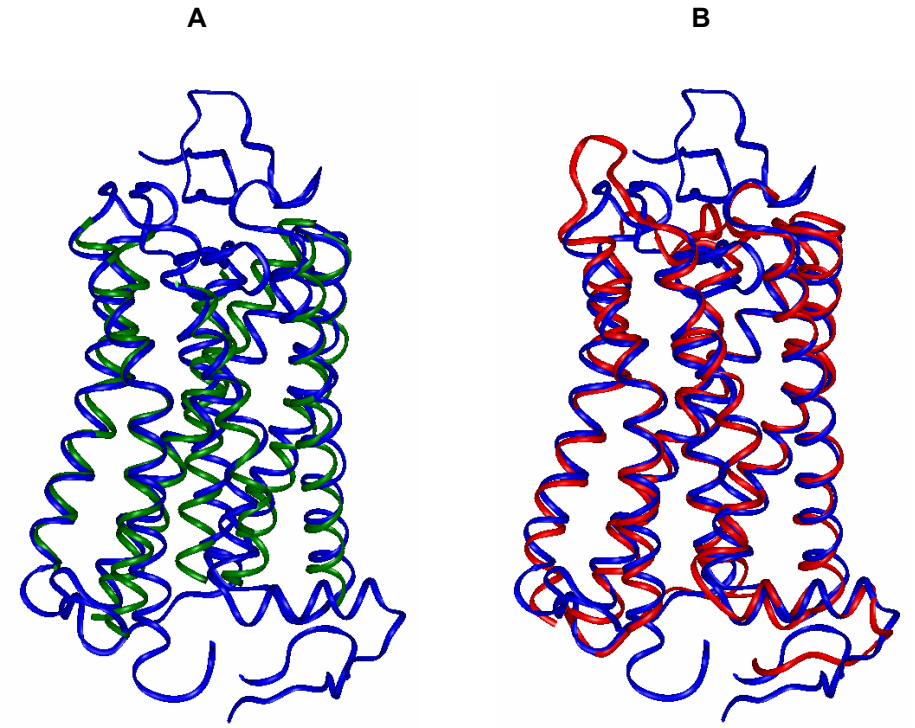


Figure 4.1.27: Comparison between the topology models and the homology models superimposed to the trace of bovine rhodopsin (blue). A) 5-HT_{2A} helixes bundle built by topology (in green) superimposed to Rh crystal structure (1F88, in blue) and B) 5-HT_{2A} model with loops built by homology (in red) superimposed to Rh (1F88, in blue). The figure shows the improvement in the quality of models.



Models assessment

Model assessment comprised different steps, most of them being the same applied in the topology modelling approach, as described in the previous section. Anyway, the assessment of the new models comprised additional steps and it is worth a detailed description of all of them. In general, we could define the model assessment as the evaluation of different parameters, mainly geometrical and physico-chemical. The resulting models were considered reliable and of acceptable quality if they satisfied all the following requirements:

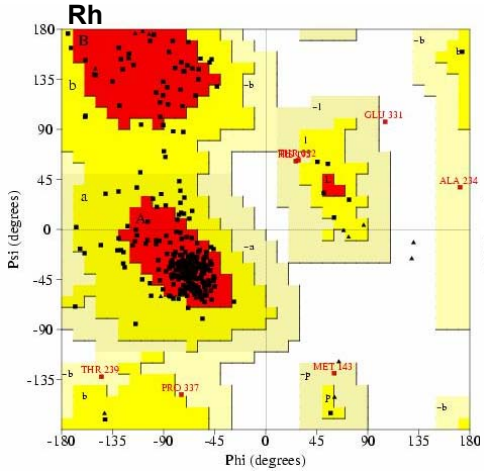
a. **Rh fitting.** The RMS between the backbone atoms of all the models and those of Rh should be lower than 1.30 Å.

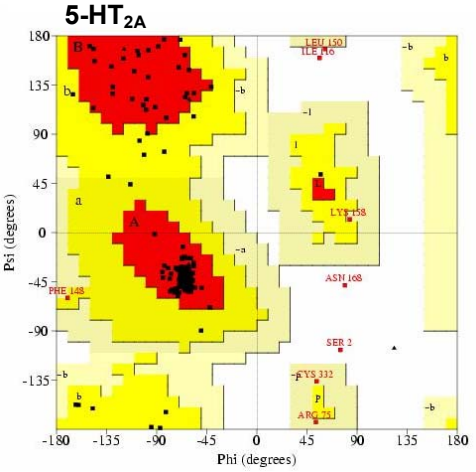
b. **Geometrical parameters.** The models should show a good dihedral angles distribution, according to the Ramachandran plot. An analysis of the backbone dihedrals present in the model structures was performed using PROCHECK software (Laskowski R.A. 1993). The analysis was performed simultaneously on the template and on the models, resulting in good quality parameters. The Ramachandran plot showed an excellent distribution of φ and ψ angles and the models have good percentages of residues in allowed regions: for 5-HT_{2A} receptor the percentage was of 98,7% and for 5-HT_{2C} receptor 98,4% (Figure 4.1.28), which are consistent and coherent with the correspondent value for Rh (99,7%)

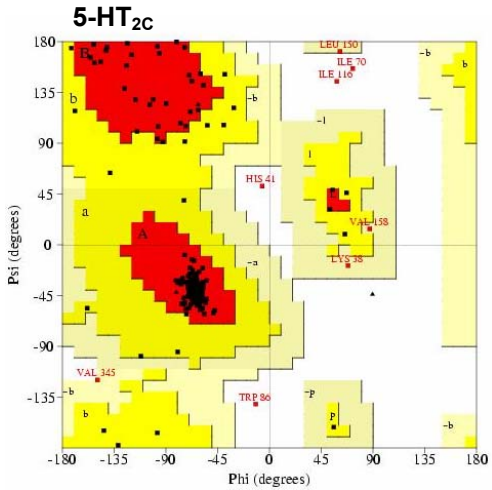
c. **Hydrophobic and hydrophilic profiles**. For the models to be considered correct, the distribution of hydrophobic and hydrophilic regions in the molecular

surface should match the lipophilic environment. This check was performed also for models obtained with the topology approach, and when compared the two analyses, the models built by homology show a better distribution of the hydrophobic / hydrophilic patterns according to visual inspection. The distribution of such regions was calculated using program GRID v.22 (Goodford PJ 1985) with different chemical probes: the hydrophobic probe (DRY) for identifying hydrophobic regions compatible with regions embedded into the lipidic membrane, the water probe (O2H) for identifying hydrophilic regions corresponding to residues located into the extracellular and intracellular spaces and the anion phosphate probe (PO4H) to check the position and orientation of positively charged residues (mainly lysines and arginines) that are expected to interact with the negatively charged phosphate groups of the membrane phospholipids. The visual inspection of these regions was found correct for all the candidate structures, as shown in the figure (Figure 4.1.29)

Figure 4.1.28. Ramachandran Plot (from right to left and from up to down): Rh crystal structure, 5-HT_{2A} and 5-HT_{2C} receptor models. The percentage of dihedrals in allowed and disallowed regions is shown in the table beside the pictures.

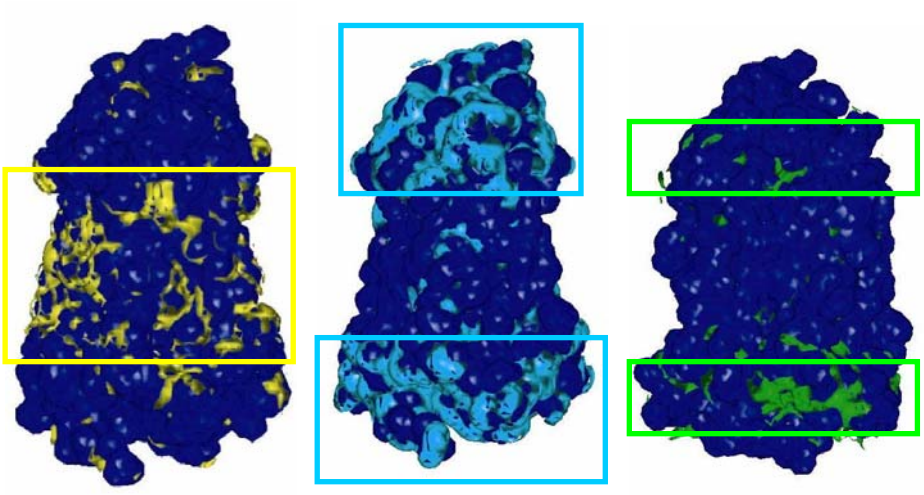






Model	Allowed regions	Disallowed regions
Rh	99,7 %	0,3 %
5-HT_{2A}	98,7 %	1,3 %
5-HT_{2C}	98,4 %	1,6 %

Figure 4.1.29: hydrophobic and hydrophilic profiles of the selected receptors (5-HT_{2A} in the image, taken as example). From left to right: DRY, O2H and PO4H. Areas with residues favourable for the interaction with the selected probe are shown with a different colour with respect to the rest of the protein. The analysis of the picture allows appreciating the concentration of hydrophobic areas in the central part of the receptor (DRY probe, in yellow) that is also a definition of the transmembrane part. Conversely, Hydrophilic areas are located at the top and bottom of the models (O2H probe, in cyan). The boundaries between the two zones are defined by the PO4H anion (green areas).



TM regions DRY probe

Extracellular & intracellular regions O2H probe

TM boundaries PO4H probe

d. **Hydrogen bond network.** The position and conformation of the most conserved residue in the GPCR family were carefully checked. Models should reproduce the correct orientation of the side chains for this set of highly conserved amino acids as these have been suggested to participate in a complex hydrogen bond network. Such network is supposed to play an important role in the conformational changes related with activation and signal transduction (Zhou W. 1994; Sealfon S.C. 1995; Scheer A. 1996; Govaerts, Lefort et al. 2001; Govaerts C. 2001) . It is located between transmembrane helixes I, II, and VII (TMHI, TMHII, and TMHVII), and involves highly conserved residues (N1.50, D2.50, P7.50). In all the models these interactions were properly reproduce, probably due also to the optimization performed automatically by MODELLER on the resulting models

e. **Consistency with mutagenesis data**: The 5-HT_{2A} and 5-HT_{2C} models were built with the purpose to help the discovery of new drugs or pharmacological tools for these receptors. Therefore the orientation of residues that play a key role in ligand biding should be in agreement with the experimental data available, especially with the mutagenesis data. Table IV lists residues which mutagenesis experiment (Kristiansen K. 1996; Edvardsen O. 1997; Beukers M.W. 1999; Beukers, Kristiansen et al. 1999) have identified as highly relevant for the binding of serotonin and other ligands. A preliminary visual inspection of the binding site proposed in the model is consistent with these experimental data.

Selection of the structure for the further studies

As cited previously, among all the models output of the program MODELLER, a single model was selected for each receptor subtype (5-HT_{2A} and 5-HT_{2C}) out of the 25 candidates per receptor generated in order to perform docking simulations with serotonin and some selected conformationally constrained butyrophenones (Brea J. 2003).

The structure of these models was then refined by reorienting the side chains of certain residues, conserved in Rh and the modelled serotonin receptors, forcing the conformation found in Rh. This operation was performed with the aim of improving the consistency of the result in the docking simulations, thus avoiding differences due to the use of slightly different starting points. Considerable efforts were put in limitating the biases due to casual conformational differences.

As the template Rh was crystallized in the inactive conformation, all the resulting models can be considered to be in the inactive conformation. In the case of GPCRs, as reminded in the "Introduction" (section: "1.1 GPCR receptors superfamily", subsection: "Activation: what is known, what is supposed, the disulfide bond between TMHIII and EL2"), several studies were performed, both experimentally and computationally, about the activation of GPCRs and some key residues were identified, even if the details of the conformational changes are still unknown.

In particular, recently published research articles proposed that the conformational changes of two aromatic residues (W6.48 and F6.52) are fundamental to determine the active or inactive state of some GPCR (Shi, Liapakis et al. 2002) and showed that rotamer changes among W6.48 and F6.52 are highly correlated, representing a rotamer toggle switch that might modulate the TMHVI Pro-kink (Shi, Liapakis et al. 2002).

A change of W6.48 from *g+* (gauche +) to *t* (trans) must be accompanied by a corresponding change of F6.52 to *t*

to avoid steric clashes. The function of the aromatic cluster in TMHVI as "rotamer switch" was extensively reviewed by Visiers (Visiers, Ballesteros et al. 2002) and Bissantz (Bissantz C. 2003) (see also the chapter "2.1.1 GPCR superfamily" for details).

Thus the position of these residues was carefully revisited. The starting positions of W6.48 and F6.52 were found in *gauche* + conformation, corresponding to the "inactive" state of the receptor, which is coherent with the conformation of the template (Rh, crystallized in the inactive state).

The starting conformation was therefore manually modified in order to generate also the "active conformation" by means of setting this "rotamer switch" in the (*t*) *trans* conformation. Eventually, 4 structures were used in the docking studies: 5-HT_{2A} in "active" and "inactive" state, and 5-HT_{2C} in "active" and "inactive" state.

3.2 Docking, energy minimization and molecular dynamic simulations

The docking studies were performed using a simpler version of the protocols used previously ("Topology modelling"), which not comprised the use of the software QXP.

Serotonin and ketanserin were docked with a two-steps protocol by means of Autodock and GOLD, while the butyrophenones were docked using GOLD with the docked position of ketanserin as reference to guide the conformational search within the active site.

In the end, serotonin, ketanserin and some butyrophenones selected for their interesting pharmacological properties were submitted to molecular dynamic simulations.

Docking of serotonin in the active and inactive conformation of 5-HT$_{2A}$ and 5-HT$_{2C}$ receptors

For each one of the receptors and for each one of the conformations ("active" and "inactive" conformations as previously described), the docking of serotonin showed good agreement with the mutagenesis data available.

This new data demonstrated the improvement of the models with respect with the first generation of models, and justified the efforts spent in re-building the models.

In general, D3.32, S3.36, Ser5.43 and S5.46 anchor the polar groups of the ligand, while Phe6.44, W6.48, F6.51 and F6.52 forms a hydrophobic environment that surrounds the indolic ring of 5-HT.

The agreement with the mutagenesis data and the strength of the interactions with the residues of the active site are expressed as percentage of hydrogen bond during 1ns MD (molecular dynamic) simulation. All the results are summarised in the graph below (Figure 4.2.30).

Figure 4.2.30. Hydrogen bond statistics for serotonin in complex with 5-HT$_{2A}$ and 5-HT$_{2C}$ (in the "active" and "inactive" conformations, 1 ns of simulation)

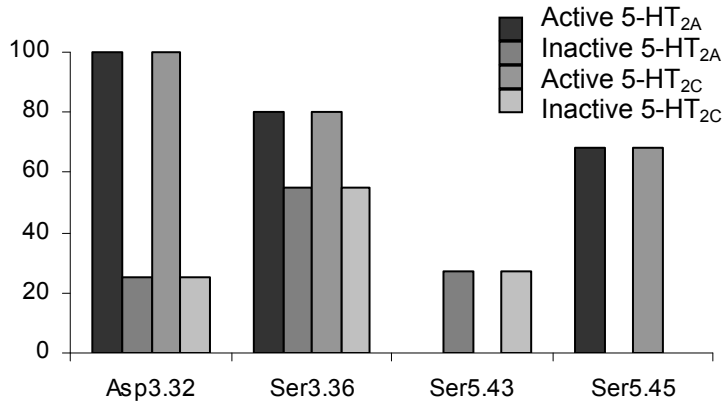
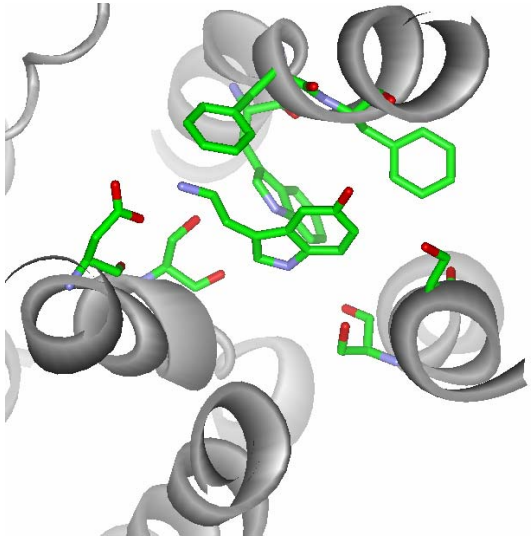


Figure 4.2.31. Serotonin docked into the active site of 5-HT$_{2A}$ model (active conformation, view from the extracellular part)



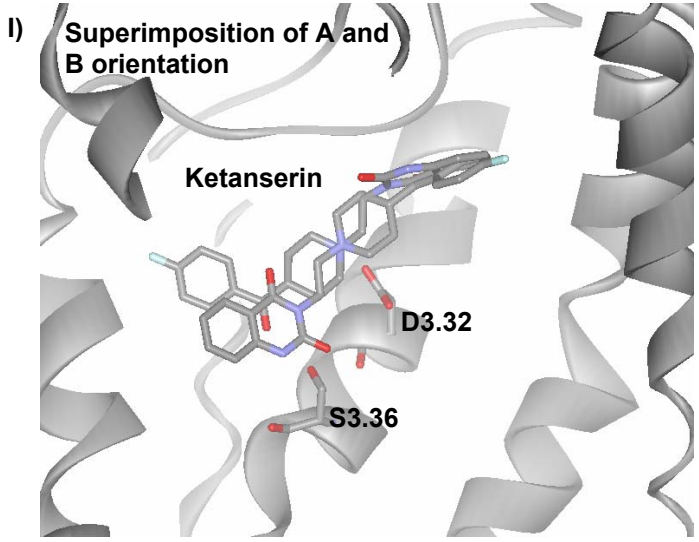
Docking of ketanserin in the active and inactive conformation of 5-HT$_{2A}$ and 5-HT$_{2C}$ receptors

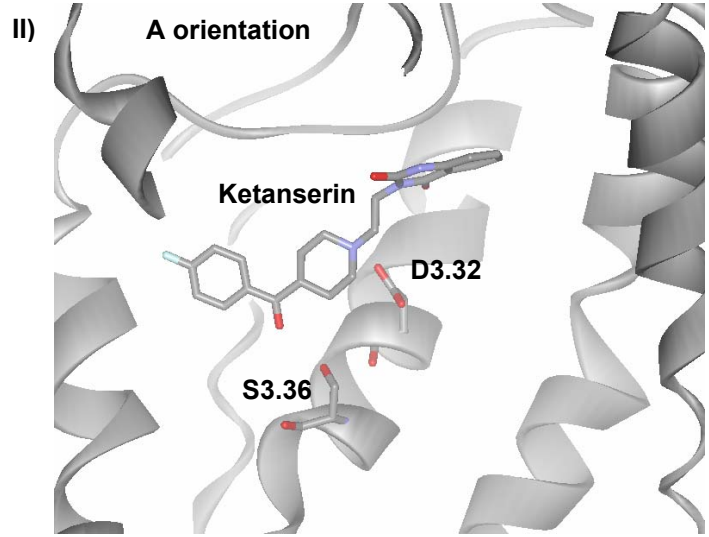
The new models overcame the limitations of the topology models described above (chapter 4, "Discussion and results", "Topology modelling of 5-HT$_{2A}$ and 5-HT$_{2C}$ receptors") since the docking studies with these new receptors showed agreement with the mutagenesis experiments (Table IV) for ketanserin.

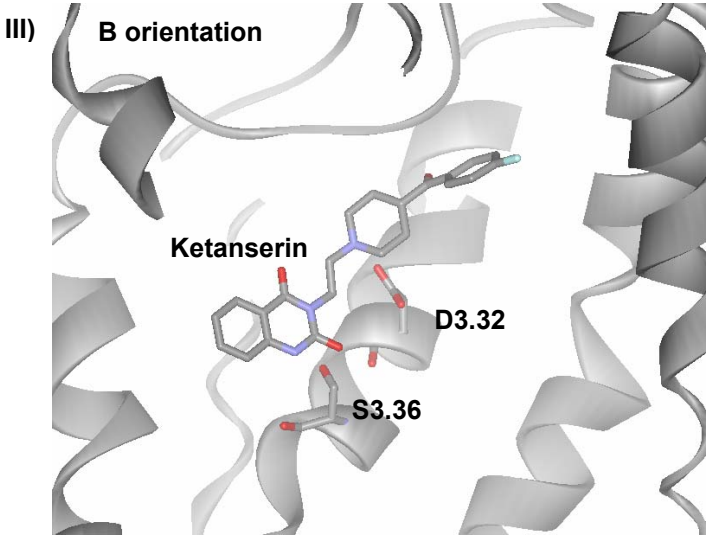
The results confirmed the hypothesis that ketanserin could bind the receptor in a cavity parallel to the 7TMH axis,

suggesting also two possible binding modes: in fact the *p*-fluorobenzoyl moiety of ketanserin could be located at the bottom of the binding site, towards the intracellular part of the receptor, or at the top, towards the extracellular part (Figure 4.2.32).

Figure 4.2.32. I) superimposition of two possible orientations (A and B) of ketanserin in the 5-HT_{2A}, II) A orientation and III) B orientation in 5-HT_{2A} active site shown separately (from top to bottom).







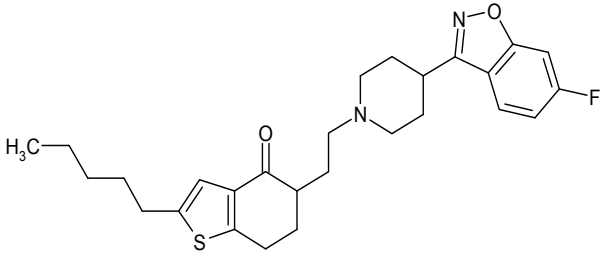
The possibility of a double binding orientation for ketanserin and ketanserin - like compounds (for example the butyrophenone series) is not completely new, as it was described also in previous studies by Raviña and co-workers (Raviña, Negreira et al. 1999). In this previous work (Raviña, Negreira et al. 1999) a much earlier version of the present series, containing only 25 compounds, was docked into the binding site of a 5-HT_{2A} receptor model, even if the receptor modelling and the ligand docking followed a completely different methodology. The authors of this work justified this duplicity due to the presence of the same pharmacophoric groups at both sides of the protonated amino group and suggested that this finding "hints the possibility of multiple binding modes".

Ketanserin and the butyrophenone series share a common scaffold of two aromatic moieties with polar groups separated by a 6 atoms ring with a charged nitrogen (mainly a piperidine or a piperazine in the butyrophenones considered). In these early studies, due to the lack of data supporting one orientation with respect to the other, the authors choose the orientation on the basis of the agreement with the generated 3D-QSAR model. On the basis of the docking results and scores carried out in our group with the new 5-HT_{2A} and 5-HT_{2C} models, both orientations (A and B) appeared reasonable and realistic for ketanserin and the butyrophenone series, and no objective reason to reject one of them was found.

In order to define which bioactive orientation the compounds preferred for binding, the new ligand QF0620b (figure 4.2.33) was synthesized as a tool to force one orientation with respect to the other. The idea behind the design of this tool was that if the ligand would have been found active in binding studies with serotonin receptors, its activity would have provided some insight about the double binding orientation. Interestingly, QF0620b not only was found active and selective towards the 5-HT_{2A} receptor, but also its pharmacological profile was similar to QF0601b (K_i for 5-HT_{2A} = 7,68; K_i for 5-HT_{2C} = 6,95), when measured in competitive binding experiments with $[H]^3$ - ketanserin.

This result supported the double binding orientation at least for compound QF0601b (figure 4.2.32): in fact the bulky aliphatic tail is supposed to reduce the number of possible binding modes by forcing QF0620b to locate the butyrophenone group in the outer part of the binding site. This result could be probably extended to other compounds of the series.

Figure 4.2.33: QF06020b, developed in our laboratory as a tool to determine the binding mode of ketanserin and ketanserin - like compounds. The ligand activity supports the hypothesis of a double binding mode parallel to the axis of transmembrane helixes.



QF0620b

Apart from the results with QF0620b, the "A" orientation for ketanserin also showed good agreement with mutagenesis data. Position A was therefore chosen for further molecular dynamics simulations (1 ns) and was used as a starting point for butyrophenones docking studies.

Molecular dynamics simulation of ketanserin - 5-HT_{2A} and ketanserin - 5-HT_{2C} complexes (both in the "active" and

"inactive" conformations) showed great stability, in terms of the equilibrium reached by the total energy, the temperature, and the RMS of the ligand (Figure 4.2.35 and Figure 4.2.36).

This equilibrium reflected the stability of the docking position, due to a rich network of polar and hydrophobic interactions between the ligand and the residues of the active site. In fact, the dynamical behaviour of the complexes, showed that the ligand is kept in the binding site through a complex hydrogen bond network (Figure 4.2.34): the anchoring interaction is the ionic bond with D3.32, which is reinforced by an hydrogen bond between the piperidine nitrogen of the ligand and the aspartic side chain; other important interactions involve most of the residues identified by mutagenesis as essential for binding (W6.48, S3.36, Y7.43) (Table IV).

S3.36 interacts with the carbonyl of the ligand through an hydrogen bond, while Y7.43 interacts with the piperidine ring and at the same time with the D3.32 side chain, keeping the D3.32 in the optimum position for the interaction with the ligand.

W6.48 interacts with the ligand especially in the "active" conformation: it competes with S3.36 for the hydrogen bond with the carbonyl group of heteroaromatic ring and could also establish aromatic interactions with the *p*-fluorobenzoyl group.

We should bear in mind that W6.48 belongs to the so called "aromatic cluster" which is highly conserved through GPCRs and is probably involved in the conformational changes during activation; therefore its involvement in the docking of ketanserin is particularly important. Moreover this data is coherent with the pharmacological profile of "inverse agonist" of ketanserin (even if other studies consider ketanserin as an antagonist).

The analysis of the hydrogen bond also showed the involvement of the second extracellular loop (EL2) in stabilizing the docking position: in fact the ligand interacts especially with

K155 and E159. The interaction with K155 is stronger in the 5-HT_{2A} complexes than in the 5-HT_{2C} complexes.

The interaction with E159 is not present in 5-HT_{2C} complexes as the residue in the correspondent position is an Arginine. The sequence 155 – 161 (Figure 4.1.25) in the 5-HT_{2A} is KVFKEGS, while in 5-HT_{2C} is KVFVNNT, and this difference is sufficient to avoid the establishment of stronger interactions with the loop EL2. In most experimental results EL2 is not thought to take part in the interaction between ligands and receptors for Class A GPCR; despite this assumption, and in the light of later results with butyrophenone ligands that will be discussed later on the text, we consider that these interactions could represent an early stage in the docking of ketanserin, maybe related with the entrance of the ligand into the active site. Anyway, this intuition should be confirmed through more sophisticated models of these complexes in explicit environment and longer molecular dynamic simulations.

With respect to the differences between the ligand – $5HT_{2A}$ and ligand – 5-HT_{2C} complexes, the analysis of the molecular dynamic simulations suggested that ketanserin prefers the interaction with the active form of the receptor 5-HT_{2A}. The data supporting this hypothesis are the lower energy during the simulation (Figure 4.2.35, part A) and the lower RMS of the ligand in the complex with the "active" conformation (Figure 4.2.35, part C).

We could conclude that the higher RMS is related with a less stable and more flexible complex, which would evolve in a more stable conformation that is not observable within the time frame embraced (1 ns). As mentioned before, these results are also in agreement with the pharmacological profile of ketanserin: even if in early studies ketanserin is described as an antagonist, in more recent studies this well known ligand is treated as an inverse agonist. Inverse agonism, as described before (Figure 2.1.2.11) decreases the basal activity of the receptors moving the equilibrium between active and

inactive receptor conformations towards the inactive conformation. Therefore the inverse agonism activity of ketanserin is coherent with a major stability of the complex with 5-HT_{2A} active form, preventing additional conformational changes that would lead to the full activation of the receptor and to the second messenger cascade. These results also reflect that ketanserin is slightly more selective for 5-HT_{2A} than 5-HT_{2C}.

Figure 4.2.34. Hydrogen bond statistics of ketanserin in complex with 5-HT_{2A} and 5-HT_{2C} receptors ("active" and "inactive" conformations (1ns of simulation)

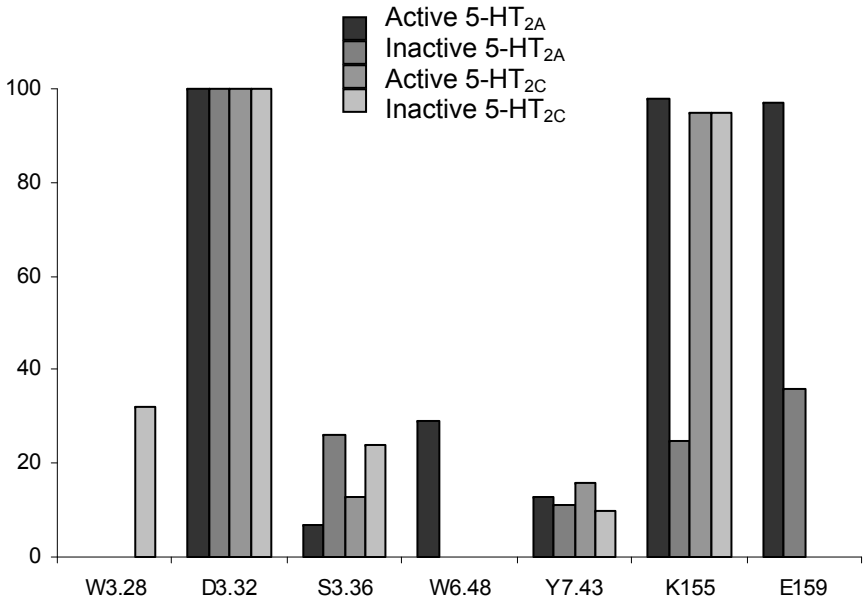
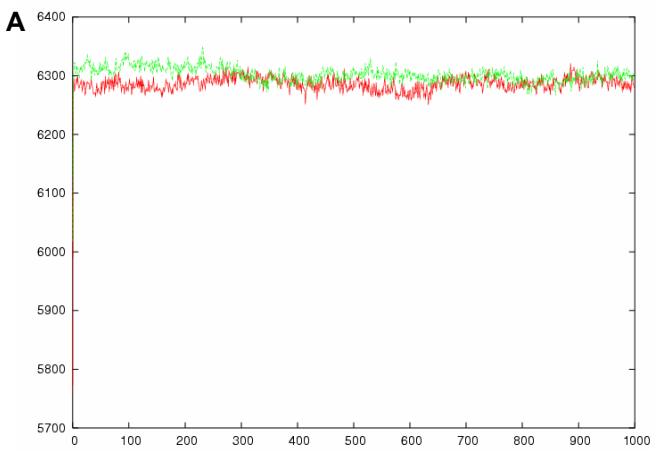


Figure 4.2.35. Ketanserin in complex with 5-HT_{2A} (active and inactive conformation). From top to bottom: graphs for total energy (A), temperature (B) and RMS of the ligand (C) are shown for the ketanserin in complex with 5-HT_{2A} (in active and inactive conformation) for 1 ns MD. The measures for the active conformation are in red, while the measures for the inactive conformation of complexes are in green.



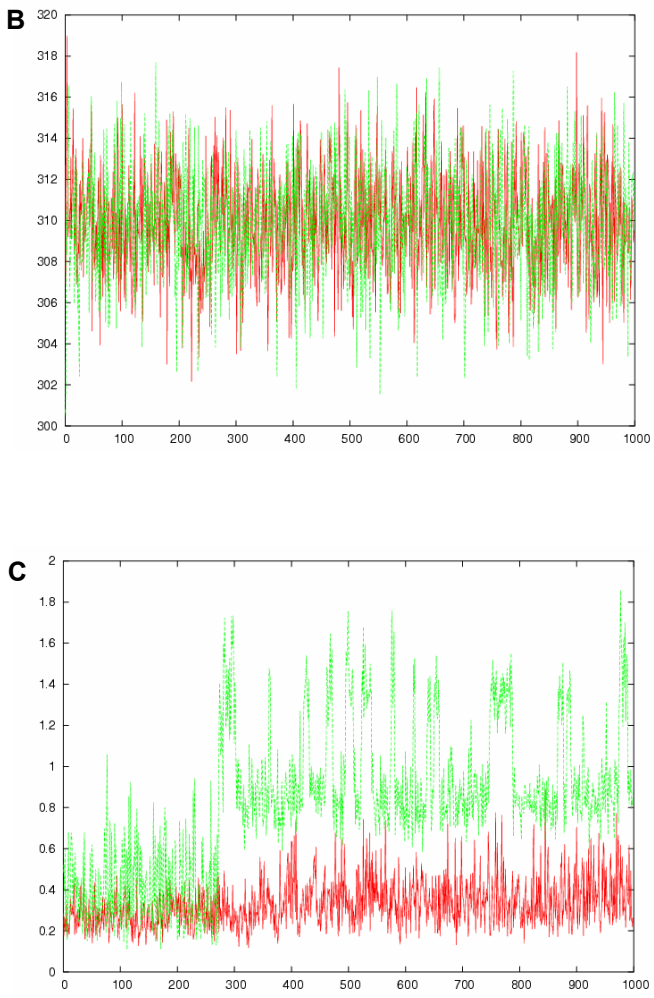
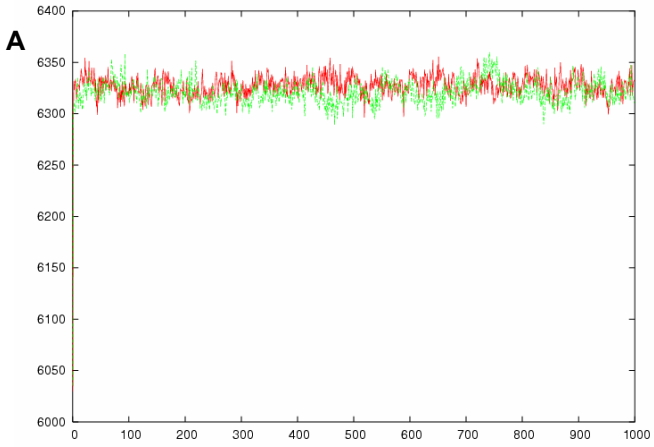
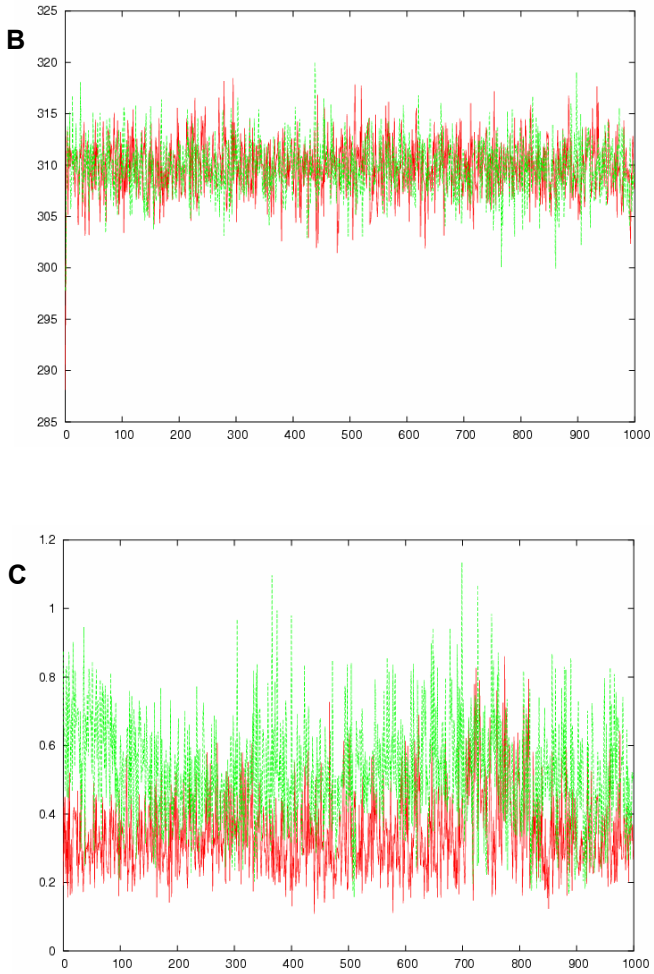


Figure 4.2.36. Ketanserin in complex with 5-HT$_{2C}$ (active and inactive conformation). As for the previous complexes, from top to bottom: graphs for total energy (A), temperature (B) and RMS of the ligand (C) are shown for the ketanserin in complex with 5-HT$_{2C}$ (in active and inactive conformation) for 1 ns MD. The measures for the active conformation are in red, while the measures for the inactive conformation of complexes are in green.





All in all, after analyzing the results from the docking simulations of ketanserin complexes, three major conclusions could be drawn:

- The hydrogen bond statistics seem to be in agreement with the major selectivity of ketanserin for the 5-HT_{2A} receptor than for the 5-HT_{2C} receptor

- The preferred docking mode seems to locate the *p*-fluorobenzoyl group at the bottom of the binding site (Figure 4.2.32 I), even if alternatives could be possible

- At the moment, current “active” and “inactive” models were not sufficient to determine the preferred protein conformation for binding; the explanation would probably require more sophisticated models. In fact other changes produced upon receptor activation are not reflected into the structure. Anyway, the complex ketanserin – 5-HT_{2A} “active form” seemed more stable than the “inactive form”, stability that is in agreement with the inverse agonism activity of the ligand.

- The EL2 loop seemed to play some role in the ligand binding: this result is surprising for class A GPCRs that bind amines, but in the light of the results that will be discussed later on, the interactions with EL2 could be related with the entrance of the ligand into the receptor and could correspond to an early stage of the docking process.

As the results with the new 5-HT_{2A} and 5-HT_{2C} models were encouraging, further docking studies with the selected butyrophenones (QF0601b, QF0610b, QF0703b and QF1004b) were carried out, using the binding position of ketanserin as a starting point.

The docking studies and molecular dynamic simulations with QF0601b, QF0610b, QF0703b and QF1004b aimed at gaining further insights and a better description of the atomic forces governing the binding process in order to:

1 – understand the role of the aromatic cluster (W6.48, F6.51, F6.52 and Y7.43) in the "active" and "inactive" form of the receptors in complex with different ligands and whether the current models could describe properly the difference between the "active" and "inactive" conformations

2 – understand if the "A" orientation chosen for the binding position of ketanserin could be generalized to all the butyrophenones. To achieve some hints about this question, two different sets of experiments were designed: the first one with A orientation and the second one with B orientation. The results are presented forward in the text.

3 – understand if the Rand S enantiomers show differences in the interaction with the receptors and if the differences are related to the pharmacological profile

The following sections will present the results obtained with the docking of QF0601b, QF0610b, QF0703 and QF1004b in the active and inactive forms of 5-HT_{2A} and 5-HT_{2C} receptors, which will answer some of the questions.

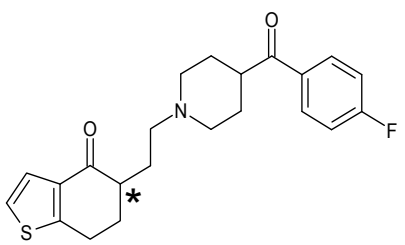
Results with butyrophenones QF0601b, QF0610b, QF0703b and QF1004b

As previously explained (section "2.2 Docking studies"), the docking of ketanserin was used as a starting point for the docking of the butyrophenones series presented in Chart I and II. The main reasons are that they are structurally related and compete for the same binding site: in

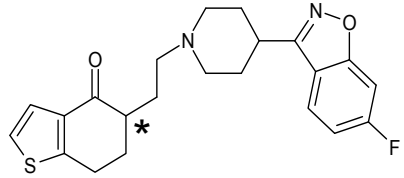
fact the binding constants of butyrophenones were measured in competitive binding experiments with [H^3]-ketanserin.

Among all the butyrophenones docked, four of them were chosen for further studies because of their interesting pharmacological profile: QF0601b, QF0610b, QF0703b and QF1004b. Their 2D structures are illustrated in figure 4.2.35, and their pKi towards human 5-HT_{2A} and 5-HT_{2C} are presented in Table VII.

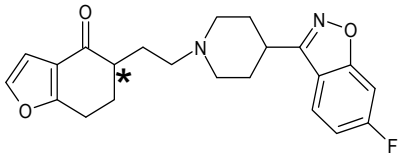
Figure 4.2.37. 2D structures of butyrophenones selected after docking studies for molecular dynamic simulations. Each one of the compounds has a chiral centre, which is indicated by the asterisk (*).



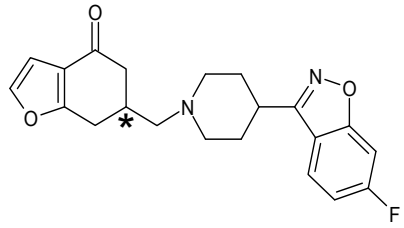
QF0601b



QF0610b



QF0703b



QF1004b

Table VII. pK_i of each one of the selected butyrophenones measured in competitive binding studies against [H^3]-ketanserin. The pKi were calculated for the racemic mixtures.

Compound	pK_i (human)	
	5-HT_{2A}	5-HT_{2C}
QF0601b	8.84±0.17	6.48±0.70
QF0610b	8.56±0.20	6.90±0.05
QF0703b	8.97±0.09	7.16±0.01
QF1004b	7.97±0.03	4.50

The interest in the pharmacology of these compounds resided in their selectivity towards 5-HT_{2A} receptor subtype than 5-HT_{2C}. As expressed in the first chapter "Introduction – Schizophrenia: ethiology, symptoms and therapy", selectivity for 5-HT_{2A} receptor subtype is considered an important characteristic of new atypical antipsychotic drugs (APDs). For this reason, 5-HT_{2A} selective ligands could be used as useful tools to understand the minimum structural requirements for the discovery of novel APDs.

Another interesting feature of these compounds was the relation between their structural similarity and the differences in the activation of secondary messenger response. Small chemical changes in their structures are related with important differences in the binding constant and in the pharmacological profile. One of the most evident examples is the case of compounds QF0703b and QF1004b (figure 4.2.37).

The only difference at the structural level between these two ligands is the length of the aliphatic chain linking the piperidine and the butyrophenone ring. This link is one carbon shorter in QF1004b than in QF0703b (figure 4.2.37) and this small structural difference is enough to change the pKi for

human 5-HT_{2C} from 7.16±0.01 to 4.50 (Table VII). Moreover, this difference causes important modifications in the curve of inhibition of the phospholipase A2 and C paths and in the typical / atypical profile. QF0703b's activity is characterized by a biphasic curve in the inhibition of the phospholipase A2 path (Table II), and a monophasic curve in the inhibition of the phospholipase C path (Table II), and its Meltzer index is 1.17 (fall within the Meltzer index for atypical antipsychotic)**.**

Conversely, QF1004b's pharmacology describes a monophasic curve in the inhibition of both PLA2 and PLC paths, and the Meltzer index for this compound is 0.99 (this value falls within the Meltzer index for typical antipsychotic). For these reasons, the models were submitted to MD (molecular dynamic) simulations in order to understand ligand – receptor interactions at the molecular level and to discover the relationship between structural changes and pharmacological differences.

QF0601b, QF0610b, QF07603b and QF1004b were docked into the active site of 5-HT_{2A} and 5-HT_{2C} receptors, both of which were modelled in the "active" and "inactive" form. The docking simulations showed that these compounds could bind the receptor in the same position as ketanserin.

The ligands, modelled in the extended conformation, are located in a cavity that lies in parallel to the axis of the transmembrane bundle. The compounds, as ketanserin, could occupy the binding site in two possible orientations: the butyrophenone moiety could be located at the top (A orientation of ketanserin, Figure 4.2.32) or at the bottom of the binding site (B orientation of ketanserin, Figure 4.2.32).

As for ketanserin, the orientation with the butyrophenone moiety towards the extracellular loops (A orientation) was chosen for the subsequent simulations. At the end, eight ligand – receptor complexes were generated, energy minimized and submitted to 1 ns MD simulations.

In this run of experiments, the chirality of the ligands was not taken into consideration. As mentioned before, all these

compounds contain a chiral centre (Figure 4.2.37) but R and S enantiomers were not modelled. At this stage of the project the experiments were not focussed on understanding the role of chirality, therefore the compounds were modelled in 2D and converted to 3D by CORINA 2.4, without providing any information to the program to generate R or S enantiomers.

In summary, the issue of chirality was postponed to other experiments, as explained forward in the text, while this run of experiment mainly focussed on:

- understanding the structural requirements for the selectivity towards 5-HT_{2A}
- understanding the role of the aromatic cluster and the conformational changes that it undergoes upon ligand binding
- studying whether the "A" orientation was the preferred binding mode also for butyrophenones.

As in the case of ketanserin, all the complexes showed high stability during the simulation, which is reflected in the equilibrium of the total energy, the temperature, the RMS of the ligands and the high percentage of hydrogen bond statistics. The analysis of these parameters is presented in the following sections.

QF0601b in complex with 5-HT_{2A} and 5-HT_{2C}

QF0601b docked into the binding site of 5-HT_{2A} and 5-HT_{2C} showed stability in all the complexes, both in the active and inactive form of the receptors.

As in the case of ketanserin, the stronger and anchoring interaction is the ionic interaction with D3.32, reinforced by an hydrogen bond. In fact this interaction is strongly present in all the complexes (Figure 4.2.38). Other polar interactions that contribute to stabilize the complexes are the hydrogen bond with S3.36 and W3.28. S3.36 is essential also for the interaction of the natural ligand 5-HT (serotonin) as demonstrated in mutagenesis experiments (Table IV). W3.28 is in the upper part of the binding site, near the second extracellular loop (EL2): during the molecular dynamic simulations interacts with the ligand via hydrogen bond or by establishing π-π interactions with the butyrophenone ring.

The complexes with active and inactive forms of the receptors show comparable stability and equilibrium of temperature, total energy and RMS of the ligand. While when comparing the 5-HT_{2A} and 5-HT_{2C} receptors, the complexes with 5-HT_{2A} appear to be more stable than the complexes with 5-HT_{2C}, result that is in agreement with the selectivity of these ligand for 5-HT_{2A} receptor. More over, as QF0601b and other butyrophenones behave as antagonist to serotonin receptors, the docking of these compounds with the inactive forms of the receptors should be preferred with respect to the active forms.

Figure 4.2.38. Hydrogen bond statistics of QF0601b in complex with 5-HT_{2A} and 5-HT_{2C} receptors ("active" and "inactive" conformations, 1ns of simulation)

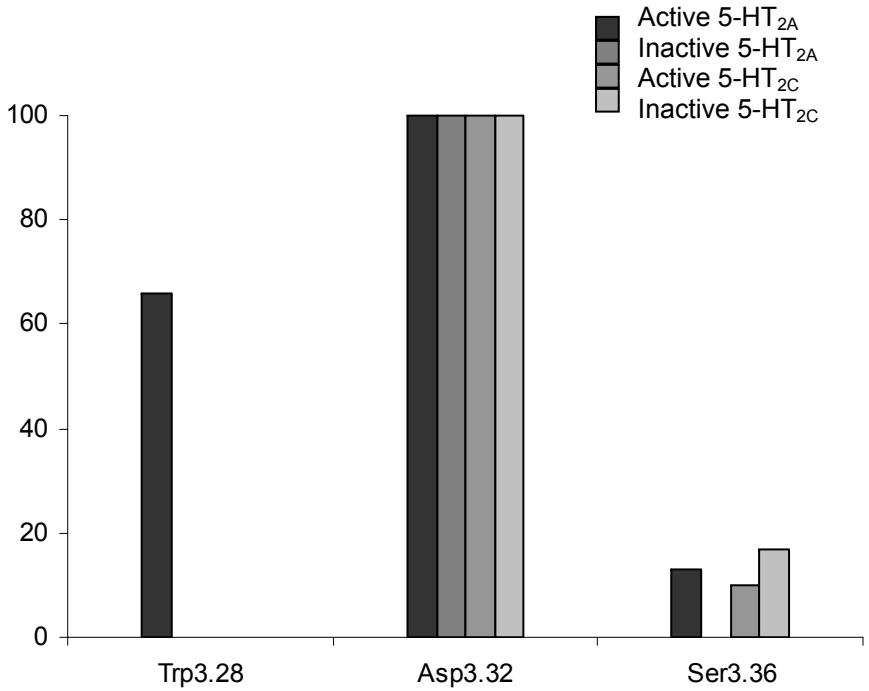
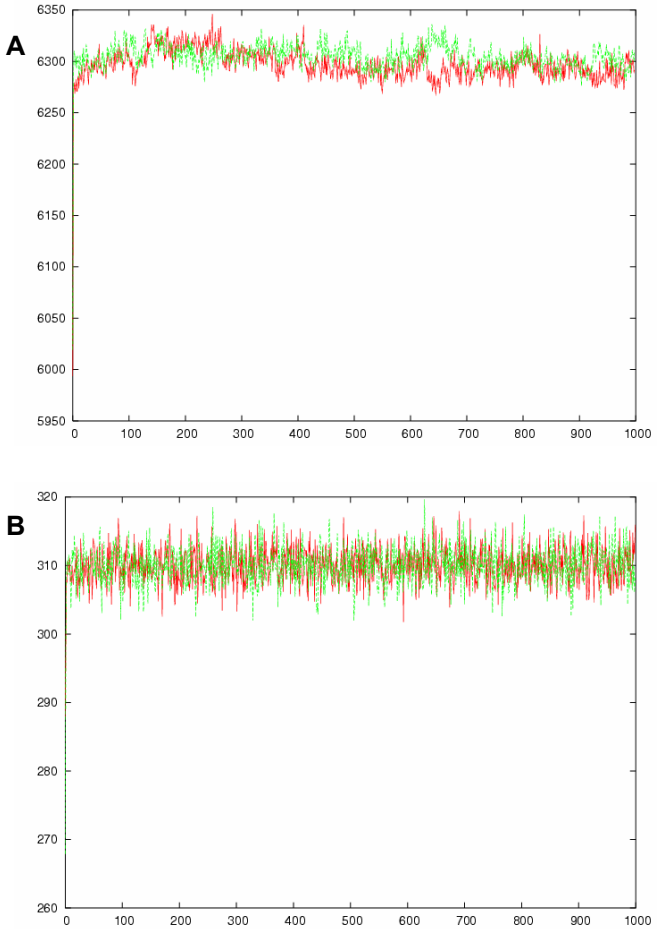


Figure 4.2.39. QF0601b in complex with 5-HT_{2A} (active and inactive conformation). As for the previous complexes, from top to bottom: graphs for total energy (A), temperature (B) and RMS of the ligand (C) are shown for QF0601b in complex with 5-HT_{2A} (in active and inactive conformation) for 1 ns MD. The measures for the active conformation are in red, while the measures for the inactive conformation of complexes are in green.



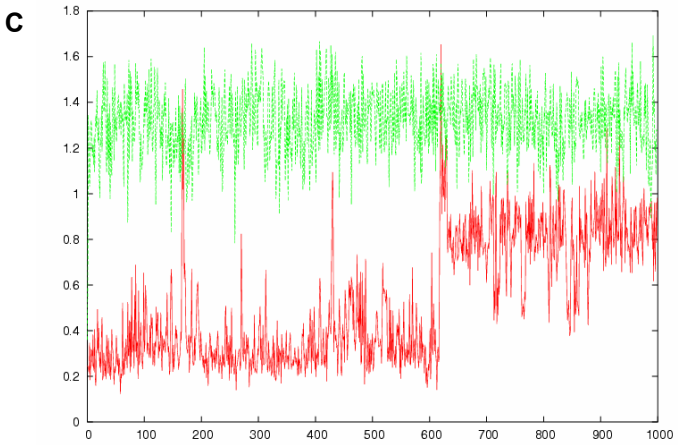
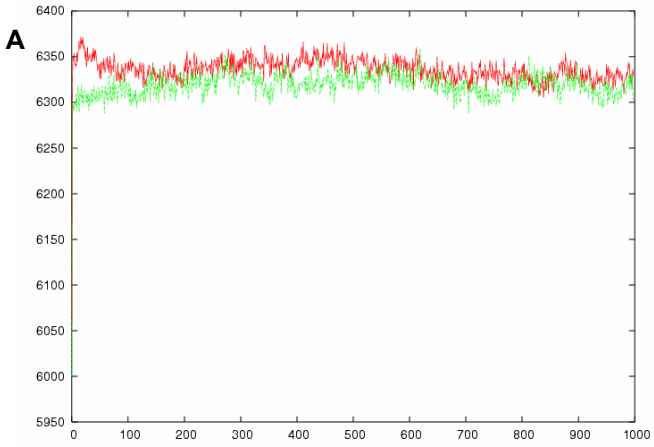
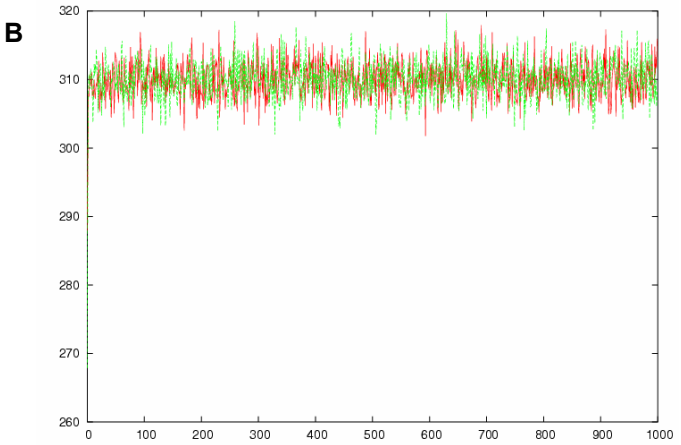
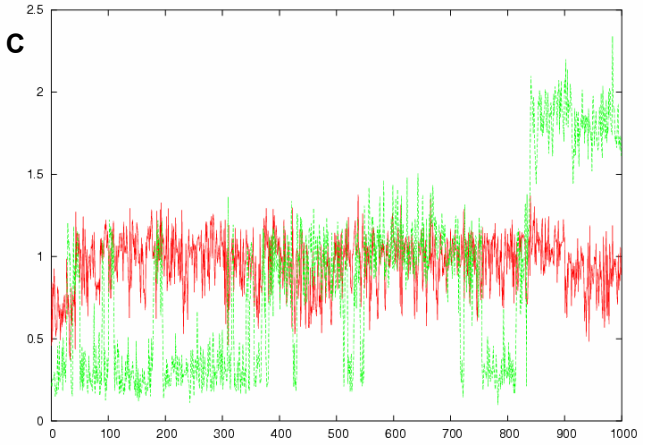


Figure 4.2.40. QF0601b in complex with 5-HT$_{2C}$ (active and inactive conformation). As for the previous complexes, from top to bottom: graphs for total energy (A), temperature (B) and RMS of the ligand (C) are shown for QF0601b in complex with 5-HT$_{2A}$ (in active and inactive conformation) for 1 ns MD. The measures for the active conformation are in red, while the measures for the inactive conformation of complexes are in green.







QF0610b in complex with 5-HT$_{2A}$ and 5-HT$_{2C}$

The ligand QF0610b is kept in the docking position through a network of hydrogen bonds that involve: D3.32, S3.36, T3.37, W3.28, S5.46, W6.48, and Y7.43. D3.32 is the strongest interaction, followed by W6.48 (Figure 4.2.41).

W6.48 belongs to the aromatic cluster of residues probably involved in the activation conformational changes of these GPCRs, and interacts with the ligand especially in the "active" conformation. In fact, the dihedral angle of W6.48 side chain governs the position of the indolic nitrogen, and therefore its polar interactions with the carbonylic group of the ligand. During the molecular dynamic simulation, the carbonylic oxygen of QF0610b could interact with S3.36 or with W6.48: both hydrogen bonds do not occur in the same frames, and the carbonyl swings between the S3.36 and W6.48 during the entire simulation.

T3.37 could interact either with the ligand or with S3.36 side chain, helping S3.36 side chain to maintain the optimum conformation for the interaction with QF0610b. A similar role is played by Y7.43 that establishes hydrogen bonds with the ligand and with D3.32, maintaining the aspartic side chain in the optimum position for the interaction with the positively charged piperidinic nitrogen.

S5.46 is one of the residues important for the interaction with the natural ligand serotonin (5-HT) and other ligands (as show in Table IV).

All in all, these complexes show approximately the same stability as the complexes with QF0601b: this result is not surprising as QF0601b and QF0610b pKi are very similar (and both ligands are selective for the 5-HT$_{2A}$ receptor). QF0610b, due to the presence of the isoxazolic group can explore different zones of the bottom of the binding site by establishing interactions with residues S5.46 and T3.37, while the butyrophenone ring establish the same interactions as QF0601b.

Figure 4.2.41. Hydrogen bond statistics of QF0610b in complex with 5-HT$_{2A}$ and 5-HT$_{2C}$ receptors ("active" and "inactive" conformations, 1ns of simulation)

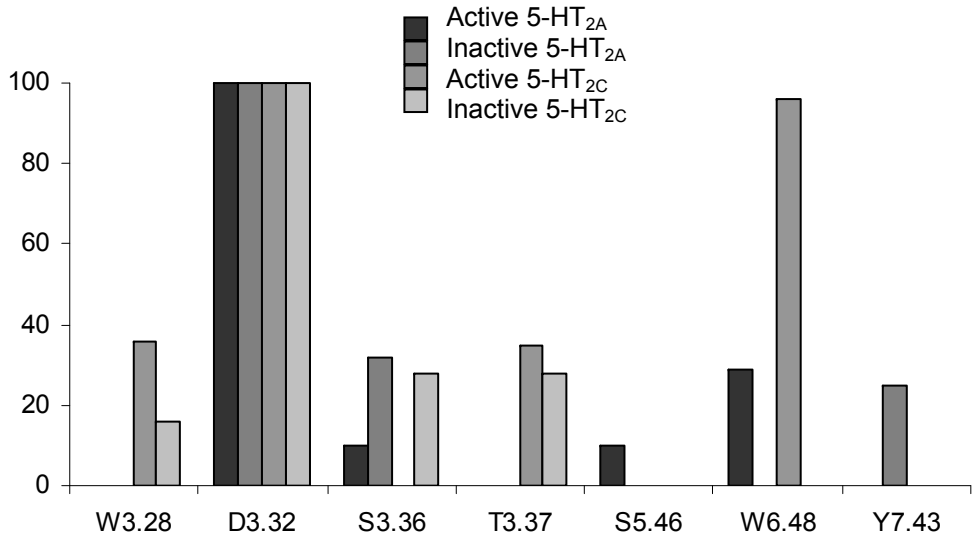
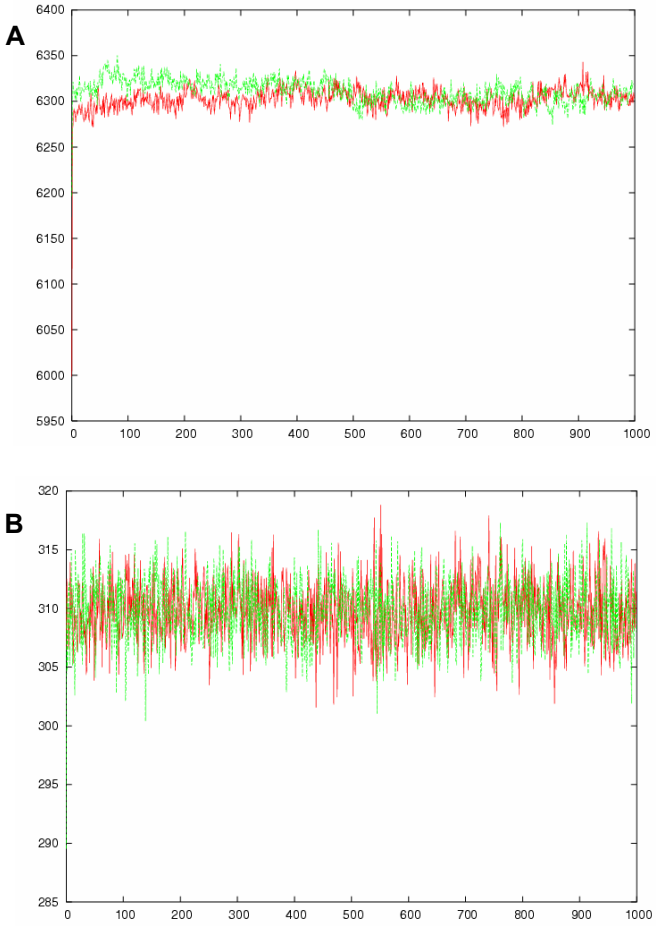


Figure 4.2.42. QF0610b in complex with 5-HT_{2A} (active and inactive conformation). As for the previous complexes, from top to bottom: graphs for total energy (A), temperature (B) and RMS of the ligand (C) are shown for QF0610b in complex with 5-HT_{2A} (in active and inactive conformation) for 1 ns MD. The measures for the active conformation are in red, while the measures for the inactive conformation of complexes are in green.



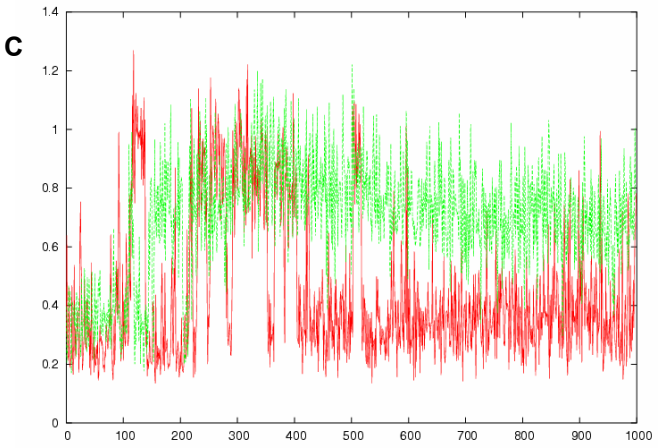
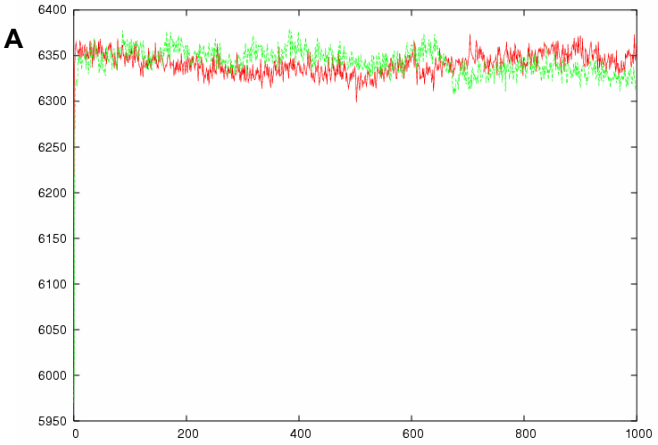
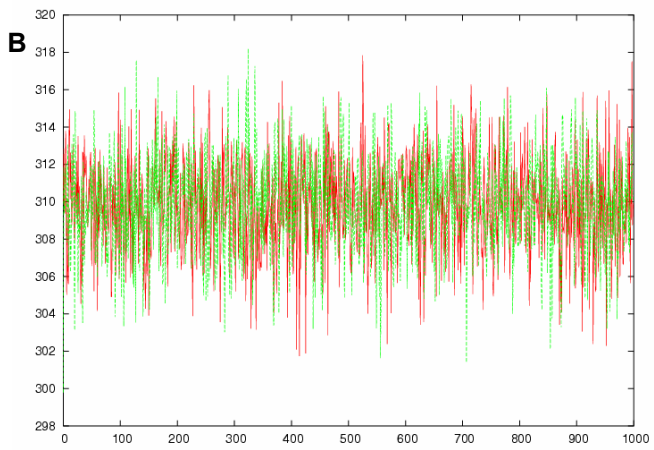
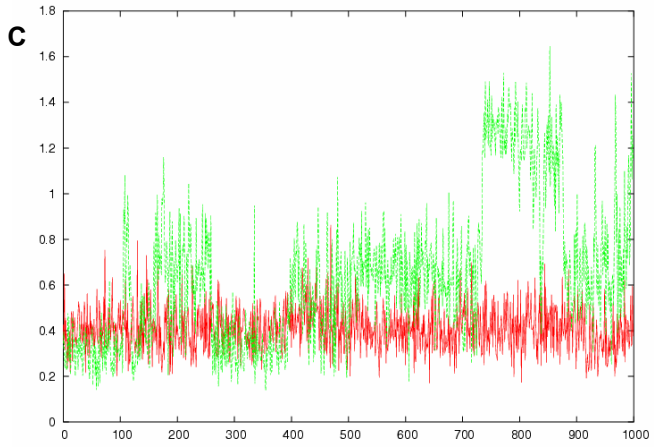


Figure 4.2.43. QF0610b in complex with 5-HT_{2C} (active and inactive conformation). As for the previous complexes, from top to bottom: graphs for total energy (A), temperature (B) and RMS of the ligand (C) are shown for QF0610b in complex with 5-HT_{2C} (in active and inactive conformation) for 1 ns MD. The measures for the active conformation are in red, while the measures for the inactive conformation of complexes are in green.







QF0703b in complex with 5-HT$_{2A}$ and 5-HT$_{2C}$: hydrogen bond statistics

As for the other compounds studied, QF0703b interacts with most of the residues already identified as important in previous modelling simulations or mutagenesis experiments: W3.28, D3.32, S3.36, T3.37, W6.48, N7.36, Y7.43, and E159. The interactions not identified in previous docking experiments with ketanserin, QF0601b and QF0610b were the hydrogen bonds of QF0703b with the side chains of N7.36 and E159. E159 belongs to the EL2 in 5-HT$_{2C}$. The EL2 has different lengths (figure 4.1.25) correspondent residue is Q159 in 5-HT$_{2A}$. The role of EL2 residues is controversial, and could be related to an initial step in the docking process. Moreover the butyrophenone ring of QF0703b is more similar to the quinazoline-2,4-dione ring of ketanserin, and probably this is one of the reasons that allow this ligand to interact with EL2 loops as ketanserin.

Figure 4.2.44. Hydrogen bond statistics of QF0703b in complex with 5-HT$_{2A}$ and 5-HT$_{2C}$ receptors (“active” and “inactive” conformations, 1ns of simulation)

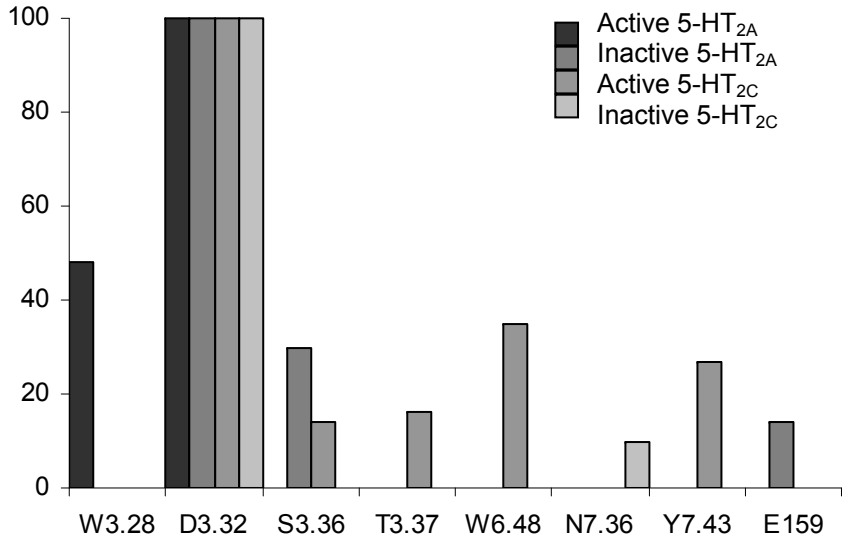
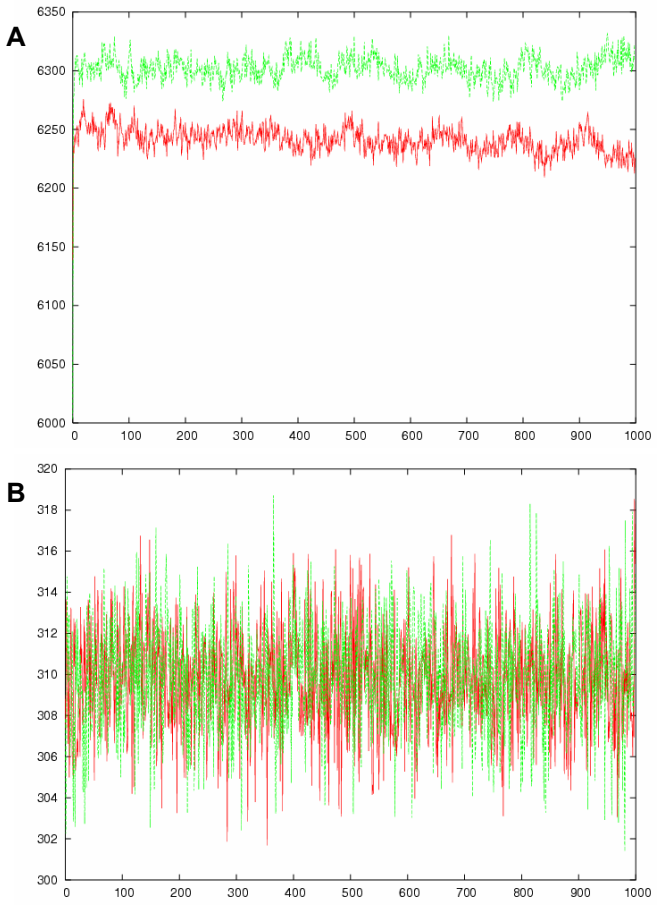


Figure 4.2.45. QF0703b in complex with 5-HT$_{2A}$ (active and inactive conformation). As for the previous complexes, from top to bottom: graphs for total energy (A), temperature (B) and RMS of the ligand (C) are shown for QF0703b in complex with 5-HT$_{2A}$ (in active and inactive conformation) for 1 ns MD. The measures for the active conformation are in red, while the measures for the inactive conformation of complexes are in green.



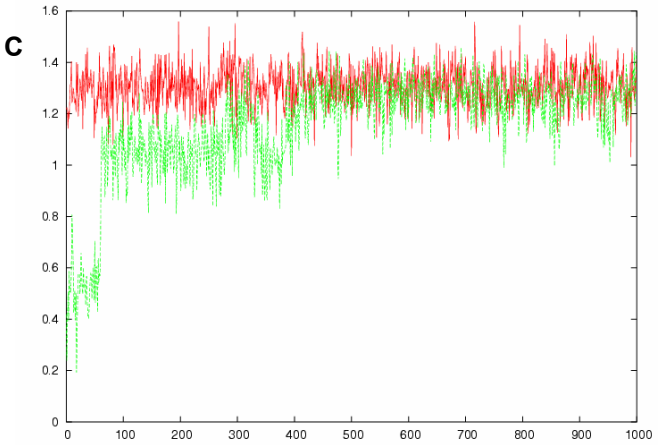
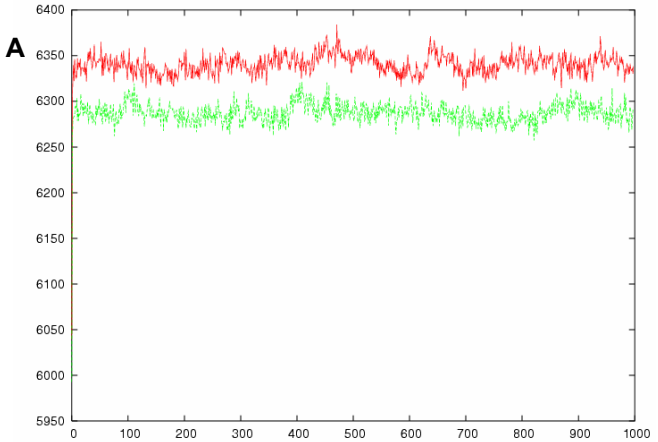
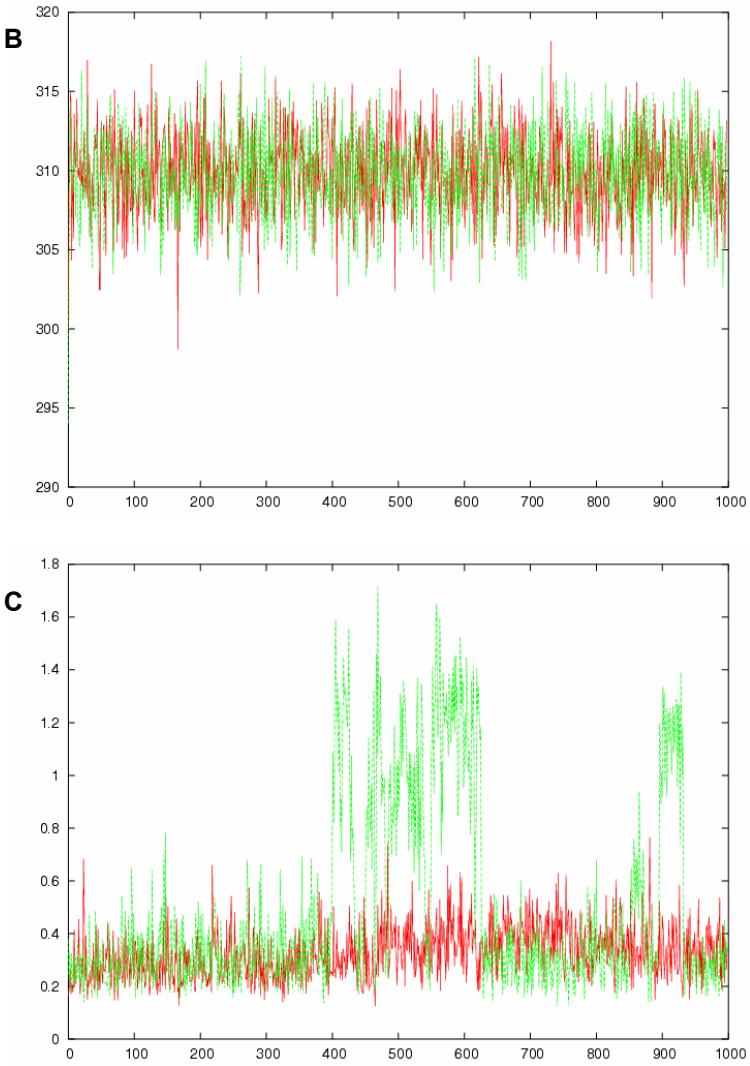


Figure 4.2.46. QF0703b in complex with 5-HT$_{2C}$ (active and inactive conformation). As for the previous complexes, from top to bottom: graphs for total energy (A), temperature (B) and RMS of the ligand (C) are shown for QF0703b in complex with 5-HT$_{2C}$ (in active and inactive conformation) for 1 ns MD. The measures for the active conformation are in red, while the measures for the inactive conformation of complexes are in green.





QF1004b in complex with 5-HT$_{2A}$ and 5-HT$_{2C}$

QF1004b interacts via hydrogen bond with: S2.45, D3.32, S3.36, T3.37, S5.46, W6.48, N7.36, and Y7.43. This ligand is structurally related with the ligand QF0703b: in fact QF1004b and QF0703b share the same butyrophenone and heteroaromatic ring, and the only structural difference is located in the linker between the butyrophenone ring and the piperidine ring. The shorter linker could be responsible of the lack of interaction with residues of the EL2, and of stronger interactions with the bottom of the binding side (as shown when comparing Figure 4.2.44 and Figure 4.2.47: the hydrogen bond statistics for residues Y7.43, N7.36 and W6.48 are higher for QF1004b than for QF0703b).

Figure 4.2.47. Hydrogen bond statistics of QF1004b in complex with 5-HT$_{2A}$ and 5-HT$_{2C}$ receptors ("active" and "inactive" conformations, 1ns of simulation)

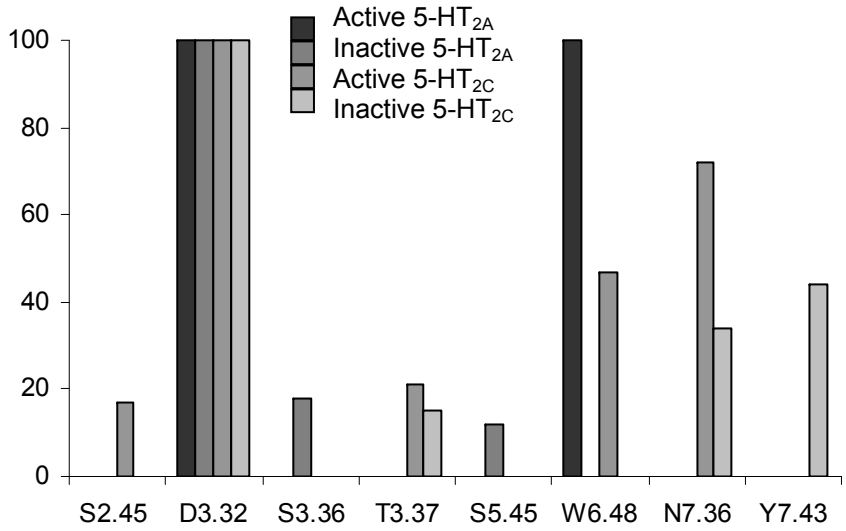
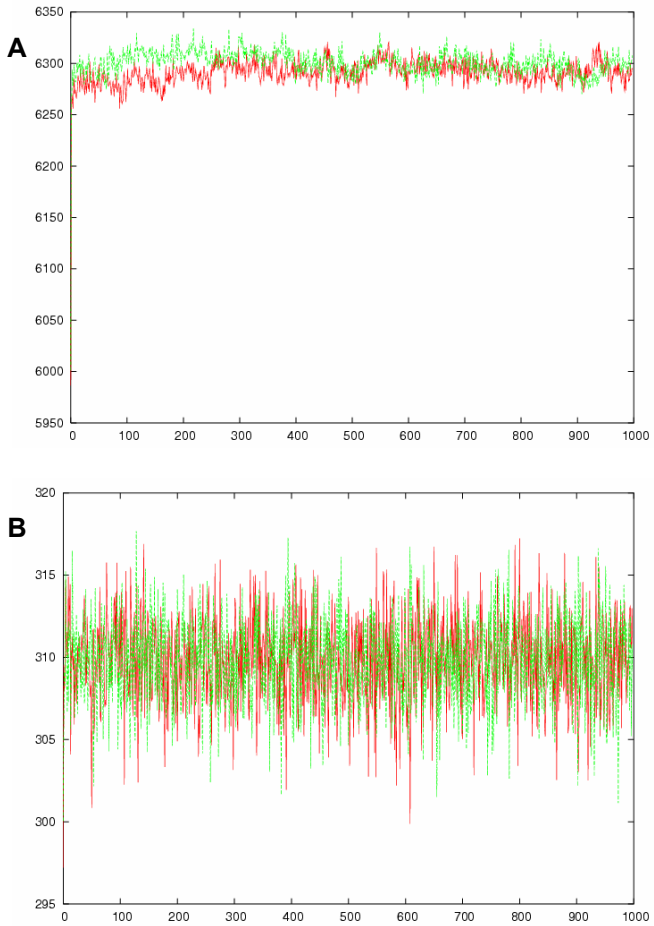


Figure 4.2.48. QF1004b in complex with 5-HT$_{2A}$ (active and inactive conformation). As for the previous complexes, from top to bottom: graphs for total energy (A), temperature (B) and RMS of the ligand (C) are shown for QF1004b in complex with 5-HT$_{2A}$ (in active and inactive conformation) for 1 ns MD. The measures for the active conformation are in red, while the measures for the inactive conformation of complexes are in green.



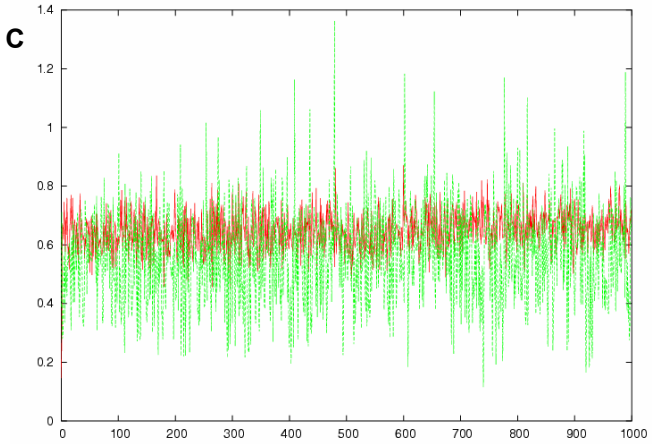
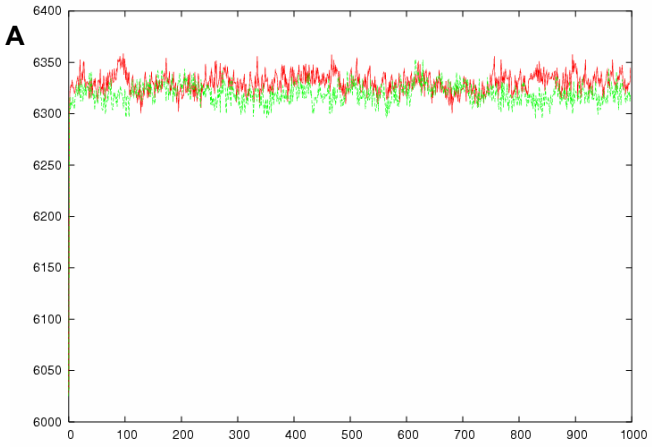
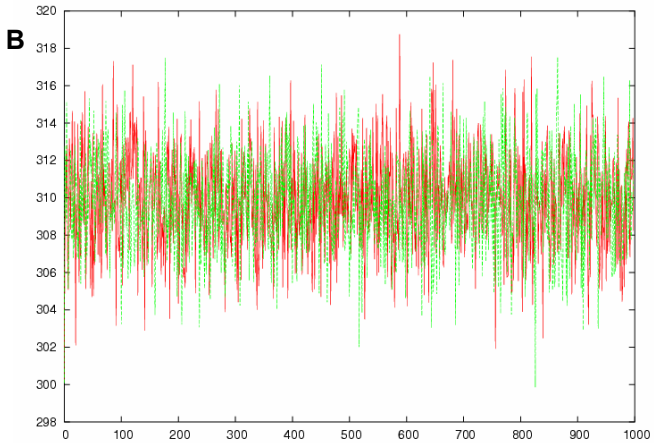
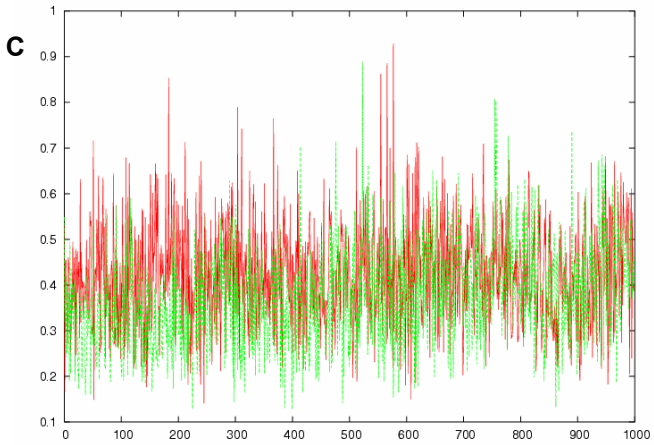


Figure 4.2.49. QF1004b in complex with 5-HT$_{2C}$ (active and inactive conformation). As for the previous complexes, from top to bottom: graphs for total energy (A), temperature (B) and RMS of the ligand (C) are shown for QF1004b in complex with 5-HT$_{2C}$ (in active and inactive conformation) for 1 ns MD. The measures for the active conformation are in red, while the measures for the inactive conformation of complexes are in green.







Figures of all the complexes (ketanserin and butyrophenones in complex with 5-HT_{2A} and 5-HT_{2C} active and inactive conformations) follow. In the first figure of the series (Figure 4.2.50), the residues involved in the most important interactions are shown and labelled. In the rest of pictures, the complexes show the same orientation, in order to identify easily the residues responsible of key polar and hydrophobic interactions.

Figure 4.2.50. Ketanserin in complex with the 5-HT_{2A} and 5-HT_{2C}. The active conformation of the receptors is shown in green, while the inactive is depicted in gray. Only the residues involved in the most relevant interactions are shown.

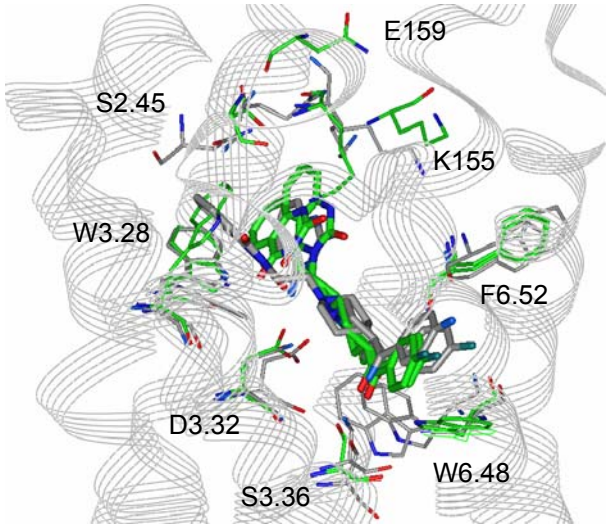


Figure 4.2.51. Ketanserin in complex with the 5-HT_{2A} active conformation: the ligand interacts preferably with D3.32, S3.36, W3.28 and W6.48. EL2 also seems to play some role in stabilizing the complex.

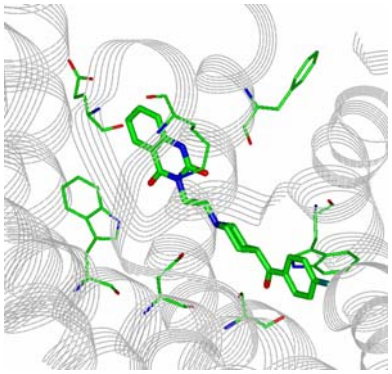


Figure 4.2.52. Ketanserin in complex with the 5-HT_{2A} inactive conformation: the different conformation of W6.48 prevents the polar interaction between this residue and the ligand.

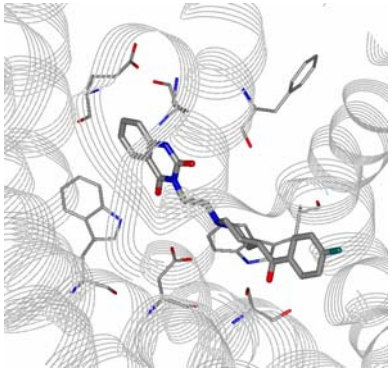


Figure 4.2.53. Ketanserin in complex with the 5-HT_{2C} active conformation: the ligand establishes less polar interactions than with 5-HT_{2A} receptor subtype, in agreement with the selectivity of ketanserin for 5-HT_{2A}.

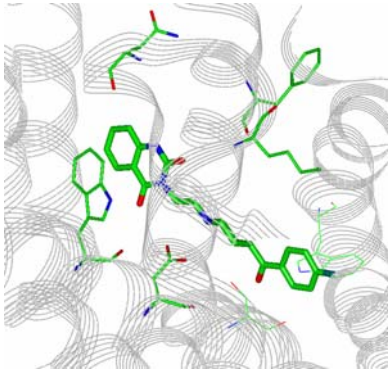


Figure 4.2.54. Ketanserin in complex with the 5-HT_{2C} inactive conformation. The main interactions between the ligand and the complex involve the residues D3.32, S3.36, Y7.43, K155, W3.28

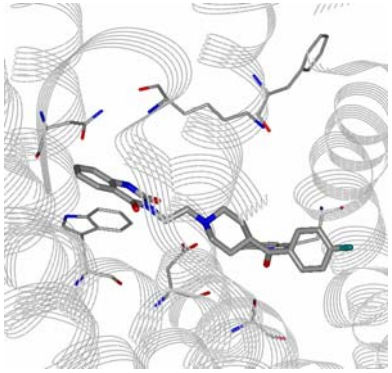


Figure 4.2.55. QF0601b in complex with 5-HT_{2A} and 5-HT_{2C}. As for the figure 44, the active conformation of the receptors is shown in green, while the inactive is in gray. Only residues involved in relevant interactions are shown.

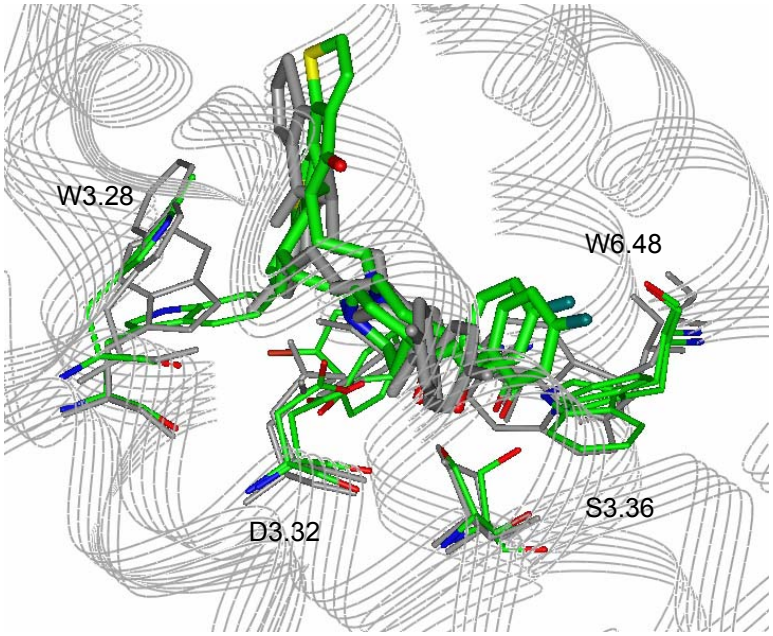


Figure 4.2.56. QF0601b in complex with 5-HT_{2A} active conformation: the ligand interacts with D3.32, S3.36, W3.28 and W6.48.

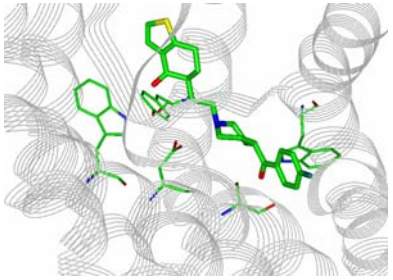


Figure 4.2.57. QF0601b in complex with 5-HT_{2A} inactive conformation. The different receptor conformation influences ligand's conformational freedom, which prevents the formation of some polar interactions.

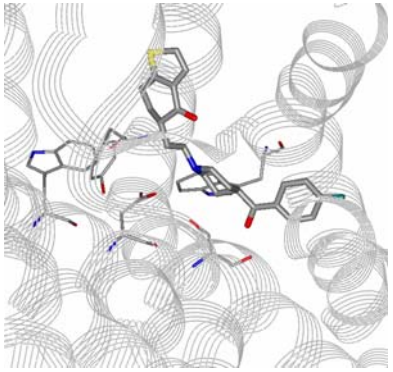


Figure 4.2.58. QF0601b in complex with 5-HT_{2C} active conformation. As in the complex with 5-HT_{2A} (active conformation), the ligand is kept in the active site through few interactions with D3.32, S3.36, W.3.28 and W6.48.

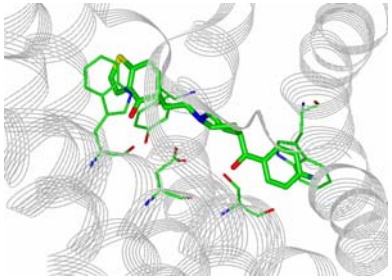


Figure 4.2.59. QF0601b in complex with 5-HT_{2C} inactive conformation: the main differences with the previous complex reside in butyrophenone ring's conformations.

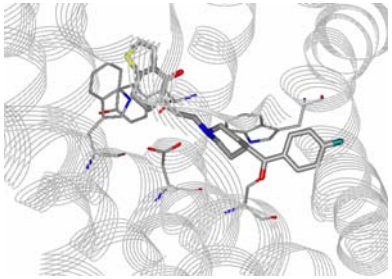


Figure 4.2.60. QF0610b in complex with 5-HT_{2A} and 5-HT_{2C} active and inactive conformations. As for the previous figures, the active conformation of the receptors is shown in green, while the inactive is shown in gray. Also, this figure presents the same orientation as figure 44, for the easier identification of residues.

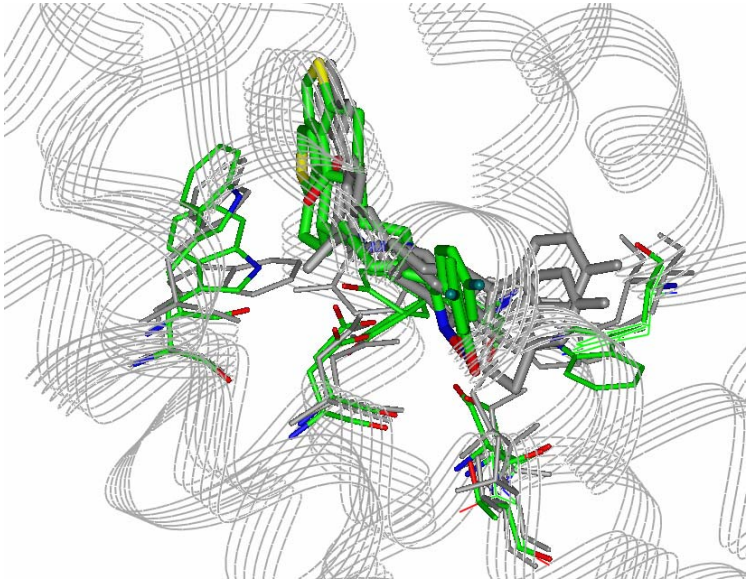


Figure 4.2.61. QF0610b in complex with 5-HT_{2A} active conformation: apart from the interactions with D3.32, S3.36, W3.28, W6.48 and Y7.43, this ligand also interacts with T3.37, hydrogen bond that is not observed in the majority of complexes.

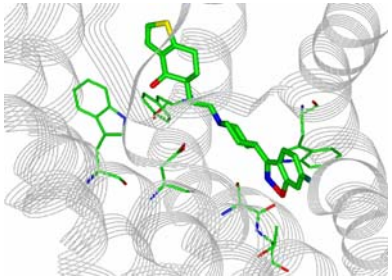


Figure 4.2.62. QF0610b in complex with 5-HT_{2A} inactive conformation: the ligand-receptor interactions, even if weaker, involve the same residues as the complex with the active conformation. The main difference between the two complexes is related with the orientation of the heteroaromatic ring.

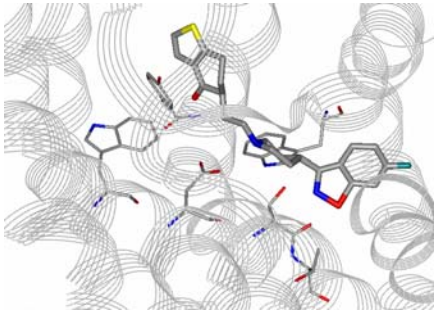


Figure 4.2.63. QF0610b in complex with 5-HT_{2C} active conformation: the hydrogen bonds and polar interactions that stabilize the complex are similar to the networks previously observed.

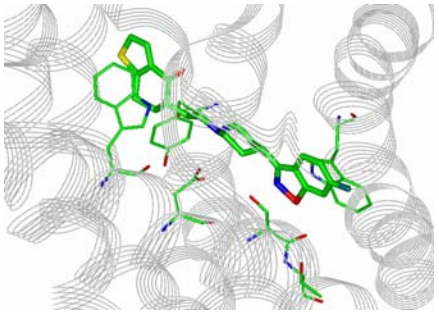


Figure 4.2.64. QF0610b in complex with 5-HT_{2C} inactive conformation: also for this complex, the ligand establishes similar interactions as previous structures.

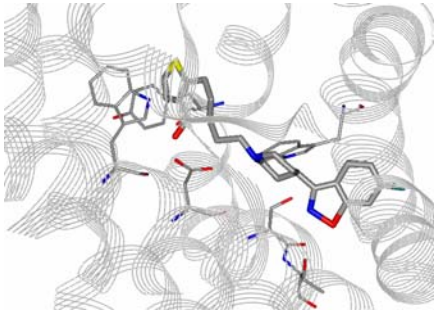


Figure 4.2.65. QF0703b in complex with 5-HT_{2A} and 5-HT_{2C} active and inactive conformation. The inactive conformation is depicted in gray, while the active conformation is shown in green. Together with the ligand, the side chains of the following residues are explicitly represented: D3.32, D3.36, T3.37, W3.28, W6.48, N7.36, Y7.43, D5.36, and N5.38.

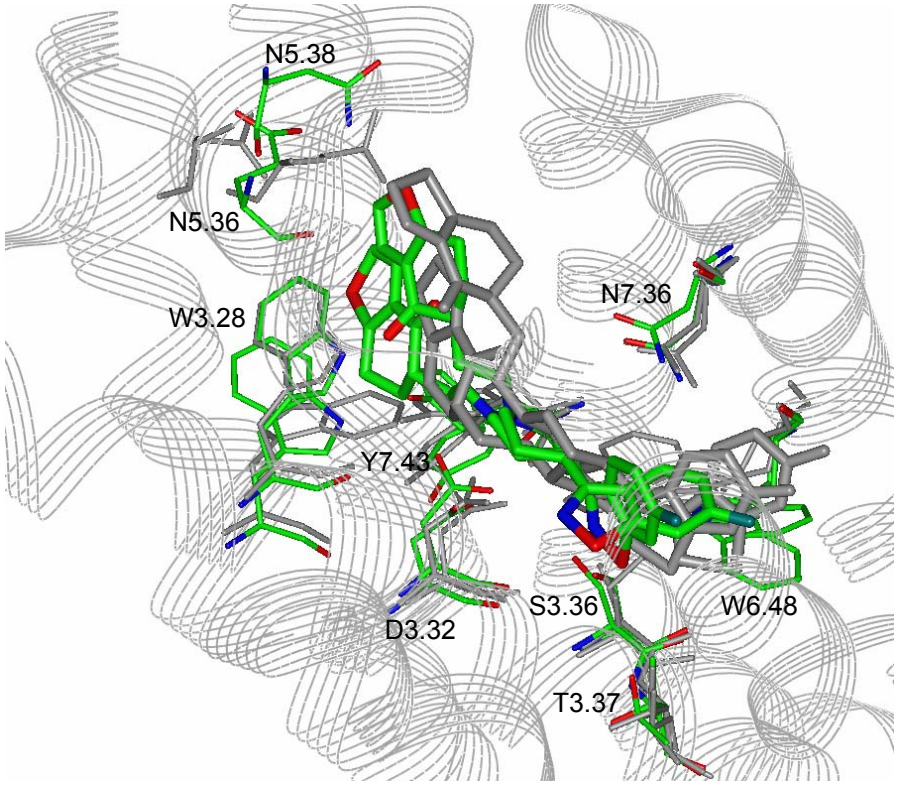


Figure 4.2.66. QF0703b in complex with 5-HT_{2A} active conformation: this complex unexpectedly forms only few polar interactions. The pKi is comparable to the previous ligands, but the strongest hydrogen bonds are with D3.32 and W3.28, even if the ligand is located near other polar residues that could interact.

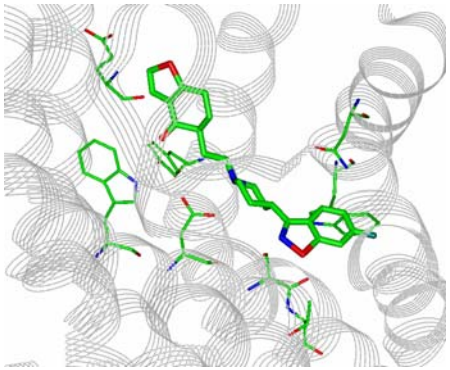


Figure 4.2.67. QF0703b in complex with 5-HT_{2A} inactive conformation: the ligand establishes interactions with D3.32, S3.36 and E159.

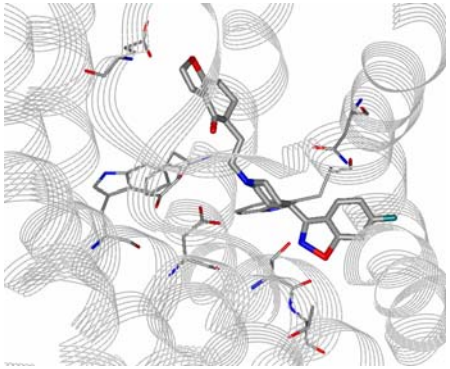


Figure 4.2.68. QF0703b in complex with 5-HT_{2C} active conformation: D3.32, S3.36 and T3.37 are the residues involved in the strongest polar interactions.

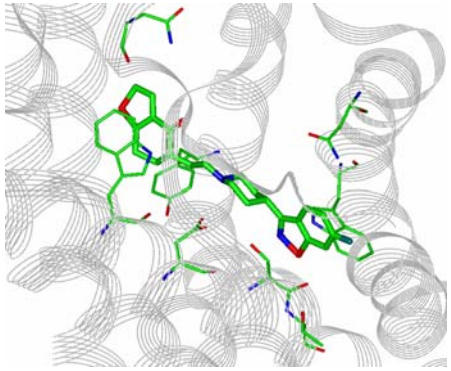


Figure 4.2.69. QF0703b in complex with 5-HT_{2C} inactive conformation: the ligand establishes stable hydrogen bonds only with D3.32 and N7.36.

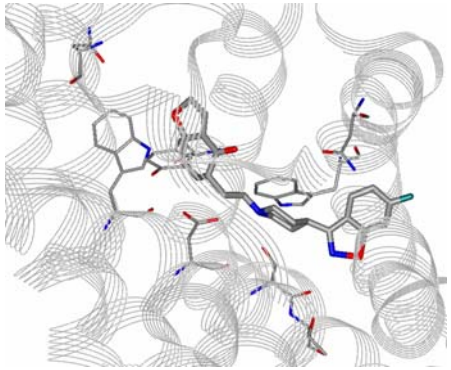


Figure 4.2.70. QF1004b in complex with 5-HT_{2A} and 5-HT_{2C} active and inactive conformations: as in the rest of figures, the active conformation is shown in green, while the inactive conformation is shown in gray. Through the superimposition of all the four complexes it is possible to appreciate the main conformational differences that occur between active / inactive complexes and between 5-HT_{2A} / 5-HT_{2C} complexes. In this latter case, the main differences reside in the orientation of the butyrophenone and heteroaromatic rings, as previously observed in other structures (figures 4.2.50 – 4.2.56).

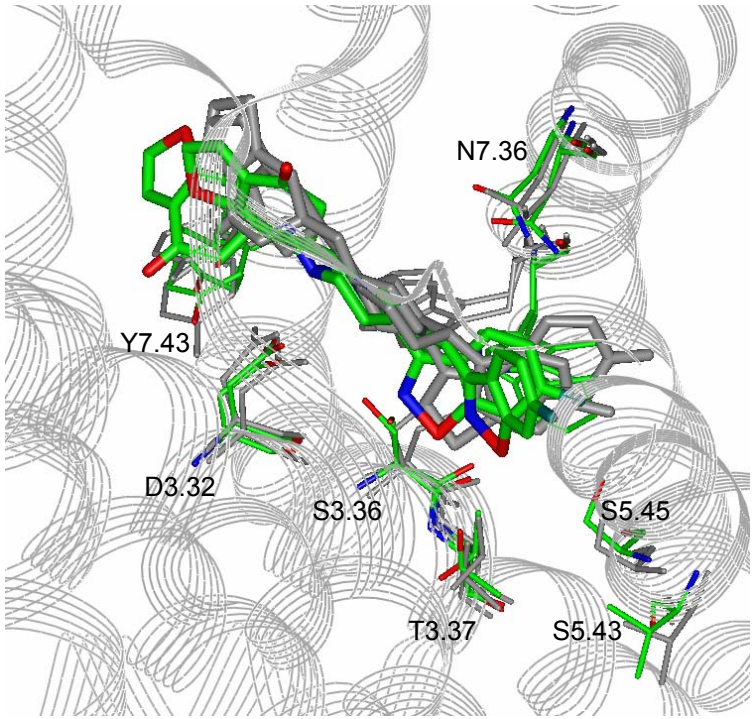


Figure 4.2.71. QF1004b in complex with 5-HT_{2A} active conformation. Also this complex, as the one in figure 4.2.60, forms very few polar interactions, as the ligand interacts especially with only two residues: D3.32 and W6.48.

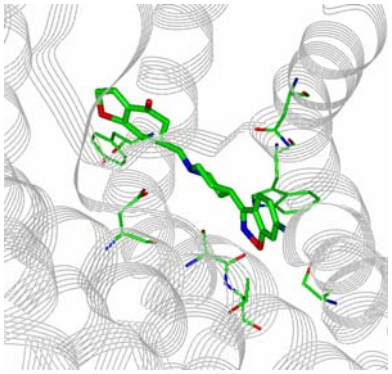


Figure 4.2.72. QF1004b in complex with 5-HT_{2A} inactive conformation. In this complex, the ligand is hydrogen bonded to S2.45, D3.32, S3.36, S5.45 and W6.48.

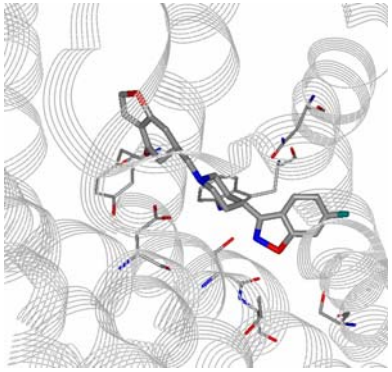


Figure 4.2.73. QF1004b in complex with 5-HT_{2C} active conformation. D3.32, T3.37 and N7.36 stabilize the ligand in the active site through hydrogen bonds with the polar groups of QF1004b.

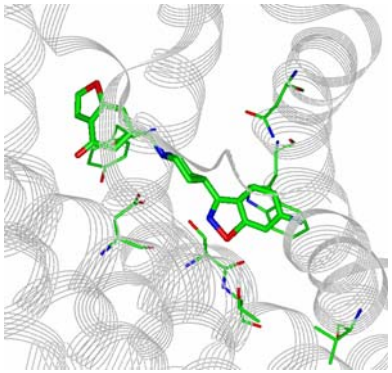
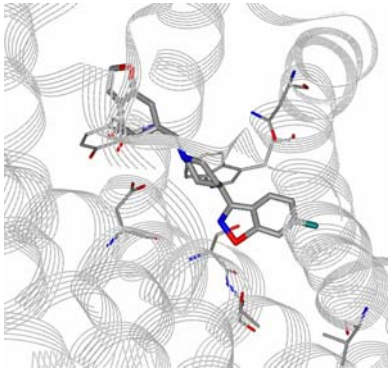


Figure 4.2.74. QF1004b in complex with 5-HT_{2C} inactive conformation. The most relevant hydrogen bonds are established with D3.32, T3.37, N7.36 and Y7.43.



By comparing the ligand – $5HT_{2A}$ complexes versus ligand – 5-HT_{2C} complexes, and also the complexes with ketanserin versus the complexes with butyrophenones, interesting conclusions could be drawn. Ketanserin establishes with both receptors polar interactions within the binding pocket with the side chains of D3.32, S3.36, W6.48 and Y7.43. Interestingly, this ligand is able to interact with residues located in the second extracellular loop (EL2), such as K155 and E159 (even if in 5-HT_{2C} complex the hydrogen bond with K155 is weaker and with E159 is absent due to differences in the sequence).

QF0601b and QF0610b interact with the same residues in the binding pocket of both receptors: D3.32, S3.36, W3.28, W6.48, for QF0601b and D3.32, S3.36, W3.28, W6.48, Y7.43 and T3.37, for QF0610b. During the MD simulations, both ligands explore with the heteroaromatic substituents two different hydrophobic pockets. The first one, located between helices TMHIV, TMHV, and TMHVI, is formed by the side chains of I4.56 (V4.56 in 5-HT_{2C}), I4.60, V5.39, L5.40, F5.47, W6.48, F6.51, F6.52 and I6.56. The second one is located between helices TMHVI, TMHVII and TMHII and it is formed by the side chains of W6.48, F6.51, F6.52, I6.56, L7.35, V7.39, W7.40, and I2.48. F6.52, which belongs to the aromatic toggle switch, seems to interact with the ligands by forming a cation – π stacking interaction with the piperidine ring of butyrophenones.

F6.51 and F5.47 delimit the bottom of the binding pocket, and their conformational changes during the MD simulation are coordinated, maintaining the π- π stacking interaction between their aromatic rings. This interesting behaviour could be related with some role of these residues in the conformational changes triggered by receptor activation.

The similarities between the docking of QF0601b and of QF0610b are not easily justifiable when comparing the two ligands on a structural level: in fact these compounds share the same butyrophenone scaffold but have very different

heteroaromatic substituents, in terms of shape, polarity, hydrophobicity and steric occupancy.

QF0703b and QF1004b form hydrogen bonds with D3.32, S3.36, W3.28 and W6.48, like the previous compounds, but also with other residues that do not seem to interact with the ligand in the previous complexes, such as: E159, S5.45, T3.37, N7.36, and Y7.43. Interestingly, the docking of these two ligand show the major differences, in term of polar interactions established, than the previous dockings. QF0703b interacts with E159 when complexed with 5-HT_{2A}, but not in complex with 5-HT_{2C}, due to the sequence differences in the two receptors. In complex with 5-HT_{2C} this ligand interacts with T3.37 (that follows S3.36, deeper in the binding pocket) and N7.36 (in the outer part of the binding pocket).

Instead QF1004b interacts with S5.45 (one of the residues important for the docking of the natural ligand, 5-HT, see Table IV) when docked into 5-HT_{2A}, and with T3.37, N7.36 and Y7.43 when docked into 5-HT_{2C}. These ligands explore preferably the hydrophobic pocket located between helices TMHIV, TMHV, TMHVI.

When comparing the docking of ketanserin and of butyrophenones, the major differences are in the hydrogen bonds with K155 and Y7.43, established by ketanserin but not by butyrophenones, and with S5.45, T3.37, N7.36, established by butyrophenones but not by ketanserin.

All in all and, after analysing the results for the four ligands considered, we could conclude that:

1. The docking and molecular dynamic simulations identify the following residues as the most important for the ligand receptor interaction: D3.32, S3.36, S5.45, and Y7.43, which establish strong hydrogen bonds with the ligands,

2. The docking and molecular dynamic simulations also identify W6.48, F6.51 and F6.52, as important to stabilize the ligand-receptor complex. F6.52 seems to establish interactions cation - π with the positive charge of the piperidine ring and π – π interactions with other aromatic residues (such as F5.47).

3. Notably, hydrogen bond with W6.48 occur only in the "active" conformations of receptors, while hydrogen bond with Y7.43 could be established in both "active" and "inactive" conformations, as well as the putative cation – π stacking interaction between F6.52 and the piperidine ring of butyrophenones.

4. In some of the ligands, a certain preference to interact with the inactive form of the receptors has been identified, as shown by the hydrogen bond statistics that are correlated with the stability of the complexes during the molecular dynamic simulations. This result is also coherent with the activity as antagonists to 5-HT_{2A} and 5-HT_{2C} receptors of these compounds.

5. In the same sense, the results indicate that the complexes with 5-HT_{2A} receptor were slightly more stable than the complexes with 5-HT_{2C} receptor, evidence that is in agreement with the selectivity of these compounds for the serotonin subtype 2A receptor.

In summary, some understanding was achieved with these experiments about the role of the aromatic cluster (W6.48, F6.51, F6.52 and Y7.43) in the "active" and "inactive" form of the receptors in complex with different ligands, but some other questions remained unclear, as if there is a preferred bioactive ligand orientation within the binding pocket ("A" or "B") and if the R and S enantiomers show differences in the interaction

with the receptors, and which is the relation with the pharmacological profile, if there is any. These results were therefore used as the starting point of further studies, more focussed only on the aforementioned questions, which will be the subject of the next sections.

Therefore, in an attempt to complete the study by answering all these questions, longer molecular dynamics simulations (2 ns) were carried out with the complexes ligand – 5-HT_{2A} receptor in the inactive form: the results are presented in the following chapters.

Molecular dynamic simulations of the 5-HT_{2A} receptor (inactive form) in complex with ketanserin, QF0601b, QF0610b, QF0703b, QF1004b ("B orientation")

The MD simulations reported above have produced interesting, even if not definitive, outcomes that provide some clues about the role of the aromatic cluster in ligand binding.

Anyway, some questions remained unanswered about the docking of butyrophenones into the active site:

1. whether the "A" orientation chosen for the binding position of ketanserin could be generalized to all the butyrophenones

2. whether the R and S enantiomers exhibit consistent differences in the interaction with the receptors and if these differences are related to their pharmacological profile

Aiming at answering these questions, additional 2 ns MD calculations were carried out with the ligands QF0601b, QF0610b, QF0703b and QF1004b in complex with 5-HT_{2A} (inactive conformation), starting with the "B orientation" of ligands (as depicted in figure 4.2.32). Thus, the butyrophenones were inserted into the binding site with the

heteroaromatic substituent towards the extracellular side of the receptor.

These new runs of experiments allowed to compare the energetic stability of the ligand-receptor complexes based on the “B orientation” with the previous one, based on the “A orientation”, and therefore to understand if there was a preferred bioactive position, and which one was it. The energetic stability of the complexes was described and evaluated in a quantitative way, as for previous molecular dynamics simulations, by measuring total energy, temperature, hydrogen bond statistics, hydrophobic contacts, and rms of the ligand.

Complexes with “B orientation” of butyrophenones were obtained from the complex ketanserin – 5-HT_{2A} (“B orientation” of ketanserin and inactive form of the receptor) through a cycle of MD simulations and energy minimizations. In summary, the MD frame representing the conformation at lowest energy of ketanserin – 5-HT_{2A} was retrieved from the MD trajectory and energy minimized. The equilibrated structure was therefore used as the starting position for the docking of butyrophenones using GOLD, following the protocol previously described for the docking of ligand in the “A orientation”.

The new complexes were then submitted to MD simulations according to the established protocol, and reached the equilibrium of total energy, temperature and RMS of the ligands after 200 ps approximately. In general, the analysis of these MD runs showed rather similar results as the previous MD simulations, being the main differences in the number of polar interactions established in the active site. Butyrophenones interacted with a minor number of residues, even if the hydrogen bonds were more persistent (see Figure 4.2.75, and figures 4.2.76 – 4.2.80).

Figure 4.2.75. The graph shows the hydrogen bond network for the 5-HT_{2A} (“inactive”) in complex with ketanserin and butyrophenones.

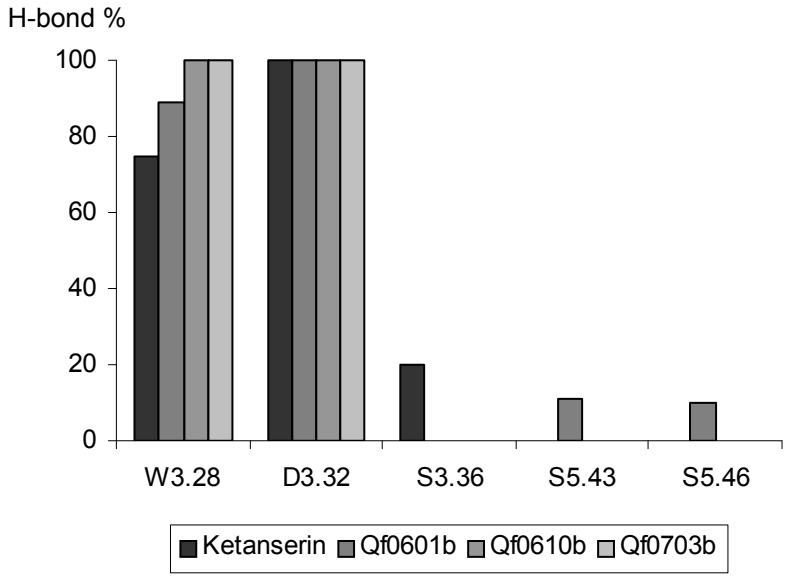
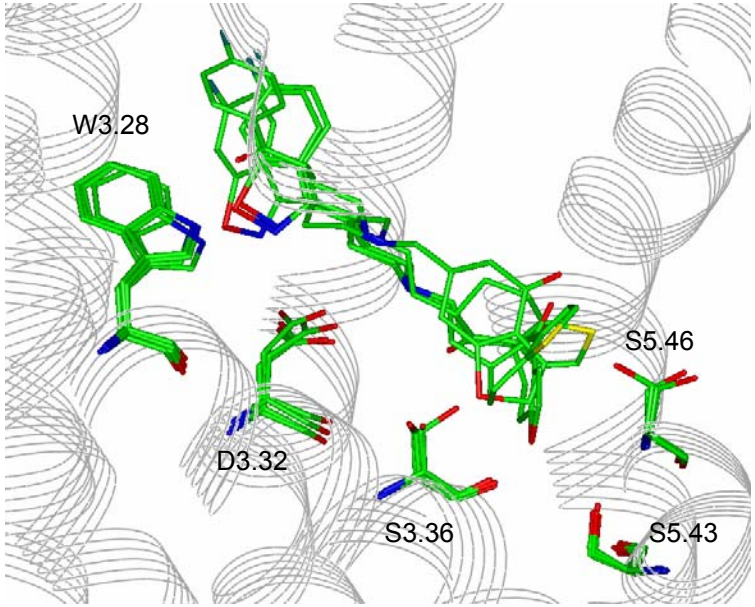


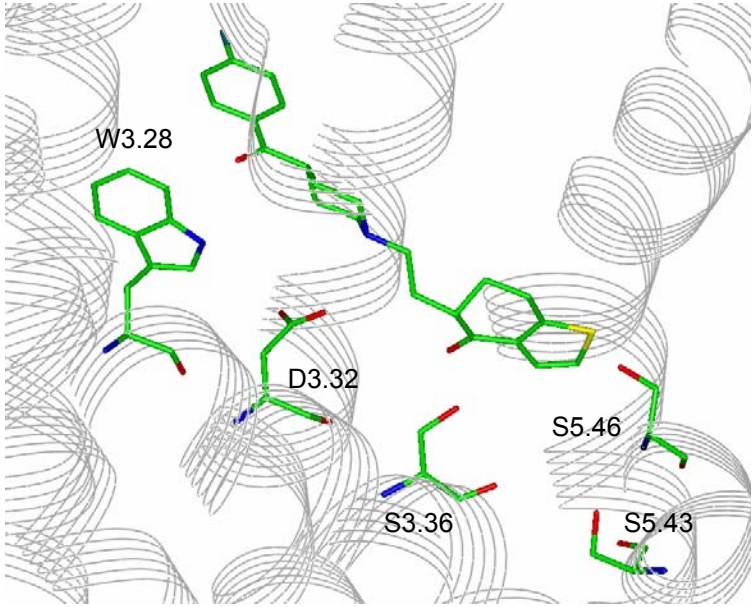
Figure 4.2.76. QF0601b, QF0610b, QF0703b and QF1004b in complex with 5-HT_{2A} (inactive form). The complexes minimized are superimposed and the most important residues for docking interactions are shown.



5-HT_{2A}R-QF0601b in B orientation complex

As found in previous studies, the most important interaction between the ligand and the receptor is the ionic bond between the charged nitrogen of the piperidine ring and the negative side chain of D3.32. The butyrophenone head interacts with the side chain of W3.28, and other important hydrogen bonds involve the two serines of helix V S5.43 and S5.46 (that also stabilize the natural agonist in the active site, according to mutagenesis).

Figure 4.2.77. QF0601b in complex with 5-HT_{2A} inactive form.

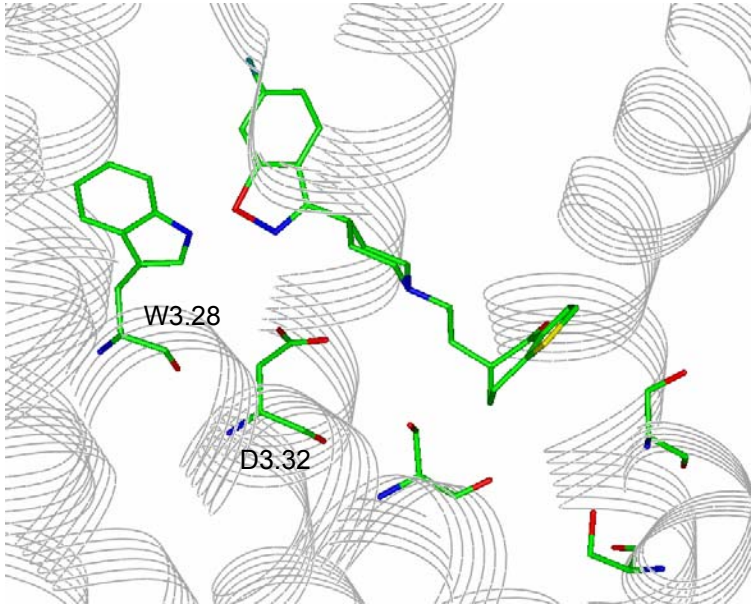


To make the comparison easier, it could be useful to remind that in the complexes with “A orientation” and receptor active form, this ligand could interact with D3.32, and W3.28 as well, but also with S3.36 and W6.48. Instead for QF0601b “A orientation” in complex with 5-HT_{2A} inactive conformation, the different receptor conformation influences ligand’s conformational freedom, that prevents the formation of some polar interactions.

5-HT_{2A}R-QF0610b in B orientation complex

In the case of this complex, the ligand is kept in the active site through only two hydrogen bonds, with the side chains of D3.32 and W3.28. Other residues are shown for clarity to facilitate the comparison with the dockings of the other butyrophenones.

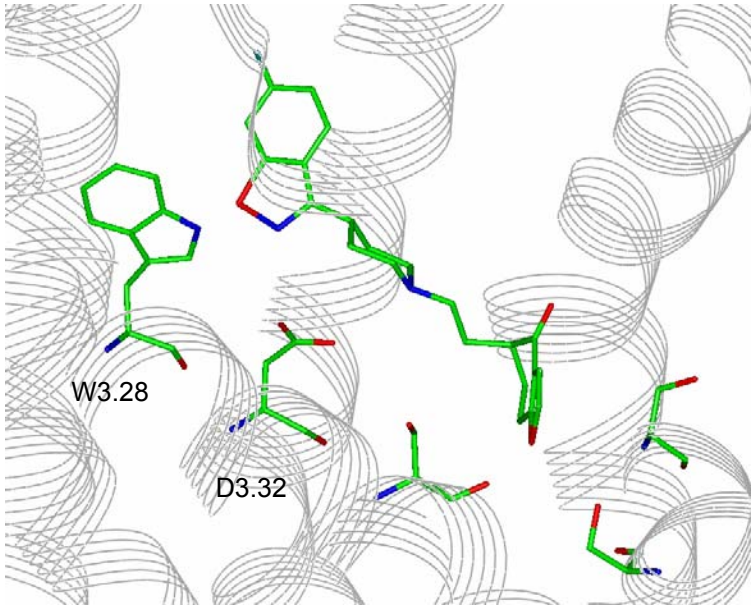
Figure 4.2.78. QF0610b in complex with 5-HT_{2A} inactive form.



5-HT_{2A}R-QF0703b in B orientation complex

QF0703b and QF1004b show very similar interactions with the active site. Also for 5-HT_{2A} in complex with QF0703b, the ligand is kept in the active site through only two hydrogen bonds, with the side chains of D3.32 and W3.28. In the figures, other residues are shown for clarity to ease the comparison with previous dockings.

Figure 4.2.79. QF0703b in complex with 5-HT_{2A} inactive form. The ligand interact mainly with W3.28 and D3.32: other side chains that define the active site are shown to allow a better comparison with other figures.



5-HT$_{2A}$R-QF1004b in B orientation complex

The ligand QF1004b, similarly to other butyrophenones, interacts strongly with the side chains of W3.28 and D3.32, and in minor extent with S3.36. Probably, the extension of this ligand, that is the shortest of the group used in these experiments, is responsible for the lack of other polar interactions as well as for the flexibility shown during the MD simulation.

In fact, during the MD simulation the ligand could adopt two different conformations: the first one with the butyrophenone moiety oriented toward the external part of the receptor (Figure 4.2.80) and another with this moiety rotated by 180°, allowing the interaction with S3.36. This behaviour is also reflected in the ligand RMS, which is noticeably less equilibrated than for other longer compounds (Figure 4.2.81).

Figure 4.2.80. QF1004b in complex with 5-HT$_{2A}$ inactive form: main polar interactions involve D3.32, W3.28 and S3.36

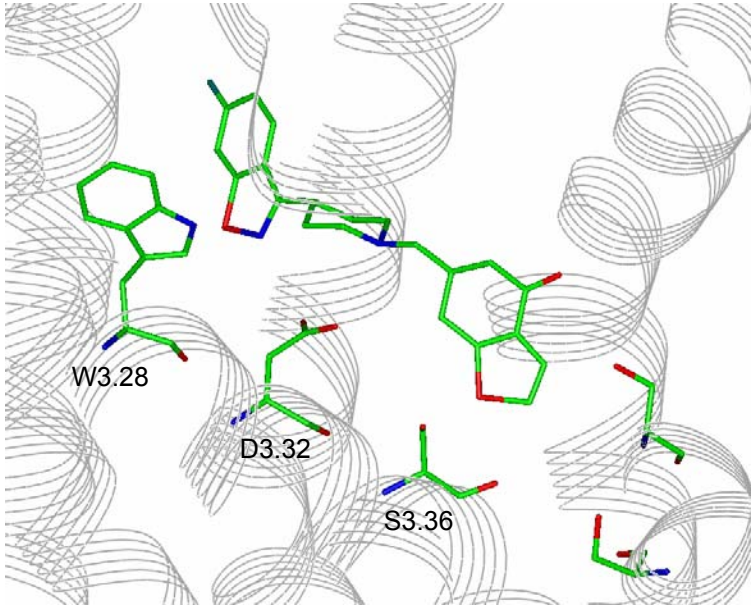
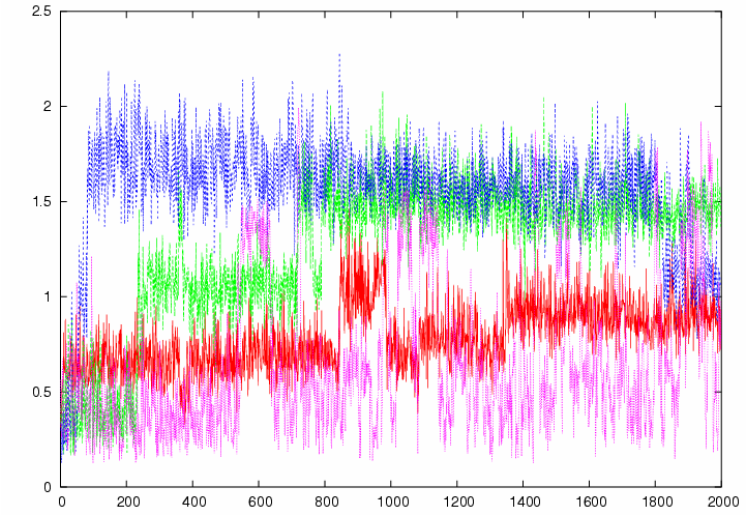


Figure 4.2.81. RMS of ligands QF0601b (blue line), QF0610b (green line), QF0703b (red line) and QF1004b (pink line) in the MD simulations analysed (2 ns). The graphs provide a qualitative picture of the relative stability of ligands during the simulation.

Even if RMS of QF1004b is the lowest, it is also the less equilibrated: during the 2 ns of MD, two abrupt changes are observable after 500 ps and around the 800 ps, while after 900 ps the ligand position is rather more unstable.



The changes in QF1004b RMS after 500 and 800 ps, are related with conformational changes in the ligand: the butyrophenone moiety turns 180° in order to establish a π-stacking interaction with F5.47 and F6.52. F5.47 and F6.52 (that belongs to the “aromatic toggle switch)” are also interacting via π- π stacking.

At the same time, the fluor-benzoyl ring also turns to establish a π-stacking interaction with the aromatic ring of W3.28. The two movements, the turn of butyrophenone and the turn of the heteroaromatic substituent, seem therefore to be correlated, and to disturb the interaction of F5.47 with F6.52.

In summary, from this second run of MD simulations some general conclusions could be drawn:

1. The key residues for interaction with the receptor are clearly identified and confirm the results of previous MD simulations.

2. "A" and "B" ligand orientations within the active site are both accessible, at least according to the results achieved with the techniques applied. This result reinforce previous findings and is also in agreement with a multistructure 3D-QSAR study of the butyrophenone series docked into the current 5-HT_{2A} ("inactive") model (Dezi C., Brea J. et al. 2007), which arose from this thesis and is presented later in the text.

3. Essential requirements for activity towards these serotonin receptors appear to be the presence of two aromatic rings with polar groups, at a certain distance (at about 4 – 5 Å), with a positively charged nitrogen located in the middle (about 9 Å).

4. Since the ligands have almost the same pKi for the binding with 5-HT_{2A}, we would not expect relevant differences in the ligand – receptor interactions for complexes with 5-HT_{2A}, as it was confirmed by these MD simulations. Anyway, the models were unable to explain the differences in the pharmacological behaviour of QF0703b and QF1004b, which probably would require more sophisticated computational or experimental techniques

5. The networks of hydrogen bonds formed by QF0703b and QF1004b are quite similar. This result is not surprising, due to the structural similarities between these two ligands, but doesn't allow explaining the differences observed in their pharmacology (Table X, pag. 109). Such differences are probably due to other factors, not recognisable by MD simulations because of the limitations of the method, including the afforded time scale (2 ns). Hypothetically, explanation of their behaviour could reside into the dimerization of the receptors, the bioactive conformation of the ligands, their orientation in the active site (the compounds could interact with the receptor in "A" or "B" orientation or in a combination of both), or their chirality (as the majority of binding experiments for butyrophenones were performed with the racemic mixture of R and S enantiomers). The complete explanation of these phenomena was beyond the scope of the present thesis, and more experimental data are needed to fully understand the atomic interactions below the pharmacology. Anyway, the methods applied, provided some results that could be an interesting support to further research.

To sum up, only the question related with chirality remained unsolved: whether R and S enantiomers show differences in the interaction with the receptors and the relationship between such differences and ligand's pharmacology.

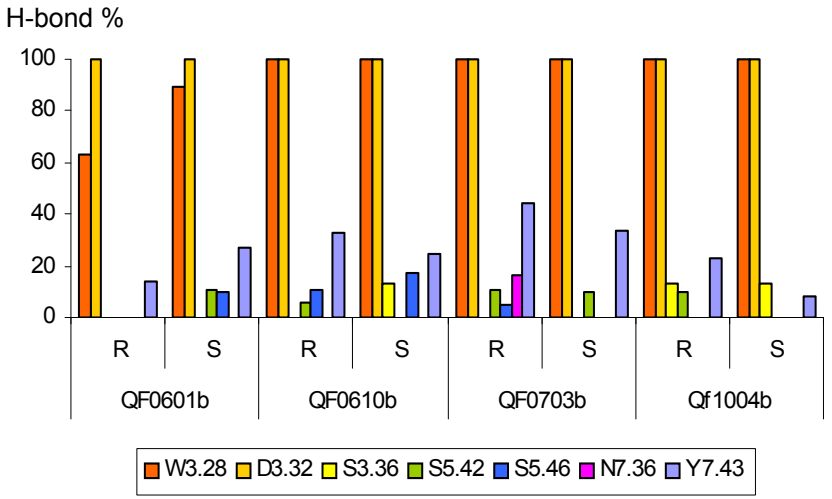
As mentioned before, unlike ketanserin, all the compounds in the butyrophenone's series contain a chiral centre and, therefore, either their R or S enantiomers could interact with the binding site. To address this issue, a computational study was carried out with both enantiomers. R and S enantiomers

were built in 2D and converted to 3D structures using CORINA 2.4. The structures were all modelled with a formal positive charge centred on the piperidine ring nitrogen. QF0601b, QF0610b, QF0703b and QF1004b, modelled in R and S enantiomers are shown in figure 4.2.83: both enantiomers were docked into the 5-HT_{2A} ("inactive") binding site and the resulting 8 complexes were submitted to 1 ns MD calculation.

Molecular dynamic simulations of the 5-HT_{2A}R (inactive form)-QF0601b, 5-HT_{2A}R-QF0610b, 5-HT_{2A}R-QF0703b, 5-HT_{2A}R-QF1004b complexes (R and S enantiomers)

As in previous simulations, these complexes reached early the equilibrium of total energy, temperature and RMS of the ligands, probably because they were derived from well equilibrated structures. The major differences observed between these complexes and the previous, resided in both the polar and hydrophobic interactions established, as some of the ligands explored different pockets during the simulation. A comprehensive picture illustrating the main polar interactions is provided in figure 4.2.82.

Figure 4.2.82. Hydrogen bond network for the 5-HT_{2A} ("inactive") in complex with R and S enantiomers of the butyrophenones.



The observation of Figure 82 allows the easy identification of the key interactions, for all the ligands, and of some general trend in polar interactions shared between R and S enantiomers. For example, the interactions with W3.28, D3.32 and Y7.43 are common to all the dockings, for R and S enantiomers. Instead the interaction with S3.36, found relevant in previous studies, seems important only for QF0610b (S) and QF1004b (R and S).

The interaction with S5.43 and with S5.46 is present in all the dockings, at least with one entantiomer, suggesting that these residues are essential for binding. Only QF1004b does not interact with S5.46, but both R and S enantiomers bind S3.36.

Probably due to its length, this compound establishes stronger interactions near D3.32 and is unable to reach the

bottom of the binding pocket, interacting with residues located there, as longer compounds do. In general, analyzing the graph, S enantiomers of QF0601b and QF0610b seem more stable, while the contrary happen for QF0703b and QF1004b.

Instead, R enantiomers interact stronger with Y7.43 side chain (apart from QF0601b). It is also remarkable that these compounds (R and S enantiomers) are unable to interact simultaneously with S3.36, S5.43 and S5.46, which instead are considered essential for the docking of the natural agonist 5-HT. Providing explanation for this finding is not an easy task: maybe the length of the ligands that make them rather less flexible than the short 5-HT or maybe the limitation of the method used (time scale of MD simulations or receptor rigidity). Alternatively, this result could be related with the antagonist / partial agonist nature of these butyrophenones.

In summary we could conclude that for QF0601b and QF0610b the R enantiomer seems to establish stronger interactions than the S enantiomer, while for the QF0703b and QF1004b the contrary happens. Despite these differences all ligands establish strong hydrogen bonds with W3.28, D3.32 and Y7.43. Other residues are involved in polar interactions to some extent, but these three residues seems to be the most relevant for binding, as also shown in previous MD simulations.

It is remarkable that for QF1004b (for both enantiomers), significant conformational changes, similar to the previously observed turns, occur during the MD simulations: after 600 and 800 ps, the butyrophenone moiety turns 180° in order to establish a π-stacking interaction with F5.47 and F6.52 (that belongs to the “aromatic toggle switch)”.

These two phenyalanines are also interacting via π- π stacking. At the same time, the fluor-benzoyl ring also turns to establish a π-stacking interaction with the aromatic ring of W3.28. The two movements, the turn of butyrophenone and the turn of the heteroaromatic substituent, seem therefore to be correlated, and to disturb the interaction of F5.47 with F6.52.

Figure 4.2.83. R and S enantiomers of QF0601b, QF0610b, QF0703b and QF1004b. The occupancy of the binding site is approximately parallel to the helices, slightly tilted with respect to the axis of the helices bundle and nearly parallel to the helix TMHIII.

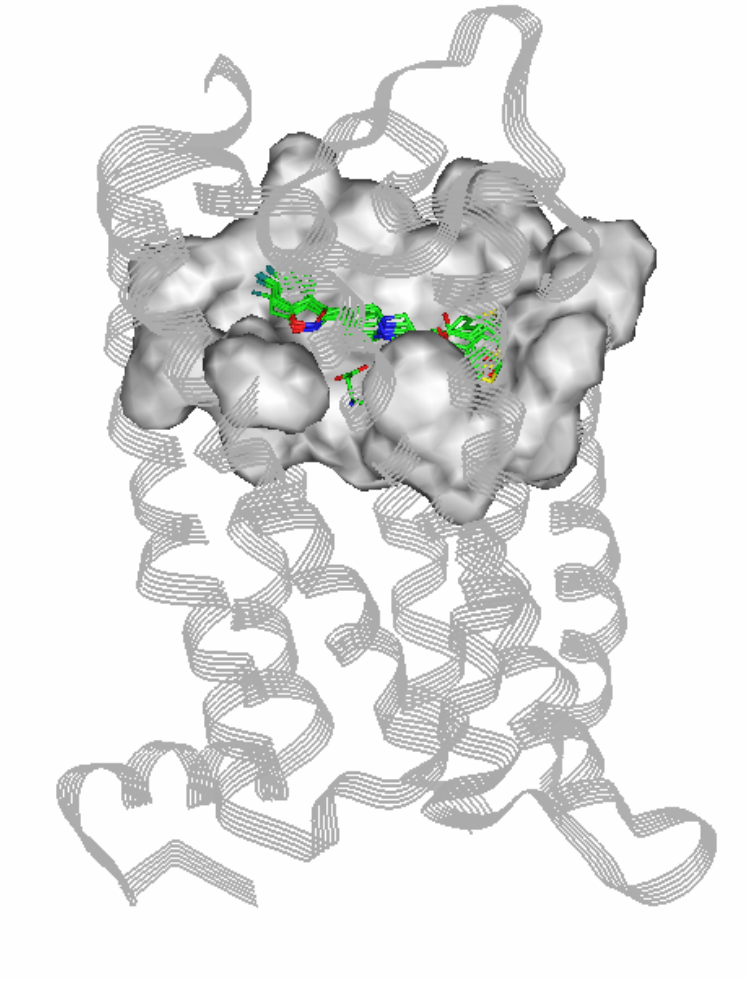


Figure 4.2.84. R and S enantiomers of QF0601b, QF0610b, QF0703b and QF1004b.

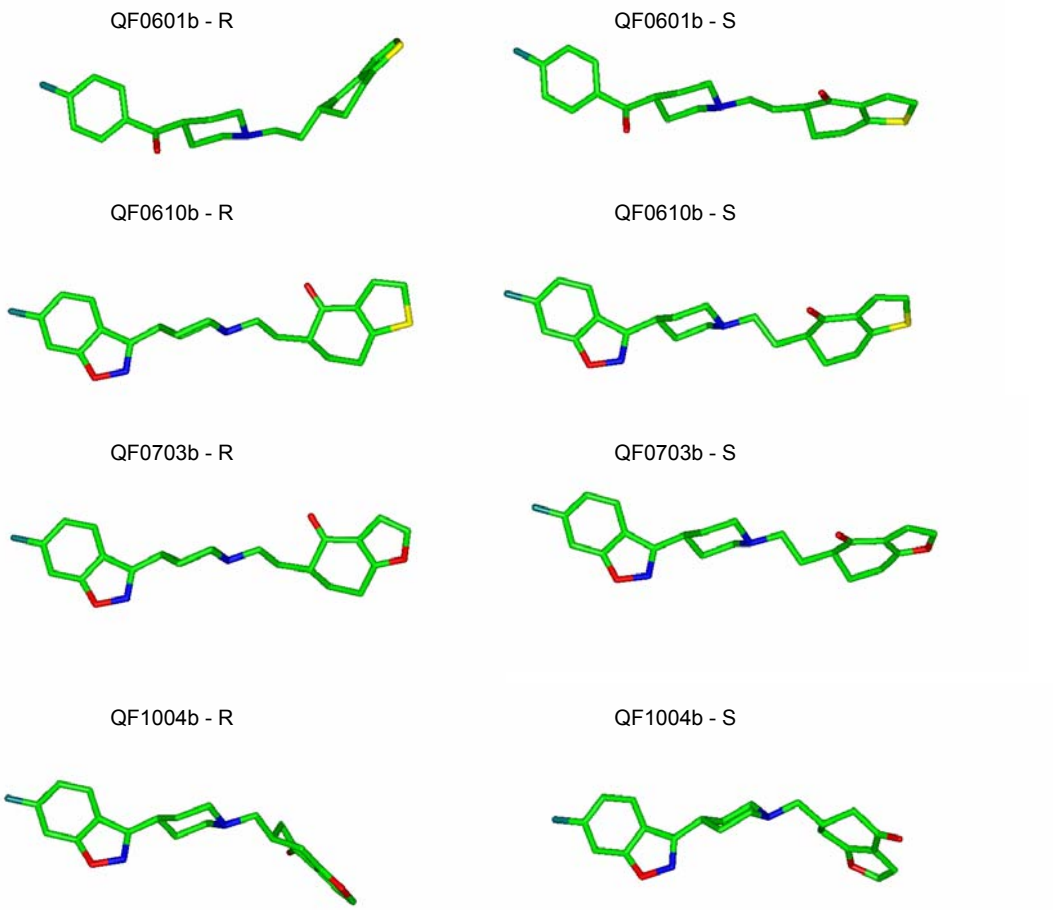
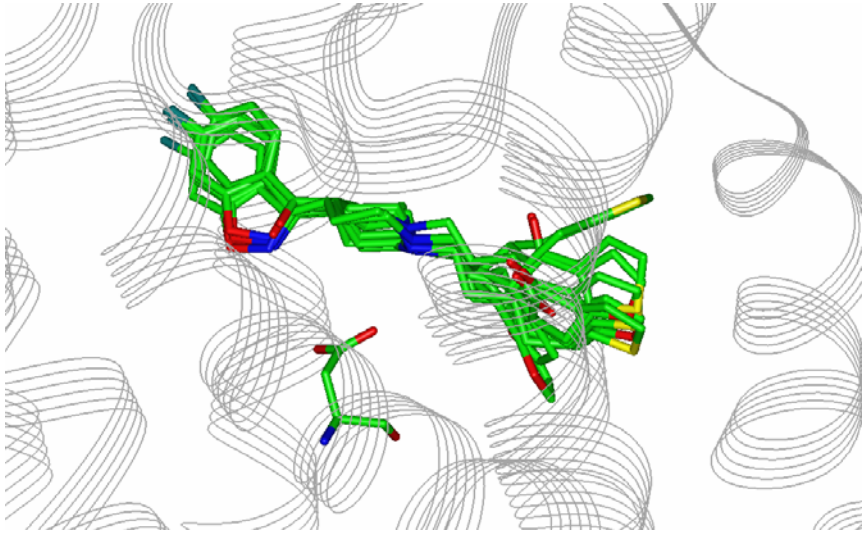


Figure 4.2.85. R and S enantiomers of QF0601b, QF0610b, QF0703b and QF1004b superimposed in the active site of 5-HT_{2A} (inactive form). D3.32 is shown to help the identification of the binding site. It is significant the conformational diversity of ligands at the bottom of the binding pocket.



In summary, the chirality of these compounds seem to play a certain role in their binding to 5-HT_{2A}, and this result could be generalized to other serotonin receptors. In enantiomers R and S, the butyrophenone ring adopts different conformations (as shown in figure 4.2.85): when butyrophenones are docked into the receptor model, such differences influence the hydrophilic and hydrophobic interactions established.

The ligand therefore has to find the optimum position to locate the carbonyl group of the butyrophenone ring within the active site. Such optimum position depends also on the length

of the ligand: to this respect, particular important seems the linker between the piperidine and the butyrophenone ring and the linker between the piperidine and the heteroaromatic substituent. The docking results and the MD simulations with the different enantiomers provided evidences for the role of these linkers (figure 4.2.82 and figure 4.2.85).

Furthermore, experimental data for some butyrophenones (not included in this study) proved that R and S enantiomers bind with different affinity to serotonin receptor (Dezi C., Brea J. et al. 2007).

3.3 GRID / GOLPE of selected complexes

One of the major problems when analyzing MD simulations is the enormous amount of data generated. This amount of data difficults the interpretation and the extraction of the most relevant information to answer the question addressed.

Fortunately, there are computational techniques that make easier the interpretation of large series of data, helping to find a correlation between experimental and structural data.

One of these techniques is the GRID/PCA method (Pastor M., Cruciani G. et al. 2000) that allows the characterization and the comparison of large set of structures.

This methodology is based on a Principal Component Analysis (PCA) of modelled interaction fields obtained using GRID to characterize the binding site of the diverse complexes obtained. In our case, GRID calculations were carried out in the binding site with certain probes that reproduced the main interacting groups in the butyrophenone series.

The purpose of the study was to identify the regions responsible for the different selectivity of ligands: GRID / PCA method allows to identifying the residues most involved in the interaction, by extracting the most relevant information for representing the data variability and then representing them into three-dimensional data. Therefore the analysis allowed

rationalizing the behaviour of the receptors in response of ligand binding, by identifying the residues whose conformation varied most during the MD simulation.

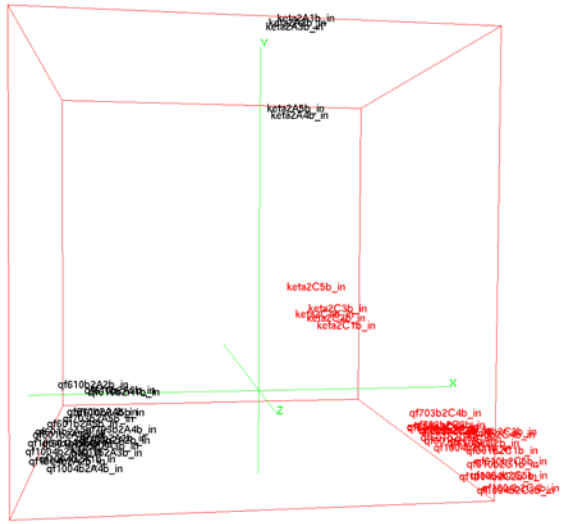
The set of molecules chosen for the study was retrieved from MD simulation trajectories of 5-HT_{2A} and 5-HT_{2C} (inactive form) in complex with ketanserin, QF601b, QF0610b, QF0703b and QF1004b. In order to obtain a proper representation of the major conformational changes that occurred during the simulations, five frames representative of the last 500 ps of each one of the MD simulations were extracted (one frame every 100 ps in the last 500 ps of simulation).

Molecular Interaction Fields (MIF) were computed with OH probe, selected as the most appropriate to map the interactions within the active site: this probe represents the phenolyc hydroxyl and could easily describe both hydrophobic and hydrophilic interactions.
In order to be sure to consider all the side chains important for ligand binding, the active site included all the residues located within a box of 16 x 26 x14 Å centred on the carboxyl group of D3.32.

The calculation was performed on 50 objects (5 frames for each one of the complexes ligand – 5-HT_{2A} and ligand – 5-HT_{2C}), resulting in 50490 variables. MIFs were then analysed, without any pre-treatment, by means of PCA in order to highlight systematic differences in the binding site structures.

In this PCA, the first component (PC1) explained the 23% of the variance, and the PC2 only the 8%. PC1 discriminates mainly between 5-HT_{2A} and 5-HT_{2C}, while PC2 discriminates between ketanserin and the complexes with butyrophenones. This result indicates that the difference between 5-HT_{2A} and 5-HT_{2C} is more relevant than the differences between ketanserin and butyrophenones series and within the butyrophenones series itself (figure 4.3.86).

Figure 4.3.86. PCA scores 3D: axis X represents the PC1, axis Y the PC2 and axis Z the PC3. Complexes with receptor 5-HT_{2A} are coloured in black, while complexes with 5-HT_{2C} are coloured in red. The frames are distributed in the PCA in four different groups, forming a tetrahedron.



As the PC1 (Figure 4.3.86) especially explains the differences between 5-HT_{2A} and 5-HT_{2C}, some of the MIF regions identify mutations between the two receptor subtypes: W3.28, S3.36, Thr3.37, and S5.46 (that is V5.46 in 5-HT_{2C}) and F6.45 (Figure 4.3.90).

The change in the conformation of W3.28 is remarkable and corresponds, probably, to a difference in ligand insertion into the active site due to the EL2. In fact, as we mentioned before (chapter 3.2 “Docking studies, energy minimization and molecular dynamic simulations”) the EL2 loop of the 5-HT_{2C} receptor is shorter than EL2 loop in 5-HT_{2A} receptor, therefore it folds deeper into the TM bundle causing a deeper insertion of the ligand within the active site.

This difference also causes a correspondent variation in W3.28 side chain conformation, which is located at the extracellular entrance of the binding site (Figure 4.3.87 – 4.3.89). In Figure 4.3.87 the GRID loadings are represented on the receptor structure: the yellow regions describe a favourable interaction between the probe and the receptor, while the blue regions represent a not favourable interaction between the OH probe and the receptor.

The loading plot allows to identify the regions responsible for the major variability explained through the PC1, which are in the most extracellular part of the receptor and partially superimposed to EL2. The analysis of the figures (Figure 4.3.87- 4.3.89) allows to appreciate the conformational and structural differences described above. EL2 is represented without the side chains, thus easing the observation of the differences in folding between the 5-HT_{2A} (coloured according to atom type) and 5-HT_{2C} (coloured in red).

GRID loadings are shown first on 5-HT_{2A} and 5-HT_{2C} superimposed (Figure 4.3.87), and then on each receptor separately (5-HT_{2A} on Figure 4.3.88 and 5-HT_{2C} on Figure 4.3.89), in order to allow a better comparison between the two receptors.

Figure 4.3.87. GRID loadings plot of PC1, the complexes with QF0601b are shown as example. The loading plot allows to identify the regions responsible for the major variability explained through the PC1. Receptor 5-HT_{2A} is coloured according to the atom types, while 5-HT_{2C} is in red. The yellow regions describe a favourable interaction between the probe and the receptor, while the blue regions represent a not favourable interaction.

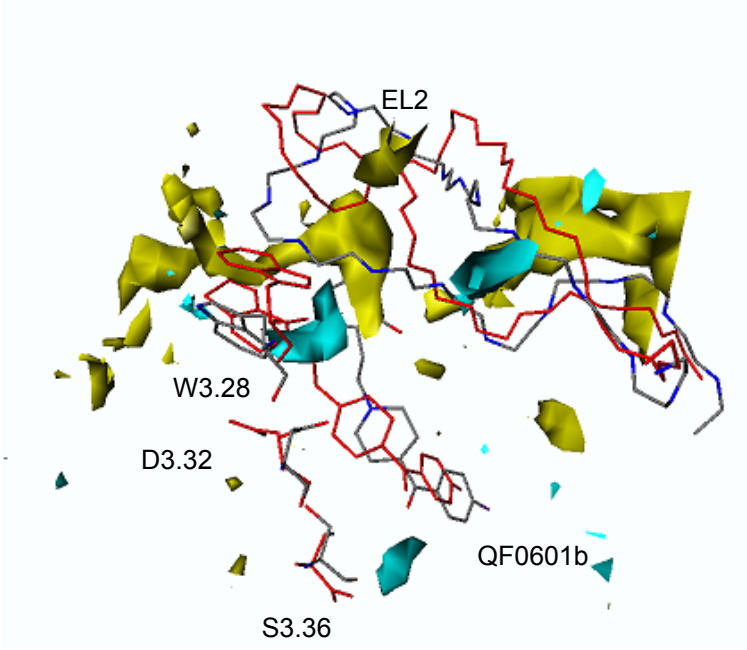


Figure 4.3.88. GRID loadings plot of PC1 represented on the receptor 5-HT_{2A}, as in the previous picture, the complex with QF0601b is taken as example. The yellow regions describe a favourable interaction between the probe and the receptor, while the blue regions represent a not favourable interaction.

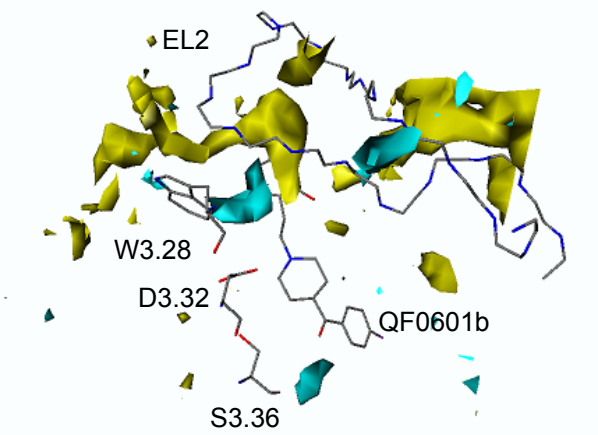
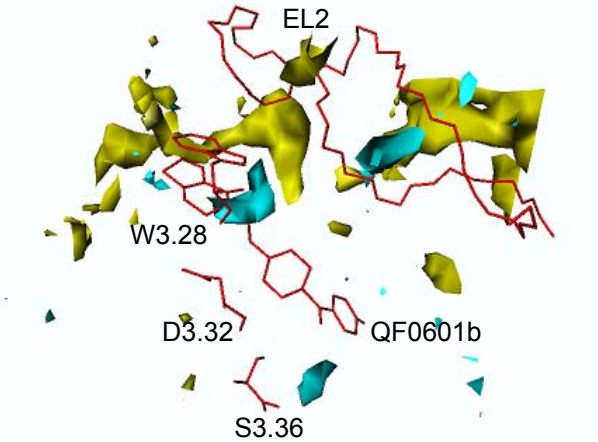


Figure 4.3.89. GRID loadings plot of PC1 represented on the receptor 5-HT_{2C}. As for the previous figures, the yellow regions describe a favourable interaction between the probe and the receptor, while the blue regions represent a not favourable interaction.



Instead, the PC2 explains mainly the differences between ketanserin – receptor complexes and butyrophenones – receptor complexes, confirming the results of previous studies (see the previous sections with the analysis of MD simulations).

The analysis centred only on QF0703b and QF1004 complexes is more interesting as it could provide some clues about the differences observed in their pharmacology. However, QF0703b complexes and QF1004b complexes are clustered in the same plot. PCA identifies as the major conformational differences the residues of the aromatic cluster involved in the activation mechanisms: W6.48, F6.51, F6.52 and Y7.43 (figure 4.3.90 – 4.3.91).

Probably, the conformational differences observed in these highly conserved residues are due to the differences in their length (QF0703b linker is one methilene longer) that allow a deeper insertion in the case of QF1004b. An interesting possibility is that these differences could also be related to the stabilization of different receptor conformations, responsible of their monophasic or biphasic behaviour.

As for PC1, GRID loadings plot for PC2, which has allowed to draw the conclusions expressed above, is shown in Figure 4.3.92.

The analysis of the picture shows as the regions identified through PC2 as the most responsible for variability are located at the bottom of the binding pocket, and correspond to the aromatic cluster (Figure 4.3.92).

Figure 4.3.90. The complexes with 5-HT_{2A} are shown in green, while 5-HT_{2C} complexes are shown in gray. QF0601b is shown as example. The superposition of the MD frames allows the appreciation of the conformational changes identified through PCA.

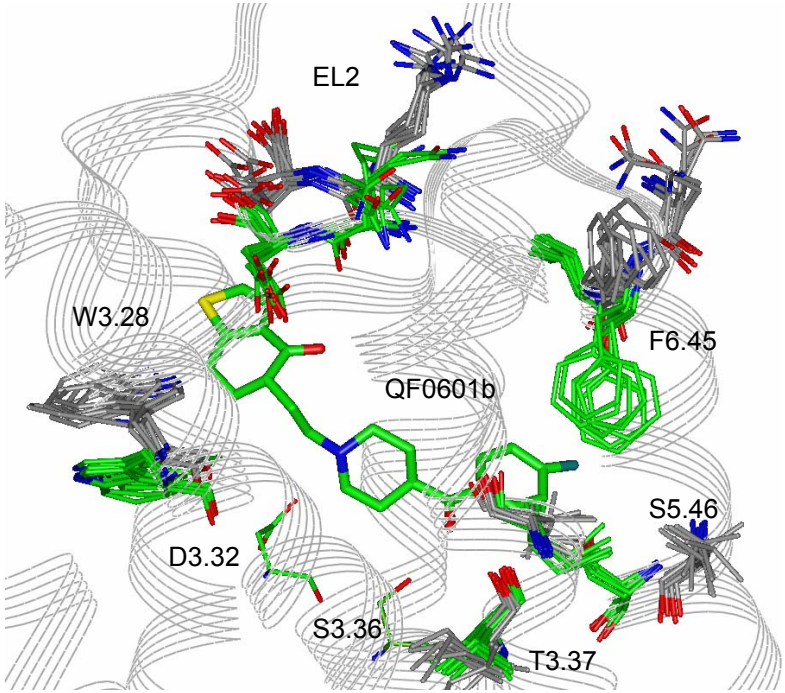


Figure 4.3.91. Complexes QF0703b – 5-HT_{2A} (green) and QF1004b – 5-HT_{2A} (gray).

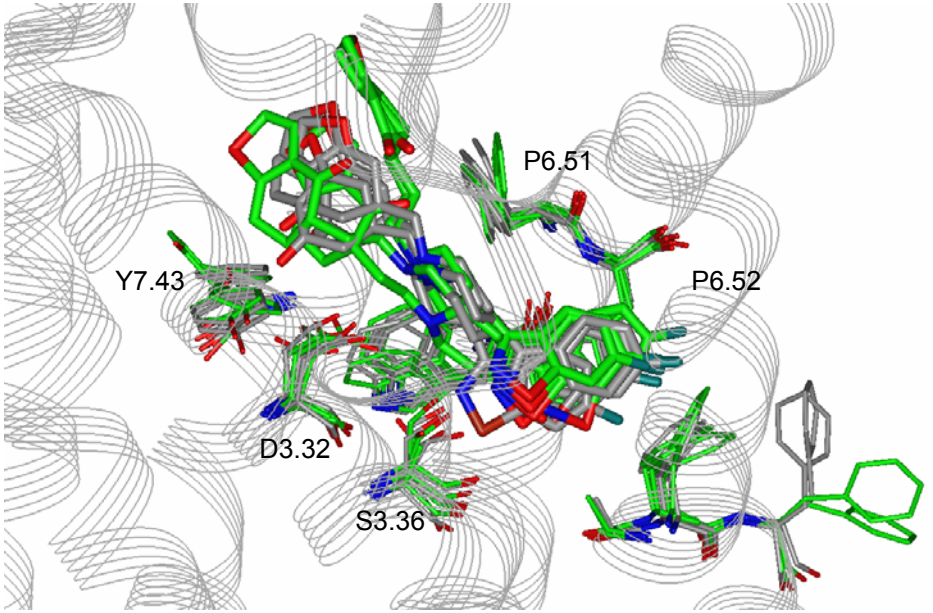
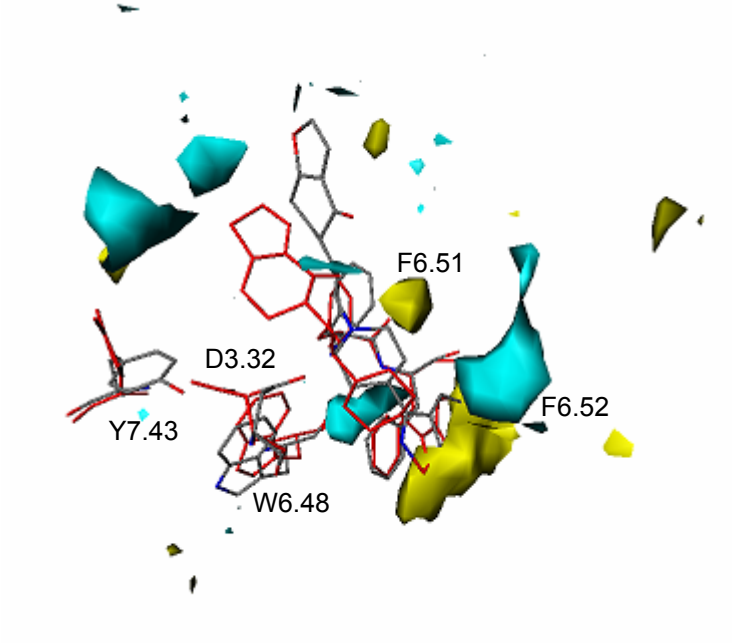


Figure 4.3.92. PC2 loadings plot for 5-HT_{2A} receptor in complex with QF0703b (coloured according to atom types) and QF1004b (in red).



All in all, the GRID/PCA analysis was an interesting tool for highlighting the more systematic differences between 5-HT_{2A} and 5-HT_{2C} when complexed with ketanserin and butyrophenones and allowed the identification of regions potentially important for drug design of selective ligands.

3.4 Multistructure 3D-QSAR Studies on a Series of Conformationally Constrained Butyrophenones Docked into a New Homology Model of the 5-HT_{2A} Receptor

The previous studies carried out on the butyrophenones QF0601b, QF0610b, QF0703b and QF1004b posed different issues related with the docking of these compounds into the active site of serotonin receptors, with the bioactive conformation, the chirality of the compounds, and the double orientation of binding.

The 4 compounds analysed, as mentioned before in the thesis (see the “Introduction” for more information) belong to a series of conformationally restricted butyrophenone analogues, still in active development, containing compounds with high affinity for the 5-HT_{2A} receptor (Chart I and II (Brea J., Rodrigo J. et al. 2002; Brea J., Masaguer C.F. et al. 2003; Brea J., Castro M. et al. 2006) and with particularly interesting pharmacological profiles (Brea J., Rodrigo J. et al. 2002; Brea J., Masaguer C.F. et al. 2003; Brea J., Castro M. et al. 2006).

The computational studies carried out aimed to unveil the structural determinants of their receptor binding affinity and selectivity and to identify the receptor binding profile, which correlates with their therapeutic usefulness. The computational methods used on the 4 butyrophenones were useful for obtaining structural models of the ligand-receptor complexes, but for a large series of compounds (as the whole series of 76 butyrophenones shown in table I) the visual inspection of the complex structures is not helpful.

Instead, 3D-QSAR methodologies can use these receptor-docked ligand structures to obtain models that extract from them the most relevant information. This use of ligand-receptor complexes as a starting point for 3D-QSAR has been previously reported (Waller C. L. and Marshall G. R. 1993; Waller C. L., Oprea T. I. et al. 1993; Cho S. J., Garsia M. L. et al. 1996; Pastor M 1997; Lozano J. J., Pastor M. et al. 2000; Sippl W. 2000; Sippl W., Contreras J. M. et al. 2001; Sippl W. 2002). One of the main limitations of 3D-QSAR methods is the difficulty of obtaining suitable ligand alignments. In this study, as ligand-receptor complexes has been built and validated, the ligands can be docked into the receptor binding site, and the resulting conformations can be used as aligned structures for the 3D-QSAR analysis.

Unfortunately, due to features of docking reminded at the beginning of the chapter (double orientation of binding, chirality, bioactive conformation) these efforts led to obtaining not one, but a set of candidate structures for every compound, from which it was not obvious how to pick a single representative structure. Most of the criteria that are used in these circumstances to make a choice seemed arbitrary or dangerous, thus, the 3D-QSAR models were built using a balanced set of structures representing the different species that bind the receptor.

The procedure is unusual but not entirely new. A similar approach has been reported in the past in a different context (Pastor M., Perez C. et al. 1997) with good results. Moreover, the peculiar characteristics of the series studied offers an excellent opportunity to validate the usefulness of the proposed approach by interpreting the results using the structure of the receptor and by comparing the results with those obtained using more conservative approaches (Ravina E., Negreira J. et al. 1999).

As for the previous studies with the 4 butyrophenones, the docking of ketanserin optimized through energy minimization and MD simulations was used as a starting point for the docking of the 76 butyrophenones. The affinities of these 76 compounds for the 5-HT_{2A} receptor have been determined experimentally and span from inactive compounds (with a pKi < 5) to a pKi > 9, with a good variability in pKi values (as shown in the article, in the Annex section).

Both enantiomers R and S of every compound of the series were docked into the modelled receptor binding site

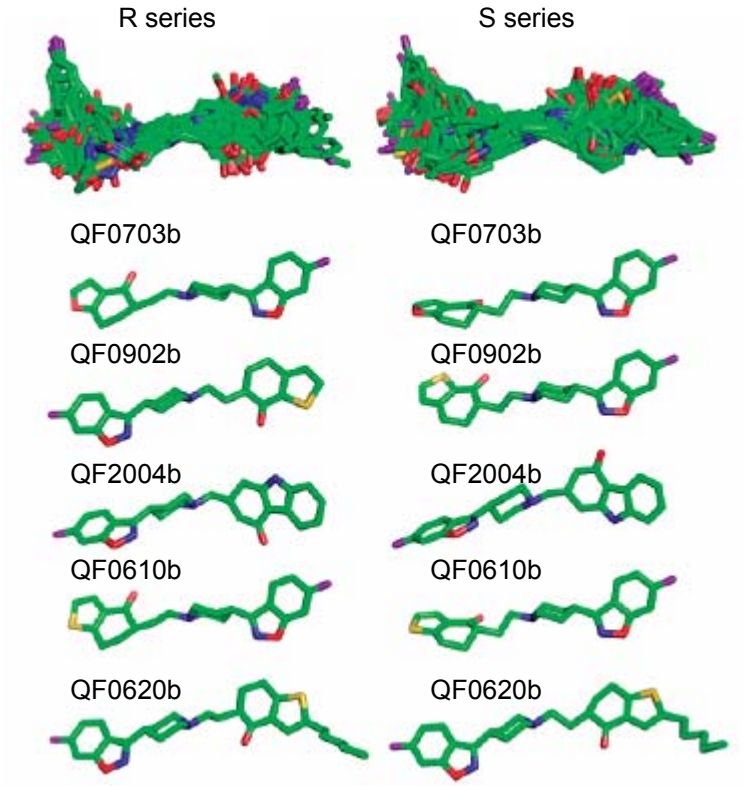
using program GOLD 2.2.49. At the end, 152 different structures were obtained and submitted to the docking protocol (see the article in the Annex for details).

The method was able to produce docked structures with a fitting score (GOLDscore) above a reasonable prefixed cut-off (35) for most of the compounds in the series (142 out of 152, 93%). Figure 93 shows, superimposed, the docked structures for the R and for the S series. In the figure, the observation of both series clearly shows both enantiomers bind in approximately the same position, but some orientations are more frequent in one series than in the other.

For example, in the R series, many compounds place the carbonyl group on the lower right side, while in the S series, the fluorine substituent of the p-fluorobenzoyl moiety is much more frequent on the upper right side. In the figure is evident that in the case of compound QF0703b, both enantiomers are inserted in the same orientation, while for QF0902b, the R and S enantiomers are inserted in opposite directions. Notice that the only difference between these two compounds is the substitution of the oxygen by a sulphur in the heteroaromatic ring of the alkanone moiety.

In compound QF2004b, both enantiomers orient the acceptor and donor groups in opposite directions. Both enantiomers of compound QF0610b are inserted in the same direction, but the incorporation of a pentane substituent to produce compound QF0620b induces a change to the alternative orientation, with no relevant change in the binding affinity.

Figure 4.4.93. From top to bottom: superimposition of all the docking solutions for the R enantiomers (left) and the S enantiomers (right); side by side view of the docking solution for the R (left) and S (right) enantiomers obtained for compounds QF0703b, QF0902b, QF2004b, QF0610b and QF0620b.



The analysis of the docking results confirmed the results of the MD simulations study reported before, that most structures can bind in two alternative orientations, one with the alkanone group oriented toward the inner part of the binding pocket and another orienting this group in the opposite direction. In both orientations, the charged nitrogen can bind the D3.32 and the polar groups can make approximately equivalent contacts.

The preference of the structures for binding in a certain orientation follows no obvious rules. In some structures (like R-QF0148v, R-QF0601b, or S-QF0602b, (see Table 1 and Chart 1 and 2, in the paper in the Annexes, for 2D structure of the compounds) both orientations were present among the three best solutions provided by GOLD, thus showing no strong preference for any orientation.

Other compounds seem to bind only in one orientation, but the introduction of a small structural change induces a preference for the other alternative. This is the case for R-QF0703b and R-QF0902b, which were inserted in opposite directions even if their structures differ only in the substitution of the furan ring by a thiophene (Figure 4.4.93). For some compounds, like QF0902b, the R- and S-enantiomers show a preference for binding in opposite orientations (Figure 4.4.93). A particularly interesting case is compounds QF0610b and QF0620b.

In the docking simulation, both QF0610b enantiomers are found to bind with the alkanone oriented toward the inner part of the binding pocket. Compound QF0620b, a derivative of QF0610b prepared ad hoc to prevent the original binding mode (by adding a bulky pentane substituent to the alkanone), was observed to dock in the opposite orientation (Figure 4.4.93).

The experimentally measured binding affinity of QF0620b was only slightly lower than the binding affinity of its parent (8.56 for QF0610b and 7.68 for QF0620b), thus suggesting that for compound QF0610b both binding orientations would also be accessible and contribute to the

experimental binding affinity, something that probably applies to most of the compounds in this series.

The output structures obtained from the docking are an excellent material for building a 3D-QSAR model, because they provide a more realistic representation of the ligand bioactive conformation than any alignment obtained by structural superimposition. When proceeding this way it must be remembered that the spatial framework for the interpretation is the structure of the receptor binding site, because this structure was used to place the ligands within a consistent reference frame.

In the present study, the biological activity values were obtained from binding affinity measurements carried out (in the vast majority of cases) using the compounds racemic mixture, thus making available only one affinity value for the R and S enantiomers present in the mixture. From a biological point of view, both enantiomers are not equivalent and their binding affinity might also be different. Indeed, in the few examples in which the activity has been measured separately for the R and S enantiomers, the values obtained were slightly different (see Table 1 in the article).

Obviously, the lack of separate affinity values for both enantiomers is a problem for the QSAR modelling. The solution proposed here is to include both enantiomers in the model as separate structures, but associate them to the single binding affinity value measured for the racemic mixture. By proceeding so, the binding affinity used corresponds to neither of the individual enantiomers and can be considered an average of both enantiomers true values.

Accordingly, the first 3D-QSAR model (M1) was built including both the R and S enantiomers for every compound, and the GRID/GOLPE methodology was used. Molecular interaction fields (MIF) were calculated with three probes, DRY, O, and N1, representing, respectively, a hydrophobic, a hydrogen bond acceptor, and a hydrogen bond donor group. These MIF were imported into the program GOLPE together

with the affinity values. The variables were pretreated and scaled before using the partial least-squares (PLS) regression analysis and the SRD/FFD variable selection method.

At the end, the PLS model obtained with three latent variables (LV) showed good statistical quality (r^2 = 0.90, q^2_{LOO} = 0.78, see Table VIII and Figure 4.4.94).

Figure 4.4.94. Scatterplots representing the experimental versus calculated pKi values for the models M1 (left) and M2 (right).

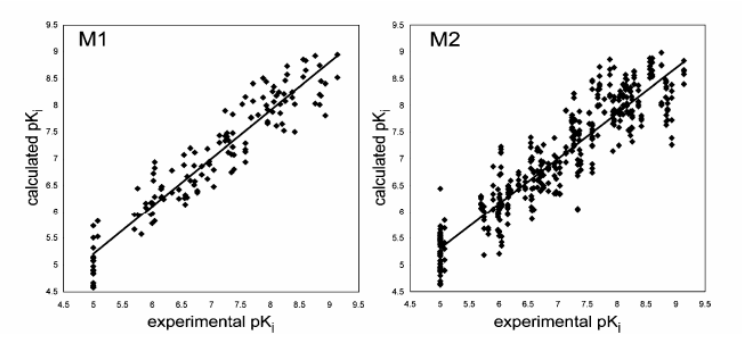


Table VIII. Statistical Parameters of the 3D-QSAR Models Obtained

model	LV	objects	Var.	r^2	q^2_{LOO}	$SDEP_{LOO}$	q^2_{5RG}	q^2_{2RG}
M1	3	142	1441	0.90	0.78	0.55	0.77	0.73
M2	3	426	1579	0.84	0.81	0.52	0.79	0.79

To further validate the model and to estimate its true predictive ability, we have split the series into a training and a

test set. The objects assigned to the test set (15 structures, approximately 10% of the whole set) were selected applying the Most Descriptive Compounds algorithm, as it was implemented in the GOLPE software on the space of the first three principal components extracted from the Principal Component Analysis of the whole series.

Then the whole 3D-QSAR modelling procedure was carried out as described above (PLS modelling and SRD/FFD variables selection) using only the compounds of the training set, while the compounds in the test set were used only for the predictions. In this case, the Standard Deviation of Error of Prediction (SDEP) obtained for the test set was of 0.53, and the external r^2 between experimental and predicted values was of 0.80. These values are perfectly comparable with the values obtained by standard cross-validation methods and further confirm the good predictive ability of M1.

Even if the model described above seems suitable for our purposes we decided to go one step further and apply to the problem of the multiple binding modes the same solution applied to address the presence of two enantiomers. In the same way that two enantiomers were used to characterize each compound, it is possible to use more than one docking solution to characterize each enantiomer.

For those compounds in which the three docking solutions are rather similar the use of multiple structures will have very little effect (the median of the RMSD was 0.62 for the solutions 1 and 2, 0.87 for the solutions 1-3, and 0.9 for the solutions 2 and 3), but in some cases, these solutions represent alternative binding modes. In our opinion, the choice of the first solution to represent the docking structure was arbitrary in these cases, much in the same sense that the choice of a certain enantiomer would be arbitrary, because the limited accuracy of the scoring functions question the reliability of the solution ranking provided by the docking programs.

Moreover, most compounds probably interact with the receptors in both orientations and in proportions that are

difficult to estimate. Therefore, the contemporary use of all these solutions will provide a much more realistic picture of their true binding ability. As in the previous case, we must be ready to accept a small decrease in the model fitting, lower r^2 values, and a less detailed description of fine effects, but hopefully, the models will be more robust and have better predictive ability, as it has been reported in other similar situations.

A second 3D-QSAR model (M2) was built using a similar methodology, but in this case, including the first three solutions provided by GOLD to characterize every enantiomer (six structures per compound) amounted to a total of 426 structures. At the end of the modelling procedure (see the paper for details), we obtained a PLS model with three LV of remarkable quality (r^2 = 0.84 and a q^2_{LOO} = 0.81, Table VIII, Figure 94). From a statistical point of view, the r^2 of M2 is lower than M1, as can be expected for the aforementioned reasons, but only slightly (from 0.90 to 0.84), and the difference is small if we consider that the number of objects included in the model has been increased from 142 to 426.

With respect to the values of q^2, they should be interpreted with care. In this particular series, the application of leave-one-out (LOO) method can produce overoptimistic results, because the removal of a single structure will not completely remove the presence of the compound from the reduced models. Therefore, it is not surprising to obtain a higher q^2_{LOO} of 0.81 for M2 than the 0.78 value obtained for M1.

A better estimation of the true predictive ability of M1 and M2 can be obtained using stricter crossvalidation utilizing either two or five randomly assigned groups (see Experimental Section in the paper). The results agree with the previous results, indicating that both models have a rather high predictive ability and that the incorporation of multiple structures does not decrease the predictive ability of the original model.

To further confirm the predictive ability of M2 and to compare it with M1, an external validation protocol identical to the one described for M1 was carried out. In this case, the test set included 45 compounds, corresponding to the three best docking solutions obtained for the 15 structures of the M1 test set.

As in the previous case, the structures in the test set were left out of the analysis, which was repeated from the beginning using only the compounds in the training set. To obtain comparable results with M1, the predictions obtained for the three docking solutions used to characterize each compound in the test set were averaged to provide a single estimate. The results of such predictions were good (SDEP = 0.52, r^2 = 0.81) and slightly better than those obtained with M1, thus confirming the improved predictive ability of M2 over M1, as suspected.

The comparison of the experimental versus calculated plots for M1 and M2 represented in Figure 94 shows that in M2 the objects exhibit a larger spread, and as it was expected, not all structures fit well into the model. The coefficient plots obtained for M1 and M2 look rather similar and contain essentially the same information. The main difference observed is that the plots for M2 are “cleaner”, devoid of the small regions produced by the effect of single compounds that usually have no general relevance.

All these results suggest that the model M2, obtained with multiple structures, can be a better alternative than M1, but the similarity observed between their coefficient plots indicate that the results of the M1 and M2 interpretation would be qualitatively similar.

Before beginning the interpretation of the M2 coefficient plots, we must remind the reader that the MIF assign negative energy values to favourable probe-ligand interactions and positive values to unfavourable (repulsive) probe-ligand interactions. Positive field values represent mainly the molecular shape, while negative values represent

regions where the ligand can make energetically favourable interactions with the binding site.

The PLS coefficient plots shown in Figure 95 contain both positive and negative values for the three probes used. Positive (yellow) coefficients might correspond to binding site regions where the presence of negative fields have an inverse correlation with activity or to regions where the presence of the molecular bulk (positive field) correlates directly with the activity. Negative (cyan) coefficients might represent regions where the presence of negative fields correlate directly with the activity or regions where the ligand positive fields correlate inversely with the activity.

These considerations add complexity to the interpretation and require a careful analysis of the coefficients, the field values produced for different compounds, and the activity contribution plots. It should also be stressed that the coefficient positions will make reference to the binding site residues (even if the binding site has not been used directly for obtaining the model), because the ligand alignment was based on the docking solutions.

Therefore, the reference framework of the whole analysis is the structure of the binding site instead of the superimposed ligand structures.

For the DRY Probe (Figure 4.4.95, Top), the highest negative coefficients were located in two areas, one at the bottom of the pocket near L163 and the other in the centre near the side chains of V7.39 and F6.51. The values of the coefficients indicate that the presence of hydrophobic groups in these positions increases the binding affinity, which is consistent with the hydrophobic nature of the neighbour residues. The main positive region obtained for this probe is located near the main chain of W6.48, indicating that the presence in this region of hydrophobic groups is detrimental for the binding. A similar interpretation can be given to a small positive region located near to the indolic nitrogen of W3.28 or to the polar atoms of the main chain near to the Gly160.

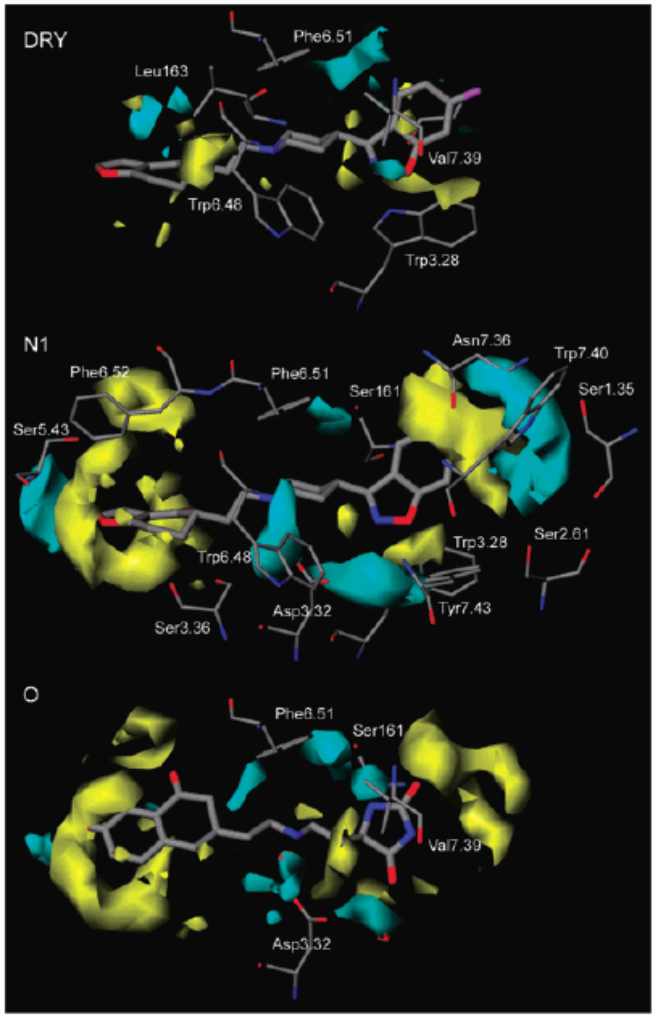
For N1 Probe (Figure 4.4.95, middle), most negative regions overlap some of the residues able to act as hydrogen bond donors. This is the case of a large region in the centre that extends between the oxygen atoms of D3.32 and the indolic nitrogen of W3.28. Also in the centre, there is a smaller region near the side chain of S161. Apart from these, there are two large negative regions in both extremes of the pocket.

One is located at the bottom near the S5.43 and presumably indicates that ligands able to make a favourable interaction with this residue have higher binding constants. On the other side of the receptor there is a much wider area enclosing the side chain of N7.36 and close to other polar side chains (S1.35, S2.61, W7.40).

Even if the regions do not overlap these residues, we must remember that the structure represented here is only a static snapshot of the receptor, and the above-mentioned residues could move their side chains to establish interactions at these positions. This can also be the case for an elongated region in the centre of the pocket near to the S3.36 and W6.48. In the conformation shown in Figure 94, the polar part of the W6.48 side chain could not overlap this region, but this residue forms part of the toggle switch that has been suggested to participate in a concerted conformational change together with F6.51 and F6.52.

During this conformational change the W6.48 rotates, orienting the polar nitrogen toward the ligand and locating this atom precisely in the region highlighted by the negative coefficients. With respect to the positive coefficients, most of them describe binding site regions that are overlapped by compounds with high binding affinity. The two main positive regions are located on the extremes of the binding site, indicating that the compounds with the highest affinity are long and fill the whole binding pocket, while shorter compounds usually show low affinity.

Figure 4.4.95. Coefficient plots obtained in model M2 for different blocks of variables representing the probes DRY, N1, and O. Positive coefficient values are represented in yellow and negative coefficient values are represented in cyan. Every coefficient plot included the structure of a representative ligand (QF0703b for DRY and N1 and QF0150v for O), as well as the structures of the more relevant residues.



Regarding the O Probe (Figure 4.4.95, bottom), the negative regions observed for this probe are smaller and less defined than those observed for N1, probably because the series contains much fewer hydrogen bond donor groups. The most important regions are located near the oxygen atoms of D3.32 and have a straightforward interpretation.

The negative regions located on the central part of the binding pocket, near the hydrophobic side chains of S161, F6.51, and V7.39, seem to represent the interactions of some compounds that contain a spiranic hydantoin ring (Chart 2 in the paper), like QF0150v) shown in Figure 4.4.95. In these compounds, one of the hydrogen bond donor atoms of the hydantoin ring generates two regions for potential hydrogen bond interactions, one near the side chain of S161 and the other oriented toward the V7.39.

The positive regions are rather similar to the positive regions observed for the N1 probe and can be interpreted in a similar way.

To summarize, the model stresses the importance of some of the polar interaction already recognized in the description of the ketanserin complex, like those with D3.32, W3.28, S161, and S 3.36, but adding some original information:

- highlights some biologically relevant interactions with hydrogen bond donor residues in both extremes of the pocket (S5.43 on one side, S1.35 and/or N7.36 on the other side) and a putative interaction with the indolic nitrogen of W6.48 when its side chain adopts an alternative conformation

- shows the importance of the hydrophobic interactions with residues located at the bottom of the pocket (L163) as well as in the centre (F6.51, V7.39)

- stresses the importance of the ligand size, indicating that ligands with high affinity tend to completely fill the cavity

- indicates that the few compounds in the series with hydrogen bond donor groups located in the alkanone moiety can make interactions that lead to an increase in the binding affinity, probably by interacting with S161.

5. Discussion

Reliable and robust models of the 5-HT_{2A} and 5-HT_{2C} receptors were built and used to dock several structures representing well known as well as novel compounds with binding affinity for the receptors. The ligand – receptor complexes were simulated using both docking and MD simulations.

The analysis of docking and MD simulations results and their correlation with the pharmacology of the ligands is not trivial, due to the large amount of data generated. Nevertheless, the results shown interesting trends and insights. Also, they could be the base of further studies, using computational or other experimental techniques.

One of the most interesting findings is the putative location of the binding site, the binding orientation of ligands and the identification of residues essential for binding. In our study, the binding site is located between helixes TMHIII, TMHV and TMHVI, slightly tilted with respect to the axis of the helices bundle, in a cavity that span from the extracellular loop EL2 to the aromatic residues W6.48, F6.51, F6.52.

The characteristics of the binding site support the design and discovery of ligands with a polar nitrogen (positively charged) located between two aromatic rings with polar groups: in fact, the negative side chain of D3.32 separates the

two aromatic boundaries of the active site. Both the top and the bottom of the binding site are delimited by aromatic residues: W3.28 and Y7.43 at the top and W6.48, F6.51, F6.52 and F5.47, at the bottom.

It is also worth to remind that the majority of the aromatic residues involved in the stabilization of the dockings are highly conserved throughout GPCR Class A family. Their role could therefore extend beyond the mere stabilization of the ligand-receptor complexes and could be related with macromolecular conformational changes involving the helixes.

In fact, one remarkable observation during the MD simulation was the behaviour of aromatic residues in presence of the different ligands. In particular, the movement of F6.51, F6.52, F5.47 and of other phenilalanines, not directly involved in binding were co-ordinated by the formation of π-π interactions.

During the MD simulations, the network of aromatic interactions seemed to evolve similarly to a zip that opens and closes, behaviour that should be confirmed with other studies and further investigated by computational methods without the restrictions applied in the current simulations.

During the MD simulations carried out with 5-HT_{2A} and 5-HT_{2C} in complex with QF0703b and QF1004b this particular behaviour of phenilalanines surrounding the active site and belonging to the aromatic cluster has been observed. Obviously, this observation could be influenced by the restriction applied to protein C-alpha trace, and therefore making hypothesis regarding the nature and the possible role of this movement could be biased by the limitations applied during the simulation. But it is not possible to exclude that these conformational changes in the aromatic side chains were giving some clues about the receptor conformational changes during the activation / inhibition mechanism, or with other supramolecular mechanisms.

When interacting with butyrophenones, these phenilalanines appear to generate a wave from the binding site

to the surface of TMHIV and TMHVI, by establishing and loosing π-π interactions. To which extent this behaviour could be related to macromolecular conformational changes, still remains a question to be answered.

Other interesting findings are related with binding orientation and chirality of ligands: apparently, both A and B orientations are possible and this duality of insertion should be considered for the design of novel compounds. In addition, both R and S enantiomers are suitable for binding 5-HT_{2A} and 5-HT_{2C} receptor, and chirality seems to influence binding affinity by conditioning the accommodation of ligands within the active site, resulting in a different network of interactions between ligands and receptors. All these issues introduce complexity in the search for novel compounds, making the drug discovery for these targets even more challenging.

In order to reduce such complexity, and attempt was made to define which bioactive orientation the compounds preferred for binding, by synthesizing the new ligand QF0620b (Figure 4.2.33) as a tool to force one orientation with respect to the other. Interestingly, QF0620b not only was found active and selective towards the 5-HT_{2A} receptor, but also its pharmacological profile was similar to QF0601b (K_i for 5-HT_{2A} = 7,68; K_i for 5-HT_{2C} = 6,95), when measured in competitive binding experiments with $[H]^3$ - ketanserin. This result supported the double binding orientation at least for compound QF0601b.

Moreover, the analysis of the docked structures and of the MD simulations was not trivial, and not successful in uncovering relevant relationships between the ligand structures and their binding affinities. Then the application of 3D-QSAR methodologies was envisaged as a potential solution. One of these techniques is the GRID/PCA method (Pastor M., Cruciani G. et al. 2000) that allows the characterization and the comparison of large set of structures. GRID / PCA method allowed the identification of the residues most involved in the interaction and the rationalization of the

behaviour of the receptors in response of ligand binding, by identifying the residues whose conformation varied most during the MD simulation.

In our series, the simultaneous use of several receptor-docked structures representing the same compound produces an interesting model that compares favourably, from a statistical point of view, with models obtained using classical methods, while being presumably more robust and having more general validity. The model is interpretable in terms similar to those used in structure based drug design studies, allowing extracting information that can be easily translated as guidelines for the design of novel compounds.

Moreover, estimation of the models predictive ability based on strict cross-validation methods and in external datasets suggests that the predictions for structurally related compounds will fall within an interval of ±1.0. Even if this cannot be considered a very accurate estimation, it is more than enough for identifying potentially active or inactive compounds.

6. Conclusions

The present thesis aimed to contribute to the research of novel antipsychotic drugs. Structural models of the 5-HT2A and 5-HT2C receptors, alone and in complex with relevant ligands were obtained. MD simulations and 3D-QSAR were also carried out in order to extract the most relevant information for the discovery of new compounds.

The application of all these methods allowed drawing the following conclusions:

1. A novel protocol for building GPCRs has been applied and validated. The new protocol presents some interesting and original features that open new

possibilities to other researches who like to apply it to other GPCRs.

2. The novel protocol was used for obtaining structural models of the 5-HT2A and 5-HT2C receptors, alone and in complex with ketanserin and a series of butyrophenones.

3. In our models, the loops were found not to play a key role in the binding the studied ligands. However, in some of the obtained models, polar residues of EL2 such as Lys155, could extend the large polar side chains within the boundaries of the binding site and establish some interaction, but these interaction seems not to be preeminent in stabilizing the docking position.

4. A set of key residues for ligand receptor binding were identified. All of the key residues previously reported as essential for binding by mutagenesis studies were confirmed by our models. Some of them are involved in the activation process and conserved throughout GPCR Class A receptors. The docking and molecular dynamic simulations identified the following residues as the most important for the ligand receptor interaction: D3.32, S3.36, S5.45, Y7.43, W6.48, F6.51 and F6.52.

5. Selectivity for the 5-HT2A receptor. The results of the computational methods indicate that the complexes with 5-HT2A receptor were slightly more stable than the complexes with 5-HT2C receptor, in agreement with the selectivity found in experimental binding affinity assays.

6. Ketanserin and members of butyrophenone series can bind serotonin receptors in a double orientation. This double orientation of ligands can potentially be exploited for the design of novel compounds.

7. Chirality of ligands could play a preeminent role in ligand binding. This interesting result introduces difficulties in drug design of more effective ligands, as chirality should be taken into account. This result encourages collecting exhaustive binding and pharmacology data for the enantiomers of additional compounds.

8. The use of multiple conformations and orientations of ligands allowed the development of a multi-structure 3D-QSAR model with a high predictive power (q^2_{LOO} = 0.78 using three latent variables)

9. The resulting multi-structure 3D-QSAR model confirms the importance of some of the polar interactions previously recognized in the docking studies, such as those with D3.32, W3.28, S161, and S3.36, as well as some other not previously recognized interactions, which can be exploited for the design of more active compounds.

7. Bibliography

Aghajanian G.K. and Marek G.J. (2000). Brain Res Rev **31**: 302-312.

Agrafiotis D.K. (2002). "Combinatorial informatics in the post-genomics era." Nat Rev Drug Discov **1**: 337-346.

Alewijnse A.E., Timmerman H., et al. (2000). "The effect of mutations in the DRY motif on the constitutive activity and structural instability of the histamine H2 receptor." Mol Pharmacol **57**(5): 890-898.

Allen F. H. (2002). "The Cambridge Structural Database: a quarter of a million crystal structures and rising." Acta Crystallogr B **58**(Pt 3 Pt 1): 380-8.

Almaula N., Ebersole B.J., et al. (1996). "Contribution of a helix 5 locus to selectivity of hallucinogenic and nonhallucinogenic ligands for the human 5-hydroxytriptamine2A and 5-hydroxytriptamine2C receptors: direct and indirect effects on ligand affinity mediated by the same locus." Mol Pharmacol **50**: 32-42.

Almaula N., Ebersole B.J., et al. (1996). "Mapping the binding site pocket of the serotonin 5-hydroxytryptamine2A receptor. Ser3.36(159) provides a second interaction site for the protonated amine of serotonin but not of lysergic acid diethylamide of bufotenin." J Biol Chem **271**: 14672-14675.

Altenbach C., Cai K., et al. (2001). "Structure and function in rhodopsin: mapping light-dependent changes in distance between residue 65 in helix TMI and residues in the sequence 306-319 at the citoplasmic end of TM7 and in helix H8." Biochemistry **40**(51): 15483-15492.

Altenbach C., Klein-Seetharaman J., et al. (2001). "Structure and function in rhodopsin: mapping light-dependent changes in distance between residue 316 in helix 8 and residues in the sequence 60-75, covering the cytoplasmic end of helices TM1 and TM2 and their connecting loop CL1." Biochemistry **40**(51): 15493-15500.

Altenbach C., Klein-Seetharaman J., et al. (1999). "Structural features and light-dependent changes in the sequence 59-75 connecting helices I and II in rhodopsin: a site-directed spin-labeling study." Biochemistry **38**(25): 7945-7949.

Altenbach C., Yang K., et al. (1996). "Structural features and light-dependent changes in the cytoplasmatic interhelical E-F loop region of rhodopsin: a site-directed spin labeling study." Biochemistry **35**(38): 12470-12478.

Altschul S. F. and Koonin E. V. (1998). "Iterated profile searches with PSI-BLAST--a tool for discovery in protein databases." Trends Biochem Sci **23**(11): 444-7.

Altschul S. F., Madden T. L., et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res **25**(17): 3389-402.

Alvarado M., Coelho A., et al. (2005). "Synthesis and binding affinity of novel 3-aminoethyl-1-tetralones, potential atypical antipsychotics." Bioorg Med Chem Lett **15**(12): 3063-6.

Ashton M., Schwarz M.K., et al. (2002). "The selection and design of GPCR ligands: from concept to the clinic." Comb Chem High Thr Scree **7**: 441-452.

Attwood T.K. and Findlay J.B. (1994). "Fingerprinting G-protein-coupled receptors." Prot Eng **7**: 195-203.

Baldwin J.M. (1993). "The probable arrangement of the helices in G protein-coupled receptors." EMBO J **12**(4): 1693-1703.

Baldwin J.M. (1994). "Structure and function of receptors coupled to G proteins." Curr Opin Cell Biol **6**: 180-190.
Baldwin J.M., Schertler G.F.X., et al. (1997). "An alpha-carbon template for the transmembrane helices in the rhodopsin family of G-protein-coupled receptors." J Mol Biol **272**(1): 144-164.
Ballesteros J., Jensen A.D., et al. (2001). "Activation of the ß2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6." J Biol Chem **276**(31): 29171-29177.
Ballesteros J., Kitanovic S., et al. (1998). "Functional microdomain in G-protein-coupled receptors." J Biol Chem **273**(17): 10445-10453.
Ballesteros J. and Palczewski K. (2001). "G protein-coupled receptor drug discovery: implications from the crystal structure of rhodopsin." Curr Opin Drug Discov Devel **4**(5): 561-74.
Ballesteros J. and Weinstein H. (1994). Integrated methods for the construction of three dimensional models and computational probing of structure-function relations in G-protein coupled receptors. Methods in Neurosciences. S. S. C. Conn P.M. San Diego, Academic Press**:** 366-428.
Baroni M., Costantino G., et al. (1993). "Generating optimal linear PLS estimations (GOLPE): and advanced chemometric tool for handling 3D-QSAR problems." Quant. Struct. Act.-Relat. **12**: 9-20.
Befort K., Zilliox C., et al. (1999). "Constitutive activation of the d-opiod receptor by mutations in transmembrane domains III and VII." J Biol Chem **275**(26): 18574-18581.
Beukers M. W., Kristiansen I., et al. (1999). "TinyGRAP database: a bioinformatics tool to mine G-protein-coupled receptor mutant data." Trends Pharmacol Sci **20**(12): 475-7.

Bissantz C. (2003). "Conformational changes of G protein-coupled receptors during their activation by agonist binding." J Recept Signal Transduct Res.
Bissantz C., Folkers G., et al. (2000). "Protein-based virtual screening of chemical databases. 1. Evaluation of different docking / scoring combinations." J Med Chem **2000**(43): 4759-4767.
Bloom F.E. and Kupfer D.J. (1995). Psychopharmacology: the fourth generation of progress. New York, Raven Press**:** 407-471.
Blundell T.L., Sibanda B.L., et al. (1987). Nature **326**: 347.
Bockaert J. and Pin J.P. (1999). "Molecular tinkering of G protein-coupled receptors: an evolutionary success." EMBO J **18**: 1723-1729.
Boeckmann B., Bairoch A., et al. (2003). "The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003." Nucleic Acid Res **31**: 365.
Boehm H.J. and Klebe G. (1996). Angew Chem int Ed Engl **35**: 2588.
Boess F.G., Monsma F.J. Jr, et al. (1998). "Identification of residues in transmembrane regions III and VI that contribute to the ligand binding site of the serotonin 5-HT6 receptor." J Neurochem **71**: 2169-2177.
Boess F.G., Monsma F.J. Jr., et al. (1997). "Interaction of tryptamine and ergoline compounds with threonine 196 in the ligand binding site of the 5-hydroxytryptamine6 receptor." Mol Pharmacol **52**: 515-523.
Brea J., Castro M., et al. (2006). "QF2004B, a potential antipsychotic butyrophenone derivative with similar pharmacological properties to clozapine." Neuropharmacology **51**(2): 251-62.
Brea J., Masaguer C.F., et al. (2003). "Conformationally constrained butyrophenones as new pharmacological tools to study 5-HT_{2A} and 5-HT_{2C} receptor behaviours." Eur J Med Chem **38**: 433-440.

Brea J., Rodrigo J., et al. (2002). "New serotonin 5-HT_{2A}, 5-HT_{2B} and 5-HT_{2C} receptor antagonists: synthesis, pharmacology, 3D-QSAR, and molecular modeling of (aminoalkyl)benzo and heterocycloalkanones." J Med Chem **45**: 54-71.
Bronowska A., Chilmonczyk Z., et al. (2001). "Molecular dynamics of buspirone analogues interacting with the 5-HT1A and 5-HT2A serotonin receptors." Bioorg Med Chem **9**: 881-895.
Browne W.J., North A.C.T., et al. (1969). J Mol Biol **42**: 65.
Bruccoleri R.E. and Karplus M. (1987). Biopolymers **26**: 137.
Bryson K., McGuffin L. J., et al. (2005). "Protein structure prediction servers at University College London." Nucleic Acids Res **33**(Web Server issue): W36-8.
Campbell M., Young P.I., et al. (1999). "The use of atypical antipsychotics in the management of schizophrenia." Br J Pharmacol **47**: 13-22.
Carlsson A. (1988). "The current status of the dopamine hypotesis of schizophrenia." Neuropsychopharmacology **1**: 179-186.
Case D.A., Pearlman D.A., et al. (1999). AMBER6. University of California.
Chanda P.K., Minchin M.C., et al. (1993). "Identification of residues important for ligand binding to the human 5-hydroxytryptamine1A serotonin receptor." Mol Pharmacol **43**: 516-520.
Chothia C. and Lesk A.M. (1986). "The relation between the divergence of sequence and structure in proteins." EMBO J **5**: 823-836.
Choundary M.S., Craigo S., et al. (1993). "A single point mutation (Phe340-->Leu340) of a conserved phenylalanine abolishes 4-[125I]iodo-(2,5-dimethoxy) phenylisopropylamine and [3H]mesulergine but not [3H]ketanserin binding to 5-hydroxytryptamine2 receptors." Mol Pharmacol **43**: 755-761.

Choundary M.S., Sachs N., et al. (1995). "Differential ergoline and ergopeptine binding to 5-hydroxytryptamine2A receptors: ergolines require an aromatic residue at position 340 fo high affinity binding." Mol Pharmacol **47**: 450 - 457.
Cohen G.B., Yang T., et al. (1993). "Constitutive activation of opsin: influence of charge at position 134 and size at position 296." Biochemistry **32**(23): 61111-6115.
Congreve M. (2005). "Structural biology and drug discovery." Drug Discov Today **10**: 895-907.
Cramer R.D., Patterson D.E., et al. (1998). "Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins." J Am Chem Soc **110**: 5939-5967.
Cuff J. A., Clamp M. E., et al. (1998). "JPred: a consensus secondary structure prediction server." Bioinformatics **14**(10): 892-3.
Cuff J.A. and Barton G.J. (1999). "Evaluation and improvement of multiple sequence methods for protein secondary structure prediction." Proteins **34**: 508-519.
Cussac D., Newman T.A., et al. (2000). Naunyn Schmiedebergs Arch Pharmacol **361**: 549-554.
Dezi C., Brea J., et al. (2007). "Multistructure 3D-QSAR study on a series of conformationally constrained butyrophenones docked into a new homology model of the 5-HT2A receptor." J Med Chem **50**: 3242-3255.
Donnelly D., Maudsley S., et al. (1999). "Conserved polar residues in the transmembrane domain of the human tachykinin NK2 receptor: functional roles and structural implication." Biochem J **339(PT1)**: 55-61.
Ebersole B.J., Visiers I., et al. (2003). "Molecular basis of partial agonism: orientation of indoleamine ligands in the binding pocket of the human serotonin 5-HT2A receptor determines relative efficacy." Mol Pharmacol **63**(1): 36-43.

Eddy R.S. (1998). "Profile hidden Markov models." Bioinformatics **14**(9): 755-763.
Edvardsen O. and Kristiansen K. (1997). TM Journal **6**: 1-6.
Espadaler J., Fernandez-Fuentes N., et al. (2004). "ArchDB: automated protein loop classification as a tool for structural genomics." Nucleic Acids Res **32**(Database issue): D185-8.
Farahbakhsh Z.T., Ridge K.D., et al. (1995). "Mapping light-dependent strucural changes in the cytoplasmatic loop connecting helices C and D in rhodopsin: a site-directed spin labeling study." Biochemistry **34**(27): 8812-8819.
Farrens D.L., Altenbach C., et al. (1996). "Requirements of rigid-body motions of transmembrane helices for light activation of rhodopsin." Science **274**(5288): 768-770.
Fernandez-Fuentes N., Querol E., et al. (2005). "Prediction of the conformation and geometry of loops in globular proteins: testing ArchDB, a structural classification of loops." Proteins **60**(4): 746-757.
Fiorella D., Rabin R.A., et al. (1995). "The role of the 5-HT2A and 5-HT2C receptors in the stimulus effects of hallucinogenic drugs. I. Antagonist correlation analysis." Psychopharmacology **121**: 347-356.
Fiser A., Do R.K., et al. (2000). Prot Sci **9**: 1753.
Flower D.R. (1999). "Modeling G-protein-coupled receptors for drug design." Biochim Biophys Acta **1422**(3): 207-234.
Fredriksson R., Lagerstrom M.C., et al. (2003). "The G protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints." Mol Pharmacol **63**: 1256-1272.
Gaddum J.H and Hameed K.A. (1954). "Drugs which antagonize 5-hydroxytryptamine." Br J Pharmacol **9**: 240-248.

Gaddum J.H, Khan A., et al. (1955). "Quantitative studies of antagonists for 5-hydroxy-tryptamine." Q J Exp Physiol **40**: 49-74.
Gaddum J.H and Picarelli Z.P. (1957). "Two kinds of tryptamine receptors." Br J Pharmacol **12**: 323-328.
Gether U., Lin S., et al. (1997). "Agonists induce conformational changes in transmembrane domains III and VI of the ß2-adrenoceptor." EMBO J **16**(22): 6737-6747.
Gether U., Lin S., et al. (1995). "Fluorescent labeling of purified ß2-adrenergic receptor." J Biol Chem **270**(47): 28268-28275.
Ghanouni G., Steenhuis J.J., et al. (2001). "Agonist-induced conformational changes in the G-protein-coupling domain of the ß2 adrenergic receptor." Proc Natl Acad Sci USA **98**(11): 5997-6002.
Glennon R.A. (1990). "Do hallucinogens act as 5-HT2 agonists or antagonists?" Neuropsychopharmacology **56**: 509-517.
Goodford P.J. (1985). "A computational procedure for determining energetically favorable binding sites of biological important macromolecules." J Med Chem **28**: 849-857.
Govaerts C., Lefort A., et al. (2001). "A conserved Asn in transmembrane helix 7 is an on/off switch in the activation of the thyrotropin receptor." J Biol Chem **276**(25): 22991-9.
Granas C. and Larhammar D. (1999). "Identification of an amino acid residue important for binding of methiothepin and sumatriptan to the human 5-HT(1B) receptor." Eur J Pharmacol **380**: 171-181.
Granas C., Nordquist J., et al. (2001). "Site-directed mutagenesis of the 5-HT1B receptor increases the affinity of 5-HT for the agonist low-affinity conformation and reduces the intrinsic activity of 5-HT." Eur J Pharmacol **421**: 69-76.

Granas C., Nordvall G., et al. (1998). "Mutagenesis of the human 5-HT1B receptor: differences from the closely related 5-HT1A and the role of F331 in signal transduction." J Recept Signal Transduct Res **18**: 225-241.
Granas C., Nordvall G., et al. (1998). "Site - directed mutagenesis of the human 5-HT1B receptor." Eur J Pharmacol **349**: 367-375.
Greasley P.J., Fanelli F., et al. (2002). "Mutagenesis and modeling of the a1B-adrenergic receptor highlight the role of the helix 3/ helix 6 interface in receptor activation." Mol Pharmacol **61**(5): 1025-1032.
Green M.F., Marshll-BD J., et al. (1997). Am J Psychiatry **154**: 799-804.
Greer J. (1981). J Mol Biol **153**: 1027.
Groblewski T., Maigret B., et al. (1997). "Mutation of Asn111 in the third transmembrane domain of the AT1a angiotensin II receptor induces its constitutive activation." J Biol Chem **272**(3): 1822-1826.
Han M., Smith S.O., et al. (1998). "Constitutive activation of opsin by mutation of methionine 257 on transmembrane helix 6." Biochemistry **37**(22): 8253-8261.
Hansch C., Rockwell S.D., et al. (1977). "Substituent constants for correlation analysis." J Med Chem **20**(2): 304-306.
Havel T.F. and Snow M.E. (1991). J Mol Biol **217**: 1.
Henderson R., Baldwin J.M., et al. (1990). "Model for the structure of bacteriorhodopsin based on high-resolution electron cryomicroscopy." J Mol Biol **213**(4): 899-929.
Herrick-Davis K., Grinde E., et al. (2000). J Pharmacol Exp Ther **295**: 226-232.
Hiber M.F., Trumpp-Kallmeyer S., et al. (1991). "Three-dimensional models of neurotransmitter G-binding protein-coupled receptors." Mol Pharmacol **40**(1): 8-15.

Hippius H. (1989). "The history of clozapine." J Psychopharm **93**: S3-S5.
Hippius H. (1999). "A historical perspective of clozapine." J Clin Psychiatry **60**: 22-23.
Hirst W.D., Blaney F.E., et al. (2003). "Differences in the central nervous system distribution and pharmacology of the mouse 5-hydroxytryptamine-6 receptor compared with rat and human receptors investigated by radioligand binding, site-directed mutagenesis, and molecular modeling." Mol Pharmacol **64**: 1295-1308.
Ho B.Y., Karshin A., et al. (1992). "The role of conserved aspartate and serine residues in ligand binding and in function of the 5-HT1A receptor: a site-directed mutation study." FEBS Lett **312**: 259 - 262.
Holtje H.D. (1995). "Construction of a detailed serotonergic 5-HT2A receptor model." Arch Pharm (Weinheim) **328**(577-584).
Homan E.J. and Grol C.J. (1999). "Molecular modeling of the dopamine D2 and serotonin 5-HT1A receptor binding modes of the enantiomers of 5-OMe-BPAT." Bioorg Med Chem **7**: 1805-1820.
Hoyer D., Clarke D.E., et al. (1994). "International Union of Pharmacology classofication of receptors for 5-hydroxytriptamine (serotonin)." Pharmacol Rev **46**: 157-203.
Ikeda M., Iwata N., et al. (2006). "Positive association of the serotonin 5-HT7 receptor gene with schizophrenia in a Japanese population." Neuropsychopharmacology **31**(4): 866-71.
Ismaiel A.M., De Los Angeles J., et al. (1993). "Antagonism of a 1-(2,5-dimethoxy-4-methylphenyl)-2-aminopropane stimulus with a newly identified 5-HT2- vs. 5-HT1C-selective antagonist." J Med Chem **36**: 2519-2525.
Jacoby E., Gerspacher M., et al. (2006). "The 7TM G-Protein-Coupled receptor target family." ChemMedChem **1**(760-782).

Janin J. and Chothia C. (1978). Biochemistry **17**: 2943.
Jensen A.D., Guarnieri F., et al. (2001). "Agonist-induced conformational changes at the cytoplasmic side of transmembrane segment 6 in the ß2-adrenergic receptor mapped by site-selective fluorescent labeling." J Biol Chem **276**(12): 9279-9290.
Johnson M.P., Loncharich R.J., et al. (1994). "Species variation in transmembrane region V of the 5-hydroxytryptamine type 2A receptor alter the structure activity relationship of certain ergolines and tryptamines." Mol Pharmacol **45**: 277-286.
Johnson M.P., Wainscott D.B., et al. (1997). "Mutations of transmembrane IV and V serines indicate that all tryptamines do not bind to the rat 5-HT2A receptor in the same manner." Brain Res **49**: 1-6.
Jones D. T., Taylor W. R., et al. (1992). "A new approach to protein fold recognition." Nature **358**(6381): 86-9.
Jones G., Willet P., et al. (1997). "Development and validation of a genetic algorithm for flexible docking." J Mol Biol **267**: 727 - 748.
Kapur S. and Remington G. (1996). "Serotonin-dopamine interaction and its relevance to schizophrenia." Am J Psychiatry **153**(4): 466-76.
Kitchen D. B., Decornez H., et al. (2004). "Docking and scoring in virtual screening for drug discovery: methods and applications." Nat Rev Drug Discov **3**(11): 935-49.
Klabunde T. and Jager R. (2006). "Chemogenomics approaches to G-protein coupled receptor lead finding." Ernst Schering Res Found Workshop(58): 31-46.
Kolakowski L.F. Jr. (1994). "GCRDb: a G-protein-coupled receptors database." Recept Channels **2**: 1-7.
Kolinski A., Betancourt M.R., et al. (2001). Proteins **44**: 133.
Krishna A.G., Menon S.T., et al. (2002). "Evidence that helix 8 of rhodopsin acts as a membrane-dependent conformational switch." Biochemistry **41**(26): 8298-8309.
Kristiansen K. (2004). "Molecular mechanism of ligand binding, signaling and regulation within the superfamily of G-protein-coupled receptors: molecular modeling and mutagenesis approaches to receptor structure and function." Pharmacol & Ther **103**: 21-80.
Kristiansen K. and Dahl S.G. (1996). "Molecular modeling of serotonin, ketanserin, ritanserin and their 5-HT receptor interactions." Eur J Pharmacol.
Kristiansen K., Dahl S.G., et al. (1996). Proteins: Struct. funct. genetic. **26**: 81-94.
Kristiansen K., Kroeze W.K., et al. (2000). "A highly conserved aspartic acid (Asp-155) anchors the terminal amine moiety of tryptamines and is involved in membrane targeting of the 5-HT(2A) serotonin receptor but does not participate in activation via a "salt-bridge" disruption mechanism." J Pharmacol Exp Ther **293**: 735-746.
Kroeze W.K., Kristiansen K., et al. (2002). "Molecular biology of serotonin receptors - structure and function at the molecular level." Curr Top in Med Chem **2**: 507-528.
Kubinyi H. (1993). 3D QSAR in Drug Design. Theory, Methods and Applications. Leiden, ESCOM, Science Publishers B.V.
Kubinyi H. and Kehrhahn O.H. (1976). "Quantitative structure-activity relationships. 3.1 A comparison of different Free-Wilson mode." J Med Chem **19**(8): 1040.
Kuntz I.D., Blaney J.M., et al. (1982). J Mol Biol **161**: 269.
Kuntz I.D., Meng E.C., et al. (1994). Acc Chem Res **27**: 117.
Kursar J.D., Nelson D.L., et al. (1994). "Molecular cloning, functional expression, and mRNA tissue distribution of the human 5-hydroxytryptamine$_{2B}$ receptor." Mol Pharmacol **46**: 227-234.

Laskowski R.A., MacArthur M.W., et al. (1993). "PROCHECK: a program to check the stereochemical quality of protein structures." J Appl Cryst **26**: 283-291.
López-Rodríguez M.L., Benhamú B., et al. (2001). "Computational model of the complex between GR113808 and the 5-HT4 receptor guided by site-directed mutagenesis and the crystal structure of rhodopsin." J Comp Aid Mol Des **15**: 1025-1033.
López-Rodríguez M.L., de la Fuente T., et al. (2005). "A three-dimensional pharmacophore model for 5-hydroxytryptamine6 (5-HT6) receptor antagonists." J Med Chem **48**: 4216-4219.
López-Rodríguez M.L., Deupi X., et al. (2002). "Design, synthesis and pharmacological evaluation of 5-hydroxytryptamine1a receptor ligands to explore the three-dimensional structure of the receptor." Mol Pharmacol **62**: 15-21.
López-Rodríguez M.L., Morcillo M.J., et al. (2005). "5-HT4 receptor antagonists: structure-affinity relationships and ligand-receptor interactions." Curr Top in Med Chem **2**: 625-641.
López-Rodríguez M.L., Morcillo M.J., et al. (2004). "Serotonin 5-HT7 receptor antagonists." Curr Med Chem **4**(3).
López-Rodríguez M.L., Murcia M., et al. (2001). "3-D-QSAR/CoMFA and recognition models of benzimidazole derivatives at the 5-HT4 receptor." Bioorg Med Chem Lett **11**: 2807-2811.
López-Rodríguez M.L., Murcia M., et al. (2002). "Benzimidazole derivatives. 3. 3D-QSAR/CoMFA model and computational simulation for the recognition of 5-HT4 receptor antagonists." J Med Chem **45**: 4806-4815.
López-Rodríguez M.L., Murcia M., et al. (2003). "Design and synthesis of S-(-)-2-[[4-napht-1yl)piperazin-1-yl]-methyl]-1,4-dioxoperhydropyrrolo[1,2-a]pyrazine (CSP-2503) using computational simulation. A 5-HT1A receptor agonist." Bioorg Med Chem Lett **13**: 1429-1432.
López-Rodríguez M.L., Murcia M., et al. (2001). "Synthesis and structure-activity relationship of a new model of arylpiperazines. 6. Study of the 5-HT1A/a1-adrenergic receptor affinity by classical hansch analysis, artificial neuronal networks, and computational simulation of ligand recognition." J Med Chem **44**: 198-207.
López-Rodríguez M.L., Porras E., et al. (2003). "Optimization of the pharmacophore model for 5-HT7R antagonism. Design and synthesis of new naphtholactam and naphthosultam derivatives." J Med Chem **46**: 5638-5650.
Lowrie J.F., Hobbs D.W., et al. (2004). "The different strategies for designing GPCR and kinase targeted libraries." Comb Chem High Thr Scree **7**: 495-510.
Lozano J.J., Pastor M., et al. (2000). "3D-QSAR methods on the basis of ligand-receptor complexes. Application of COMBINE and GRID / GOLPE methodologies to a series of CYP1A2 ligands." J Comp Aid Mol Des **14**: 341 -353.
Manivet P., Schneider B., et al. (2002). "The serotonin binding site of human and murine 5-HT2b receptors: molecular modeling and site-directed mutagenesis." J Biol Chem.
Marti-Renom M.A., Stuart A., et al. (2000). Annu Rev Biophys Biomol Struct **29**: 291.
Massotte D. and Kieffer B.L. (2005). "The second extracellular loop: a damper for g protein-coupled receptors?" Nat Struct Mol Biol **12**: 287-288.
McGuffin L. J., Bryson K., et al. (2000). "The PSIPRED protein structure prediction server." Bioinformatics **16**(4): 404-5.
McKenna D.J., Repke D.B., et al. (1990). "Differential interactions of indolealkylamines with 5-hydroxy-tryptamine receptors subtypes." Neuropharmacology **29**: 193-198.

McMartin C. and Bohacek R.S. (1997). "QXP: Powerful, rapid computer algorithms for structure-based drug design." J Comp Aid Mol Des **11**: 333-344.
Meltzer H.Y., Matsubara S., et al. (1989). "Classification of typical and atypical antipsychotics drugs on the basis of dopamine D-1, D-2 and serotonin2 pKi values." J Pharmacol Exp Ther **251**: 238-246.
Mialet J., Berque-Bestel I., et al. (2000). "Isolation of the serotoninergic 5-HT4(e) receptorfrom human heart and comparative analysis of its pharmacological profile in C6-glial and CHO cell lines." Br J Pharmacol **129**: 771-781.
Mialet J., Dahmoune Y., et al. (2000). "Exploration of the ligand binding site of the human 5-HT(4) receptor by site-directed mutagenesis and molecular modeling." Br J Pharmacol **130**: 527-538.
Morris G.M., Goodsell D.S., et al. (1998). "Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function." Journal of Computational Chemistry **19**(14): 1639-1662.
Moult J. (2005). "A decade of CASP: progress, bottlenecks and prognosis in protein structure prediction." Curr Opin Struct Biol **15**: 285-289.
Moult J. and James M.N. (1986). Proteins **1**: 146.
Nelson D.L., Glennon R.A., et al. (1994). Comparison of the affinities of hallucinogenic phenylalkylamines at the cloned human 5-HT2A, 2B, and 2C receptors. Third IUPHAR Satellite Meeting on Serotonin, Chicago, IL.
Nowak M. K., Pawlowski M., et al. (2006). "Homology modeling of the serotonin 5-HT1A receptor using automated docking of bioactive compounds with defined geometry." J Med Chem **49**: 205-214.
Okada T., Fujiyoshi Y., et al. (2002). "Functional role of internal water in rhodpsin revealed by x-ray cristallography." Proc Natl Acad Sci USA **99**(9): 5982-5987.

Okada T., Sugihara M., et al. (2004). "The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure." J Mol Biol **342**(2): 571-83.
Oksenberg D., Marsters S.A., et al. (1992). "A single amino-acid difference confers major pharmacological variation between human and rodent 5-HT1B receptors." Nature **360**: 161 - 163.
Ooms F . (2000). "Molecular modeling and computer aided drug design. Examples of their application in medicinal chemistry." Curr Med Chem **7**: 141-158.
Orry J.W.A. and Cavasotto C.N. (2006). "Structure-based development of target-specific compound libraries." Drug Discov Today **11**(5/6): 261-266.
Palczewski K., Kumasaka T., et al. (2000). "Crystal structure of rhodopsin: a G protein-coupled receptor." Science **289**(5480): 739-745.
Parker E.M., Grisel D.A., et al. (1993). "A single amino acid difference accounts for the pharmacological distinctions between the rat and human 5-hydroxytrypatmine1B receptors." J Neurochem **60**: 380-383.
Pastor M., Cruciani G., et al. (1997). "Smart region definition: a new way to improve the predictive ability and interpretability of three-dimensional quantitative structure-activity relationship." J Med Chem **40**: 1455-1464.
Pastor M., Cruciani G., et al. (2000). "Grid-Independent descriptors (GRIND): a novel class of alignment-independent three-dimensional molecular descriptors." J Med Chem **43**: 3233-3243.
Pastor M., Perez C., et al. (1997). "Simulation of alternative binding modes in a structure-based HIV-1 protease inhibitors." J Mol Graph Model **15**: 364-371, 389.
Pearson W. R. and Lipman D. J. (1988). "Improved tools for biological sequence comparison." Proc Natl Acad Sci U S A **85**(8): 2444-8.

Perlman J.H., Colson A.O., et al. (1997). "Interaction between conserved residues in transmembrane helices 1,2 and 7 of the thyrotropin-releasing hormone receptor." J Biol Chem **272**(18): 11937-11942.
Peroutka S.J. and S. S.H. (1979). "Multiple serotonin receptors: differential binding of [3H]5-hydroxitriptamine, [3H]lysergic acid diethylamide and [3H]spiroperidol." Mol Pharmacol **16**: 687-699.
Persson B. and Argos P. (1997). "Prediction of membrane protein topology utilizing multiple sequence alignments." J Protein Chem **16**(5): 453-7.
Pierce K.L., Premont R.T., et al. (2002). "Seven transmembrane receptors." Nat Rev Mol Cell Biol **3**: 639-650.
Porter J.E., Hwa J., et al. (1996). "Activation of the a1B-adrenergic receptor is initiated by disruption of an interhelical salt bridge constraint." J Biol Chem **271**(45): 28318-28323.
Rashid M., Nishio H., et al. (2003). "Identification of the binding sites and selectivity of sarpogrelate, a novel 5-HT2 antagonist, to human 5-HT2A, 5-HT2B and 5-HT2C receptor subtypes by molecular modeling." Life Sciences **73**: 193-207.
Rasmussen S.G.G., Jensen A.D., et al. (1999). "Mutation of a highly conserved aspartic acid in the ß2- adrenergic receptor: constitutive activity, structural instability, and conformational rearrangements of transmembrane segment 6." Mol Pharmacol **56**(1).
Ravina E., Negreira J., et al. (1999). "Conformationally constrained butyrophenones with mixed dopaminergic (D(2)) and serotoninergic (5-HT(2A), 5-HT(2C)) affinities: synthesis, pharmacology, 3D-QSAR, and molecular modeling of (aminoalkyl)benzo- and -thienocycloalkanones as putative atypical antipsychotics." J Med Chem **42**(15): 2774-97.

Reynolds G.P. (2004). "Receptor mechanism in the treatment of schizophfrenia." J Psychopharm **18**(3): 340-345.
Rivail L., Gastineau M., et al. (2004). "New insights into the human 5-HT4 receptor binding site: exploration of a hydrophobic pocket." British J Pharm **143**: 361-370.
Robinson P.R., Cohen G.B., et al. (1992). "Constitutively actively mutants of rhodopsin." Neuron **9**(4): 719-725.
Rosendorff A., Ebersole B.J., et al. (2000). "Conserved helix 7 tyrosine functions as an activation relay in the serotonin 5HT(2C) receptor." Brain Res(84): 90-96.
Rost B. (1996). "PHD: predicting one-dimensional protein structure by profile based neural networks." Meth Enzymol **266**: 525-539.
Roth B.L., Choundary M.S., et al. (1993). "Mutagenesis of 5-HT2 serotonin receptors: what does an analysis of many mutant receptors tell us?" Med Chem Res **3**: 297-305.
Roth B.L., Choundary M.S., et al. (1997). "High affinity agonist - binding is not sufficient for agonist efficacy at 5-hydroxytriptamine2A receptors: evidence in favor of a modified ternary complex model." J Pharmacol Exp Ther **280**: 576 - 583.
Roth B.L., Craigo S.C., et al. (1994). "Binding of typical and atypical antipsychotic agents to 5-hydroxytryptamine-6 and 5-hydroxytryptamine-7 receptors." J Pharmacol Exp Ther **268**: 1403-1410.
Roth B.L. and Meltzer H.Y. (1995). The role of serotonin in schizophrenia. Psychopharmacology: the fourth generation of progress. K. D. J. Bloom F.E., Raven press**:** 1215 - 1227.
Roth B.L., Nakaki T., et al. (1984). "Aortic recognition sites for serotonin (5-HT) are coupled to phospholipase C and modulate phosphatidylinositol turnover." Neuropharmacology **23**: 1223-1225.
Roth B.L., Shoham M., et al. (1997). "Identification of conserved aromatic residues essential for agonist

binding and second messenger production at 5-hydroxytrypatmine2A receptors." Mol Pharmacol **52**: 259 - 266.
Russell R. B. and Barton G. J. (1992). "Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels." Proteins **14**(2): 309-23.
Sali A. and Blundell T. L. (1993). "Comparative protein modelling by satisfaction of spatial restraints." J Mol Biol **234**(3): 779-815.
Samama P., Cotecchia S., et al. (1993). "A mutation-induced activated state of the β2 adrenergic receptor." J Biol Chem **268**: 4625-4636.
Sander C. and Schneider R. (1991). "Database of homology-derived protein structures and the structural meaning of
sequence alignment." PROTEINS **9**: 56-68.
Scheer A., Fanelli F., et al. (1996). EMBO J **15**(14): 3566-3578.
Scheer A., Fanelli F., et al. (1997). "The activation process of the a1B-adrenergic receptor: potential role of protonation and hydrophobicity of a highly conserved aspartate." Proc Natl Acad Sci USA **94**(3): 808-813.
Schertler G.F.X., Villa C., et al. (1993). "Projection structure of rhodopsin." Nature.
Schlegel B., Sippl W., et al. (2005). "Molecular dynamics simulations of bovine rhodopsin: influence of protonation states and different membrane-mimicking environments." J Mol Model (Online) **12**(1): 49-64.
Schreiber R., Brocco M., et al. (1994). "Blockade of the discriminative stimulus effects of DOI by MDL 100,907 and the atypical antipsychotics clozapine and risperidone." Eur J Pharmacol **1994**(264): 99-102.
Sealfon S.C., Chi L., et al. (1995). "Related contribution of specific helix 2 and 7 residues to conformational

activation of the serotonin 5-HT2A receptor." J Biol Chem **270**(28): 16683-16688.
Seifert R. and Wenzel-Seifert K. (2002). "Constitutive activity of G-protein-coupled receptors: cause of disease and common property of wild-type receptors." Naunyn Schmiedebergs Arch Pharmacol **366**: 381-416.
Shapiro D.A., Kristiansen K., et al. (2000). "Differential modes of agonist binding to 5-hydroxytryptamine(2A) serotonin receptors revealed by mutation and molecular modeling of conserved residues in transmembrane region 5." Mol Pharmacol **58**: 877-886.
Shapiro D.A., Kristiansen K., et al. (2002). "Evidence for a model of agonist-induced activation of 5-hydroxytriptamine 2A serotonin receptors that involve the disruption of a strong ionic interaction between helices 3 and 6." J Biol Chem **277**(13): 11441-11449.
Sheikh S.P., Viladarga J., et al. (1999). "Similar structures and shared switch mechanism of the ß2-adrenoceptor and the parathyroid hormone receptor." J Biol Chem **274**(24): 17033-17041.
Shenkin P.S., Yarmush D.L., et al. (1987). Biopolymers **26**: 2053.
Shi L. and Javitch J.A. (2002). "The binding site of aminergic G protein-coupled receptors: the transmembrane segments and the second extracellular loop." Ann Rev Pharmacol Toxicol **42**: 437-467.
Shi L., Liapakis G., et al. (2002). "Beta2 adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch." J Biol Chem **277**(43): 40989-96.
Shi L., Lipakis G., et al. (2002). "ß2-adrenergic receptor activation." J Biol Chem **277**(43): 40989-40996.
Shoichet B.K. and Kuntz I.D. (1993). Prot Eng **6**: 723.
Sippl M. J. (1993). "Recognition of errors in three-dimensional structures of proteins." Proteins **17**(4): 355-62.

Sippl W. (2000). "Receptor-based 3D-QSAR analysis of estrogen receptor ligands-merging the accuracy of receptor-based alignments with the computational efficiency of ligand-based methods." J Comp Aid Mol Des **14**: 559-572.
Sippl W. (2002). "Binding affinity prediction of novel estrogen receptor ligands using receptor-based 3-D QSAR methods." Bioorg Med Chem **10**: 3741-3755.
Sippl W., Contreras J.M., et al. (2001). "Structure-based 3D QSAR and design of novel acetylcholinesterase inhibitors." J Comp Aid Mol Des **15**: 395-401.
Skolnick J., Kolinski A., et al. (1988). "Monte Carlo simulations of the folding of beta-barrel globular proteins." Proc Natl Acad Sci U S A **85**(14): 5057-61.
Sussman J. L., Lin D., et al. (1998). "Protein Data Bank (PDB): database of three-dimensional structural information of biological macromolecules." Acta Crystallogr D Biol Crystallogr **54**(Pt 6 Pt 1): 1078-84.
Sylte I., Bronowska A., et al. (2001). "Ligand induced conformational states of the 5-HT(1A) receptor." Eur J Pharmacol **416**: 33-41.
Takeda S., Kadowaki S., et al. (2002). "Identification of G protein-coupled receptor genes from the human genome sequence." FEBS Lett **520**: 97-101.
Teller D.C., Okada T., et al. (2001). "Advances in determination of a high-resolution three-dimesional structure of rhodopsin, a model of G-protein-coupled receptors (GPCRs)." Biochemistry **40**(26): 7761-7772.
Thompson J. D., Gibson T. J., et al. (1997). "The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools." Nucleic Acids Res **25**(24): 4876-82.
Thompson J.D., Higgins D., et al. (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucleic Acid Res **22**: 4673-4680.
Trotov M. and Abagyan R. (2001). Drug-Receptor thermodynamics: introduction and applications. Raffa R.B. Chichester, John Wiley & Sons**:** 603.
Trumpp-Kallmeyer S., Bruinvels A., et al. (1992). "Modeling of G-protein-coupled receptors: application to dopamine, adrenaline, serotonin, acetylcholine, and mammalian opsin receptors." J Med Chem **35**: 3448-3462.
Unger V.M., Hargrave P.A., et al. (1997). "Arrangement of rhodopsin transmembrane alpha helices obtained by electron cryo-microscopy." Nature **6647**: 203-206.
Urizar E., Claeysen S., et al. (2005). "An activation switch in the rhodopsin family of G protein-coupled receptors: the thyrotropin receptor." J Biol Chem **280**(17): 17135-41.
Verdonk M.I., Cole J.C., et al. (2003). "Improved protein-ligand docking using GOLD." Proteins **52**: 609-623.
Vernier P., Cardinaud B., et al. (1995). "An evolutionary view of drug-receptor interaction: the bioamine receptor family." Trends Pharmacol Sci **16**: 375-381.
Visiers I., Ballesteros J. A., et al. (2002). "Three-dimensional representations of G protein-coupled receptor structures and mechanisms." Methods Enzymol **343**: 329-71.
Visiers I., Ebersole B.J., et al. (2001). "Structural motifs as functional microdomains in G-protein-coupled receptors: energetic consideration in the mechanism of activation of the serotonin 5-HT2A receptor by disruption of the ionic lock of the arginine cage." Int. J. Quantum Chem.
Visiers I., Hassan S., et al. (2001). "Differences in conformational properties of the second intracellular loop (IL2) in 5-HT(2C) receptors modified by RNA editing can account for G protein coupling efficiency." Protein Eng **14**: 409-414.

Waller C.L. and Marshall G.R. (1993). "Three-dimensional quantitative-structure relationship of angiotensin-converting enzyme and thermolysin inhibitors II. A comparison of CoMFA models incorporating molecular orbital fields and disolvation free energies based on active analogs and complementary-receptor-field alignment rules." J Med Chem **36**: 2390 - 2403.
Waller C.L., Oprea T., et al. (1993). "Three dimensional QSAR of human immunideficiency virus (1) protease inhibitor. 1. A CoMFA study applying experimentally determined alignment rules." J Med Chem **36**: 4152-4160.
Wang C.D., Gallaher T.K., et al. (1993). "Site-directed mutagenesis of the serotonin 5-hydroxytryptamine2 receptor: identification of amino acids necessary for ligand binding and receptor activation." Mol Pharmacol **43**: 931 - 940.
Ward S.D.C., Hadman F.F., et al. (2002). "Conformational changes that occur during M3 muscarinic acetylcholine receptor activation probed by the use of an in situ disulfide cross-linking strategy." J Biol Chem **277**(3): 2247-2257.
Weinstein H. (2006). "Hallucinogen Actions on 5-HT receptors reveal distinct mechanism of activation and signaling by G protein-coupled receptors." The AAPS Journal **7**(4): e871-e884.
Weinstein H. and Ballesteros J.A. (1992). "Analysis and refinement of criteria for predicting the structure and relative orientations of transmembranal helical domains." Biophys J **62**: 107-109.
Westkaemper R. and Glennon R. (1991). "Approaches to molecular modeling studies and specific application to serotonin ligands and receptors." Pharmacol Biochem Behav **40**: 1019-1031.
Westkaemper R. and Glennon R. (2002). "Application of ligand SAR, receptor modeling and receptor mutagenesis to the discovery and development of a new class of 5-HT2A ligands." Curr Top in Med Chem **2**: 575-598.
Westkaemper R., Glennon R., et al. (1999). "Engineering in a region of bulk tolerance into the 5-HT2A receptor." Eur J Med Chem **34**: 441-447.
White S.H. and Wimley W.C. (1999). Annu Rev Biophys Biomol Struct **28**: 319-365.
Wickelgren I.A. (1988). "A route to treating schizophrenia?" Science **281**: 1264-1265.
Wood M. D., Scott C., et al. (2006). "Pharmacological profile of antipsychotics at monoamine receptors: atypicality beyond 5-HT2A receptor blockade." CNS Neurol Disord Drug Targets **5**(4): 445-52.
Woodward N.D., Meltzer H.Y., et al. (2006). "A meta-analysis of cognitive change with haloperidol in clinical trials of atypical antipsychotics: Dose effects and comparison to practice effects." Schizophr Res.
Zhang D. and Weinstein H. (1993). "Signal transduction by a 5-HT2 receptor: a mechanistic hypothesis from molecular dynamics simulations of the three-dimensional model of the receptor complexed to ligands." J Med Chem **36**: 934-938.
Zhang D. and Weinstein H. (1994). "Polarity conserved positions in transmembrane domains of G-protein coupled receptors and bacteriorhodopsin." FEBS Lett **337**(2): 207-12.
Zhou W., Flanagan C., et al. (1994). "A reciprocal mutation supports helix 2 and 7 proximity in the gonadotropin-releasing hormone receptor." Mol Pharmacol **45**(2): 165-170.

8. Annexes

Papers published during the doctoral studies:

1. Dezi C., Pastor M., Sanz F. Multistructure 3D-QSAR study on a series of conformationally constrained butyrophenones docked into a new homology model of the $5HT_{2A}$ receptor, *J. Med. Chem.* 2007, 50, 3242-3255

2. Brea J., Castro M., Loza M.I., Masaguer C.F., Raviña E., Dezi C., Pastor M., Sanz F., Cabrero-Castel A., Galán-Rodríguez B., Fernández-Espejo E., Maldonado R., Robledo P. QF2004, a potential antipsychotic butyrophenone derivative with similar pharmacological properties to clozapine, *Neuropharm*, *51*, 2006; 251-266.

3. Brea J., Masaguer C.F., Villazón M., Cadavid MI, Raviña E, Fontaine F, Dezi C, Pastor M, Sanz F, Loza MI. Conformationally constrained butyrophenones as new pharmacological tools to study $5HT_{2A}$ and $5HT_{2C}$ receptor behaviours, *Eur J Med Chem,* 2003; 38: 43340.

4. Rodrigo J., Barbany M., De Terán H.G., Centeno N.B., de Càceres M., Dezi C., Fontaine F., Lozano J.J., Pastor M., Villà J., Sanz F. Comparison of Biomolecules on the Basis of Molecular Interaction Potentials, *J Braz Chem Soc*, 2002; 13: 7959.

Paper published during the degree thesis work:

1. Cavalli A, Dezi C, Folkers G, Scapozza L, Recanatini M., Three-dimensional model of the cyclin-dependent kinase 1 (CDK1): Ab initio active site parameters for molecular dynamics studies of CDKS, Proteins, 2001 Dec 1;45(4):47885

Dezi C, Brea J, Alvarado M, Raviña E, Masaguer CF, Loza MI, Sanz F, Pastor M.
Multistructure 3D-QSAR studies on a series of conformationally constrained butyrophenones docked into a new homology model of the 5-HT2A receptor.
J Med Chem. 2007 Jul 12;50(14):3242-55. Epub 2007 Jun 19.

Brea J, Masaguer CF, Villazón M, Cadavid MI, Raviña E, Fontaine F, Dezi C, Pastor M, Sanz F, Loza MI.
Conformationally constrained butyrophenones as new pharmacological tools to study 5-HT 2A and 5-HT 2C receptor behaviours.
Eur J Med Chem. 2003 Apr;38(4):433-40.

E, Dezi C, Pastor M, Sanz F, Cabrero-Castel A, Galán-Rodríguez B, Fernández-Espejo E, Maldonado R, Robledo P. *QF2004B, a potential antipsychotic butyrophenone derivative with similar pharmacological properties to clozapine.* Neuropharmacology. 2006 Aug;51(2):251-62. Epub 2006 May 11.