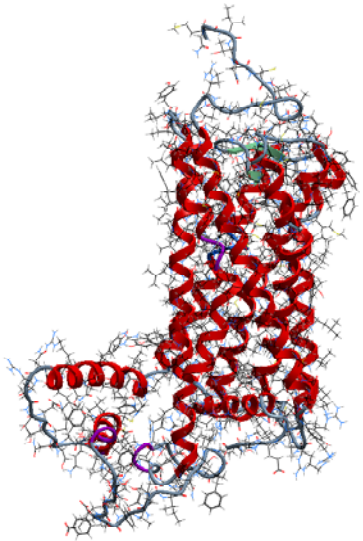
M. Sci. Thesis in Pharmacology

THE 5 HT_{1E} RECEPTOR

RECEPTOR STRUCTURE AND LIGAND INTERACTIONS

Mari Gabrielsen
October 2006



Department of Pharmacology
Institute of Medical Biology
Faculty of Medicine
University of Tromsø
Norway

ACKNOWLEDEMENTS

I would like to thank my supervisor Svein G. Dahl, and the molecular modeling group at the Department of Pharmacology at the University of Tromsø. I would especially like to thank Osman for spending endless hours helping me in the computer graphics room and the rest of the molecular modeling group for good advice and for telling me that everybody experiences problems when working with computer models.

Thank you!

TABLE OF CONTENTS

SUMMARY

The 5-HT_{1E} receptor gene was cloned in 1992 but little information on the receptor exists as no specific pharmacological tool for the receptor has been described. In this study, a model of the G-protein coupled 5-HT_{1E} receptor has been created by homology modeling using the crystallised structure of bovine rhodopsin as a template structure and a docking simulation placing 15 ligands in the putative binding site of the model has been performed in order to study structure activity relationships.

The model of the 5-HT_{1E} receptor shows that the receptor consists seven transmembrane helices forming a conserved helical bundle and one additional cytoplasmic helix in the receptor C-terminus. The putative binding site of the receptor is buried in between the transmembrane helices and residues of transmembrane helices three, four, five, six and seven are especially important for ligand binding to the receptor. The protonated amine terminal moieties of the biogenic amines are assumed to interact with the carboxyl terminus of residue D102 in transmembrane helix three.

ABBREVIATIONS

ExPASy	Expert protein analysis system
AMBER	Assisted model building and energy refinement
ICM	Internal coordinate mechanics
5-HT	5-Hydroxytryptaine, serotonin
GPCR	G-protein coupled receptor
I	Isoleucine
V	Valine
L	Leucine
F	Phenylalanine
C	Cysteine
M	Methionine
A	Alanine
G	Glycine
T	Threonine
W	Tryptophan
S	Serine
Y	Tyrosine
P	Proline
H	Histidine
E	Glutamic acid
Q	Glutamine
D	Aspartate
N	Aspargine
K	Lysine
R	Arginine
TMH	Transmembrane helix
ECL	Extracellular loop
ICL	Intracellular loop
GTP	Guanosine triphosphate

GDP	Guanosine diphosphate
RGS	Regulator of G-protein signalling
PKA	Protein kinase A
PLA	Protein lipase A
PLCβ	Protein lipase Cβ
AA	Arachidonic acid
PIP_2	Phosphatidylinositol 4,5-biphosphate
IP_3	Inositol triphosphate
DAG	Diacylglycerol
PKC	Protein kinase C
AC	Adenylyl cyclase
LH	Luteinising hormone
FSH	Follicle stimulating hormone
TSH	Thyroid stimulating hormone
GAFF	General AMBER force field
BCC	Bond charge corrections
RESP	Restrained electrostatic potential

1. INTRODUCTION

In order for the cells of the body to cooperate properly, cell signalling must occur. Cells signal to each other through compounds in the extracellular space, which the cells detect and respond selectively to. Examples of endogenous signals may be hormones, growth factors and cytokines, and the responses of a cell to external signals may be changes in gene expression, enzyme activity or in ion channel activity. When the ligands are large or hydrophilic such that they cannot penetrate the lipid plasma membrane, proteins in the plasma membrane may act as signal transducers by coupling the external signal to the biochemical responses of the cell. Hydrophobic ligands, such as some hormones, or gases penetrate the lipid bilayer directly and act on intracellular receptors.

The cloning of genes expressing receptors and ion channels, together with pharmacological studies, has shown that the diversity among the target proteins is great. This molecular diversity of the target proteins raises the possibility of discovering drugs that act selectively on the different structures. Selectivity is very important from a pharmacological viewpoint, as it means that one target structure can be targeted without affecting other structures, thereby (1) being used as a tool in determining the function and distribution of the different isoforms of the proteins and (2) later on, reducing the side effects of a drug by making the drug selective for only one isoform.

1.1 G-protein coupled receptors

There are four main types of target proteins that endogenous and exogenous compounds can act on, namely ion channels, receptors, enzymes and transport proteins. Receptors are ‘the sensing elements in the system of chemical communication that coordinates the function of all the different cells in the body’ (Rang 2003). Of the receptor types, ligand-gated ion channels (ionotropic receptors) and G-protein coupled receptors (GPCRs; metabotropic receptors) are the main receptor drug targets. Kinase-linked receptors and nuclear receptors are other receptor types.

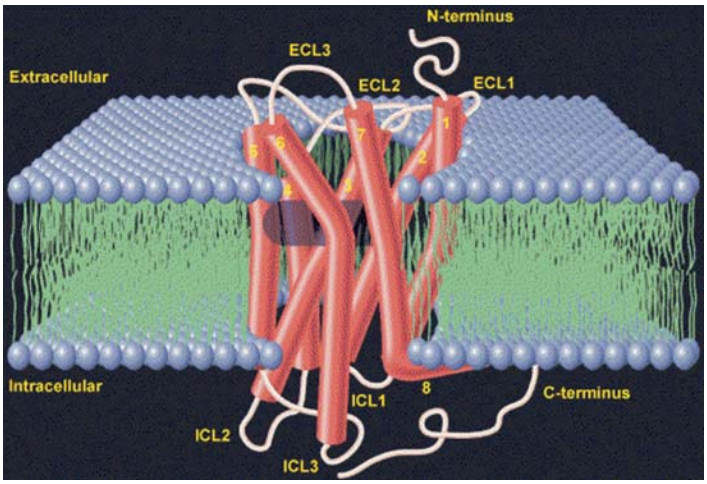


Figure 1. Schematic representation of a family A receptor in the cell membrane based on the packing arrangement of TMHs observed in the 2.6 Å crystal structure of rhodopsin (Pdb code 1L9H). The transmembrane helices are displayed as cylinders; ECL: extracellular loop; ICL: intracellular loop; 8: cytoplasmic C-terminal α helix. The putative binding site region is located between the helices. (Kristiansen 2004)

The superfamily of G-protein coupled receptor consists of proteins that share a common membrane topology, namely a nonparallel heptahelical transmembrane topology connected by extracellular and intracellular loops. The N-terminus is located extracellularly, and the C-terminus is intracellular. The helices are tilted in the membrane to obtain a proper helical packing and function of the receptor, and are arranged in an anticlockwise arrangement when viewed from the extracellular side (Kristiansen 2004). Figure 1 show a schematic representation of the helical bundle of rhodopsin in the plasma membrane.

For comparison of the family A G-protein coupled receptor sequences, a numbering system in which the positions of the residues relative to a reference residue within each transmembrane helix, is used. The reference residues of transmembrane helices 1, 3, 4, 6 and 7 are the most conserved residues in each helix, whereas the reference residues of transmembrane helices 2 and 5 are the second most conserved amino acids in these helices (Kroeze et al. 2002).

In order to determine if a sequence of unknown structure is a G-protein coupled receptor, a hydropathy plot of the receptor sequence can be generated to identify the possible segments of the sequence that may correspond to the seven helices, the hallmark of G-protein coupled receptors. A hydropathy plot is generated by calculating hydropathy indices for each amino

acid in the sequence and then plotting these indices against the residue numbers. The hydropathy index is the mean value of the hydropathy of the amino acids within a window around each position, i.e. if the window size is 9, the hydropathy index of amino acid number five in this window is calculated. Hydrophobic amino acids have high positive hydropathy scores, whereas hydrophilic amino acids have negative hydropathy scores and in transmembrane helices the hydropathy index is high for a number of consecutive positions in the sequence. The most commonly used hydropathy scale for calculation of the hydropathy indices is the scale created by Kyte-Doolittle, in which the score of each amino acid has been determined by the use of experimental data (Branden 1999).

Table 1 shows the hydropathy scores of the amino acids used to generate hydropathy plots using the Kyte-Doolittle method and figure 6 shows the hydropathy plot of the 5-HT_{1E} receptor sequence.

Table 1. Amino acid hydropathy scores based on the values reported by Kyte-Doolittle (Branden 1999).

Residue	Hydropathy score
Isoleucine (I)	4.5
Valine (V)	4.2
Leucine (L)	3.8
Phenylalanine (F)	2.8
Cysteine (C)	2.5
Methionine (M)	1.9
Alanine (A)	1.8
Glycine (G)	-0.4
Threonine (T)	-0.7
Tryptophan (W)	-0.9
Serine (S)	-0.8
Tyrosine (Y)	-1.3
Proline (P)	-1.6
Histidine (H)	-3.2
Glutamic acid (E)	-3.5
Glutamine (Q)	-3.5
Aspartic acid (D)	-3.5
Asparagine (N)	-3.5
Lysine (K)	-3.9
Arginine (R)	-4.5

1.1.1 Signal transduction by GPCRs

A suggested receptor activation mechanism involves a highly conserved motif at the intracellular end of transmembrane helix 3 (TMH3), the DRY motif, which consists of residues D3.49, R3.50 and Y3.51. This motif is present in all 5-HT receptors as well as in rhodopsin, although the residue in position 3.49 in rhodopsin is a glutamic acid, not an aspartate. Studies of the rat 5-HT_{2A} receptor suggest that the arginine in this motif is involved in a strong ionic interaction with a glutamic acid residue at the intracellular end of transmembrane helix 6, E6.30, which is thought to stabilise the inactive state of the receptor. The disruption of this interaction produces a highly constitutive active receptor with increased potency for agonists. The interaction between R3.50 and E6.30 brings the cytoplasmic ends of transmembrane helix 3 and 6 close together, primarily through the movement of the sixth transmembrane helix via a hinge at residue P6.50, the proline kink. D3.49A, E6.30Q, E6-30L, and E6.30N mutations disrupt the interaction and increase the constitutive activity of the receptor, probably by allowing the TMH3 and TMH6 to move apart (Kroeze et al. 2002). Figure 9 shows the localisation of R3.50 and E6.30 in the 5-HT_{1E} receptor.

The third intracellular loops of GPCRs (Kristiansen 2004) have been found to be important for coupling to heterotrimeric GTP-binding proteins, G-proteins in short, that mediate the intracellular actions of the receptors. Heterotrimeric G-proteins are composed of α, β, and γ subunits; the β and γ subunits being closely associated and may be regarded as one functional unit. There are at least twenty-eight distinct G-protein α subunits, five different β subunits and twelve different γ subunits which combine into a variety of functional G-proteins (Cabrera-Vera et al. 2003). The G-proteins may be inhibitory or stimulatory in their actions, and both α-GTP and the βγ complexes may target different structures in the cells, including ion channels and enzymes. Figure 3 shows some of the effects G-protein activation leads to in cells.

The activation/inactivation cycle is similar for all G-proteins. An inactive complex consisting of the three G-protein subunits is present when GDP is bound to the α subunit and this complex is associated with the receptor protein. Binding of an agonist to the receptor causes a

conformational change in the receptor and increases the affinity of the receptor for the G-protein. This leads to the rapid release of GDP from the α subunit and replacement by GTP, causing a reduction in affinity of the α subunit for the βγ complex, leading to dissociation of the G-protein heterotrimer into two subunits, namely the GTP-α subunit and the βγ subunit. These subunits can act on several effector proteins, including adenylyl cyclase, phospholipase C, tyrosine kinases, and ion channels, and activate or inactivate the proteins. The effects may be initiated directly by the G-protein subunits, or be further downstream in the signalling cascade by the actions of specific second messengers (Kristiansen 2004).

The activation cycle ends when the intrinsic GTPase of the α subunit hydrolyses GTP to GDP, causing the dissociated subunits to reassemble into an inactive complex, thereby ending the signal. This process may be accelerated by RGS (Regulators of G-protein signalling) proteins (Cabrera-Vera et al. 2003).

Figure 2 shows the G-protein activation/inactivation cycle.

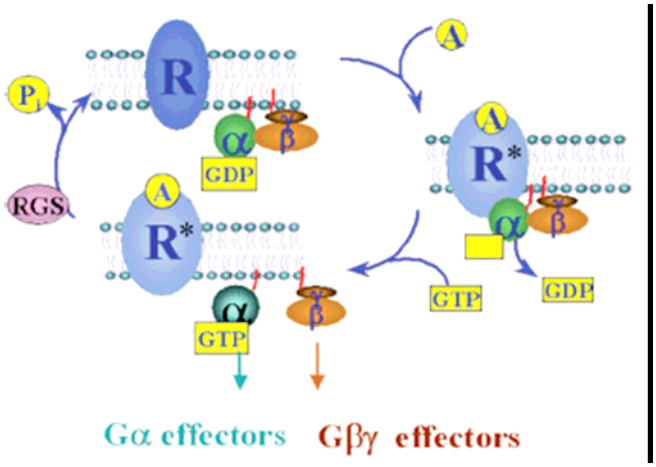


Figure 2. Receptor-mediated G protein activation. The interaction of ligand (A) with its cell surface receptor (R) facilitates the coupling of the activated receptor (R*) with intracellular heterotrimeric G proteins, which in turn promotes the exchange of GDP for GTP on the G α subunit and the dissociation of the Gα-GDP from Gβγ and the receptor. Termination of the signal occurs when the γ-phosphate of GTP is removed by the intrinsic GTPase activity of the Gα subunit, leaving GDP in the nucleotide binding pocket on Gα. Gα then reassociates with Gβγ and the cycle is complete. RGS (Regulator of G-protein Signalling) proteins accelerate the intrinsic GTPase activity of Gα subunits, thereby reducing the duration of signalling events (Cabrera-Vera et al. 2003).

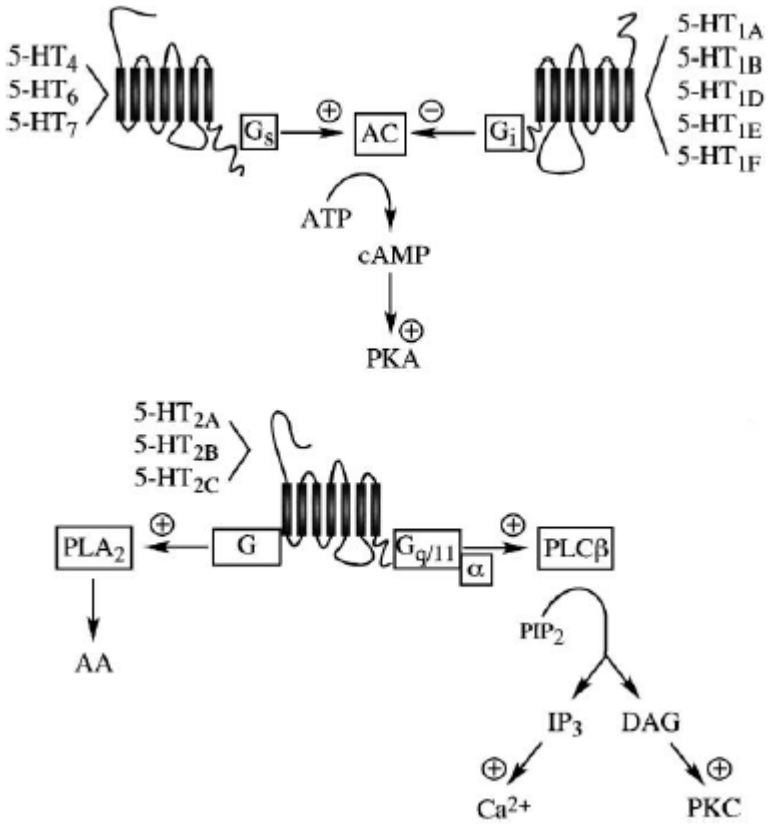


Figure 3. Prototypical signalling enzyme linkages of the G-protein coupled 5-HT receptors. AC: adenylyl cyclase; PKA: protein kinase A; PLA: phospholipase A; PLCβ: phospholipase Cβ; AA: arachidonic acid; PIP_2: phosphatidylinositol 4,5-bisphosphate; IP_3: inositol triphosphate; DAG: diacylglycerol; PKC: protein kinase C (Raymond et al. 2001)

Adenylyl cyclase is a transmembrane enzyme that catalyses the conversion of ATP to the second messenger 3',5'-cyclic AMP (cAMP), triggering the intracellular signalling pathways leading to the intracellular responses of the cells to the external stimulus. Figure 3 shows that activation of the 5-HT_1 and 5-HT_4, 5-HT_6, and 5-HT_7 receptors may stimulate or inhibit the adenylyl cyclase through G-proteins.

1.1.2 Family A of GPCRs

The superfamily of GPCRs consists of seven families of mammalian GPCRs, namely family A, B and C, large N-terminal family B-7TM (LNB-7TM), the frizzled/smoothened [F/S] family, vomeronasal 1 receptor [V1R] family, and the taste 2 receptor [T2R] family (Kristiansen 2004). Rhodopsin and the 5-HT receptors belong to the largest family, family A, which is also called the rhodopsin family. Included in this family are receptors for ligands

such as photons, odorants, biogenic amines, different hormones such as LH, FSH, and TSH, peptides, which is the largest subgroup and include neurotransmitters, hormones, and paracrines, and ligands for protease-activating receptors (Kristiansen 2004). The binding sites for most small molecule ligands of the family A GPCRs are buried in between the transmembrane helices, closer to the extracellular side than the cytoplasmic side of the membrane (Kristiansen 2004).

1.2 The 5-HT system

5-HT is the neurotransmitter for which the number of receptors is highest (Bockaert et al. 2006). The importance of 5-HT is reflected in the number of species using 5-HT as a signalling molecule –coelenterates, arthropods, molluscs, tunicates, and vertebrates all use 5-HT for signalling– and the number of 5-HT receptors found within each species. 15 human 5-HT receptors have at this date been characterised.

1.2.1 5-HT

5-HT (5-hydroxytryptamine, serotonin) is a monoamine that is acts as a neurotransmitter in the CNS and as a paracrine in peripheral vascular systems. Even though 5-HT is present in the diet, most of it is metabolised before entering the bloodstream. In the body, 5-HT is synthesised from tryptophan via 5-hydroxytryptophan by the enzymes tryptophan hydroxylase and a non-specific amino acid decarboxylase, respectively. Excretion of 5-HT in urine occurs after 5-HT is metabolised by monoamine oxidase to 5-hydroxyindole acetaldehyde and then by aldehyde dehydrogenase to five-hydroxyindoleacetic acid (5-HIAA) (Rang 2003).

In the CNS, the cell bodies of the neurons that release 5-HT are grouped into the pons and upper medulla, close to the midline (raphe) and are often referred to as *raphe nuclei*.

1.2.2 5-HT receptors

Currently, 12 human 5-HT receptors have been cloned and classified into seven 5-HT receptor families. Six families encode G-protein coupled 5-HT receptors, while one family, the 5-HT_3 family, are ionotropic channels and are not discussed here.

The 5-HT_1 receptor family is the largest 5-HT receptor family and contains five human 5-HT receptors, termed 5-$HT_{1A,}$ 5-$HT_{1B,}$ 5-$HT_{1D,}$ 5-HT_{1E}, and 5-HT_{1F}. The former 5-HT_{1C} family has been reclassified as the 5-HT_{2C} receptor based on sequence similarities to the 5-HT_{2A} and 5-HT_{2B} receptors and similar second messenger systems (Raymond et al. 2001). In addition to the 5-HT_1 receptors, three human 5-HT_2 receptors, the 5-$HT_{2A,}$, 5-$HT_{2b,}$ and 5-HT_{2C} receptors, and one human 5-$HT_{4,}$ 5-$HT_{5,}$ 5-$HT_{6,}$ and 5-HT_7 receptor, have been cloned.

Most of the G-protein coupled 5-HT receptors are expressed not only in humans or mammalians but also in many other species, such as rat and guinea pig. The 5-HT_{1D} receptor, for instance, is found in mouse, rabbit, and rat in addition to human species, whereas the 5-HT_{1E} receptor has so far only been cloned in the human and guinea pig genomes (Bai et al. 2004). The human and guinea pig 5-HT_{1E} receptors share 95 percent sequence homology.

In addition to the 12 human G-protein coupled 5-HT receptors that have been cloned, some 5-HT receptors are also modified post-genomic by alternative splicing and mRNA editing, which results in even more diversity. The 5-HT_{2C} receptor is the only known GPCR that is regulated by mRNA editing, whereas splice variants of both the 5-HT_4 receptor and the 5-HT_7 receptors have been described (Bockaert et al. 2006). The tissue specific mRNA editing involves nucleotide substitution, most frequently adenosine to inosine, or cytidine to uridine, and appears to regulate the pattern of intracellular signalling. Of today, ten functional splice variants of the 5-HT_4 receptor have been described. The different forms of the 5-HT_4 receptor vary in their C-termini, and the different variants may be in charge of the fine-tuning of signal transduction as they interact with specific intracellular proteins. The 5-HT_7 receptor is also modified by alternative splicing, also resulting in three receptors differing in their C-termini. The 5-HT_1 genes are intronless and cannot be modified by alternative splicing (Bockaert et al. 2006).

The genes encoding the 5-HT_1 receptors constitute a subgroup of genes that were diverged from a common ancestor (Bockaert et al. 2006). The 5-HT receptors vary in length from 358 to 482 amino acids, but still 33 residues are invariant and additional 27 residues that are at least 80 percent conserved among these receptors (Kroeze et al. 2002). Figure 4 shows a phylogenetic tree showing the evolutionary relationship of the human 5-HT receptors and bovine rhodopsin.

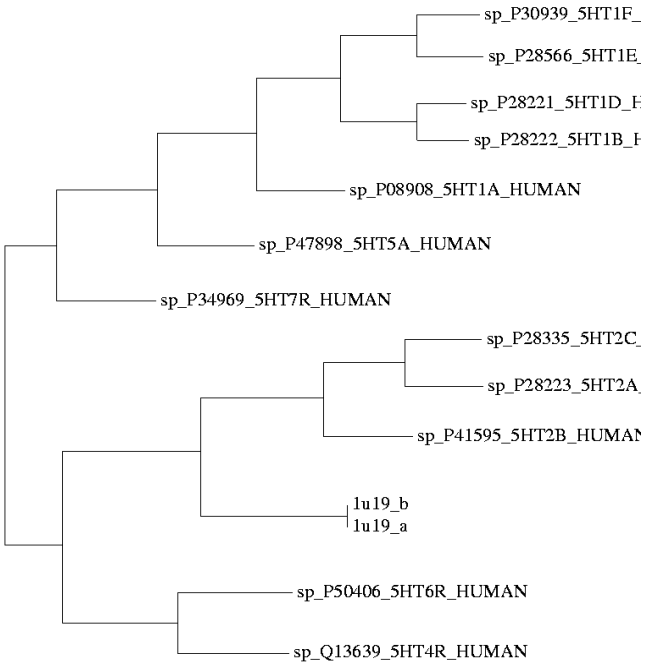


Figure 4. Evolutionary tree generated by ICM by sequence alignment of the human 5-HT receptors and heterodimeric bovine rhodopsin amino acid sequences. 1u19 a,b: rhodopsin; sp_: ExPASy accession code of human 5-HT receptors.

The 5-HT_{1E} receptor shares 39 percent sequence homology with the 5-HT_{1A} receptor, 47 percent (64 percent in transmembrane regions) with the 5-HT_{1B} and 5-HT_{1D} receptors, and 70 percent homology with the 5-HT_{1F} receptor (Raymond et al. 2001).

Table 2. Reference residues and residues with possible ligand binding functions and receptor activation functions in the transmembrane helices of 5-HT receptors. ECL: extracellular loop

Residue numbering	5-HT_{1E} receptor position	Assumed function	Selected references
1.50	N40	Reference residue	
2.50	D68	Reference residue	(Mirzadegan et al. 2003)
		Receptor activation?	(Kroeze et al. 2002)
3.25	C93	Disulfide bridge	(Kroeze et al. 2002)
		(with C173 in ECL2)	(Mirzadegan et al. 2003)
3.28	W98	Hydrophobic ligand binding pocket	(Kristiansen 2004)
3.32	D102	Important for ligand binding. Stabilisation of protonated amine moiety of ligands.	(Kristiansen 2004) (Kroeze et al. 2002)
3.49	D119	DRY motif:	(Kroeze et al. 2002)
3.50	R120	Receptor activation	
3.51	Y121	Reference residue	
4.50	W147	Reference residue	
5.45	S186	Possible hydrogen	(Kristiansen 2004)
5.46	T187	bonding residues?	
5.47	F191	Hydrophobic ligand	(Kristiansen 2004)
5.48	Y192	binding pocket	
5.50	P194	Reference residue	

Table 2 continued. Reference residues and residues with possible ligand binding functions and receptor activation functions in the transmembrane helices of 5-HT receptors.

Residue numbering	5-HT_{1E} receptor position	Assumed function	References
6.30	E285	Receptor activation	(Kroeze et al. 2002)
6.44	F300	FxxxW motif:	(Mirzadegan et al. 2003)
6.48	W304	Receptor activation? Hydrophobic ligand binding pocket	(Kristiansen 2004)
6.50	P306	Reference residue	
6.51	F307	Hydrophobic ligand	(Kristiansen 2004)
6.52	F308	binding pocket	
6.54	K310	Ligand binding?	(Kroeze et al. 2002)
6.55	E311		
7.36	D327	Ligand binding	(Kristiansen 2004) (Kroeze et al. 2002)
7.40	W331	Fingerprint residue; Ligand binding Hydrophobic ligand binding pocket	(Kristiansen 2004)
7.43	Y334	Fingerprint residue; Stabilising charged ligands/ receptor activation Hydrophobic ligand binding pocket	(Kristiansen 2004)
7.49	N340	NPxxY motif:	(Mirzadegan et al. 2003)
7.50	P341	Receptor activation	
7.53	Y344	Reference residue	

Table 1 shows residues that have been suggested to be involved in ligand binding or receptor activation. Especially important residues include residues D102 (3.32) and Y334 (7.43), which are fingerprint residues of biogenic and trace amine receptors. These residues have been suggested to play important roles in ligand binding. The side chains of these residues may function to stabilise the amine terminal moieties of the biogenic ligands that bind to these receptors and they may also act as a receptor activation switch (Kristiansen 2004). Many aromatic residues have also been proposed to form a hydrophobic ligand binding pocket around the ligands. Studies also suggest that residues R3.50 and E6.30 and residues D2.50 and N7.49 may interact through ionic interactions and hydrogen bonding, respectively, and may play roles in receptor activation (Kroeze et al. 2002). Another study suggests that two residues in the sixth transmembrane helix may account for some pharmacological differences between the 5-HT_{1E} and the 5-HT_{1B} receptors, namely the residues K310 (6.54) and E311 (6.55) in the 5-HT_{1E} receptor sequence (Kroeze et al. 2002).

The N- and C-termini of the 5-HT receptors are, together with the extracellular and intracellular loops, the least conserved regions of GPCRs. The amino acid residues in the extracellular loops vary among the receptors, as they may be important for specific ligand recognition. The intracellular loops show somewhat more similarity, suggesting a common coupling mechanism to G-proteins (Mirzadegan et al. 2003). The third intracellular loop is the most probable site of G-protein coupling in 5-HT receptors (Kroeze et al. 2002).

Of the 5-HT_1 receptors, the 5-HT_{1A} receptor has been best characterised. The reason for this is that there are many specific pharmacological tools for the receptor; in addition, the receptor was the first of the 5-HT receptors to be cloned. Many studies, including various physiological, clinical, behavioural, and pharmacological studies, show that the 5-HT_{1A} receptor plays potential roles in conditions such as depression and anxiety and also have potential roles in neuroendocrine function and thermoregulation, vasoreactive headaches, sexual behaviour, food intake, and immune function (Raymond et al. 2001). The 5-HT_{1A} receptors are found in the highest densities in the limbic system, where the receptors are

located both pre-synaptically (5-HT_{1A} auto-receptors on the soma and dendrites of 5-HT neurons) and post-synaptically. 5-HT_{1A} auto-receptors control the synthesis and release of 5-HT, and selective serotonin reuptake inhibitors (SSRIs), which currently are the most commonly used drugs for the treatment of depression (Rang 2003), are in fact efficient only after desensitisation of the 5-HT_{1A} auto-receptors.

5-HT_{1A} receptor activation stimulates neurogenesis, the creation of new neurons, in the hippocampus (Bockaert et al. 2006). Post mortem and brain imaging studies reveal that depressed or anxious patients have loss of neurons in the prefrontal cortex and hippocampus. An observation that stress, which may cause depressive episodes in humans, also decreases the hippocampal neurogenesis, suggest that this process may be involved in the pathogenesis of mood disorders (Santarelli et al. 2003). Neurogenesis has been shown to increase in response to antidepressants and 5-HT_{1A} receptor agonists are used for the treatment of anxiety. The azaspirodecanedione class of 5-HT_{1A} receptor acting drugs, which includes buspirone, is used for the treatment of anxiety. Buspirone is a partial agonist at the post-synaptic 5-HT_{1A} receptors and a full agonist at the 5-HT_{1A} autoreceptors {Felleskatalogen.no, #47}. These selective 5-HT_{1A} receptor agonists have fewer side effects than some of the other common antianxiolytic drugs, such as the benzodiazepines.

1.2.3 The 5-HT_{1E} receptor

The intronless human 5-HT_{1E} receptor gene was cloned in 1992 (Levy et al. 1992) (McAllister et al. 1992) (Zgombick et al. 1992). The human 5-HT_{1E} receptor is a 365 amino acid long G-protein coupled receptor that shares greatest sequence homology with the receptors of the 5-HT_1 receptor family. As seen in figure 4, the 5-HT_{1E} receptor protein shows the highest sequence homology with the 5-HT_{1F} receptor and lowest sequence homology with the 5-HT_{1A} receptor among the 5-HT_1 receptor family members.

Little is known about the function or distribution of the 5-HT_{1E} receptor protein as no selective ligands for the receptor yet have been described. The 5-HT_{1E} receptor has been shown to inhibit adenylyl cyclase at low concentrations in transfected HeLa and BS-C-1

cells. It has also been shown that high concentrations of agonist can stimulate the adenylyl cyclase in BS-C-1 cells (Raymond et al. 2001).

1.3 Structure determination

1.3.1 Structure determination of membrane proteins

The RCSB PDB database (Berman et al. 2000) contains the structure of every protein experimentally determined, either by x-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy or electron microscopy. The database also contains experimentally determined structures of nucleic acids (1676 structures) and of protein/nucleic acids complexes (1585 structures). 34 other structures, e.g. the structure of the glycoprotein antibiotic vancomycin, the connective tissue agarose double helix, and the carbohydrate cycloamylose, are also included in the database. Of October 3rd 2006, there are 35,909 protein structures deposited in the PDB database and nearly ninety percent of these have been determined by x-ray crystallography methods. However, of these nearly 36,000 deposited structures, under 500 are membrane proteins. As most current drug targets are membrane proteins, much work is being done in determining the structure of the membrane proteins to better understand how drugs and endogenous compounds act and designing drugs that act more selectively, thereby reducing unwanted side effects of drugs.

The three-dimensional structures of membrane proteins are usually difficult to determine using experimental methods such as x-ray crystallography. One of the reasons is that crystallisation involves solving the proteins in an aqueous solution and then removing the solution slowly in order to generate crystals. Membrane proteins come from a lipid environment and their surfaces have both hydrophobic and hydrophilic properties, making the proteins insoluble in the aqueous buffer and denature in organic solvents. In order to crystallise membrane proteins, detergents and small amphipathic molecules are added to the aqueous solution. The hydrophobic parts of the detergents bind to the hydrophobic parts of the proteins whereas the hydrophilic parts of the detergents face the surrounding solution and give the protein-detergent complex an essential hydrophilic surface while burying the hydrophobic parts. Addition of small amphipathic molecules to protein-detergent solutions may promote crystallisation of some proteins, probably by facilitation proper packing

interactions between the molecules in all three dimensions in a crystal (Branden 1999). Addition of detergent is necessary for maintaining the structural integrity of the membrane protein outside the lipid bilayer but it is also adds flexibility to the complex and this normally means that obtaining a crystal that diffracts with a high resolution is hard.

Obtaining membrane proteins in sufficient quantities for crystallisation may also be difficult. Membrane proteins are usually not expressed in great amounts in cells and must be over-expressed in expressing systems in order to obtain sufficient amounts of the protein for crystallisation. For some proteins, receptors especially, over-expression of the proteins may strongly stimulate the cells, causing the cells to undergo apoptosis. Receptors may also be difficult to over-express. In contrast to X-ray crystallography, NMR is primarily limited to relatively small proteins, usually smaller than 25 kDa, making it unsuitable for predicting the structure of most proteins.

One G-protein coupled receptor has been crystallised and the structure solved by x-ray diffraction with high resolution, namely the inactive form (the *cis* form) of the family A bovine membrane protein rhodopsin, the visual pigment in rod photo-receptor cells. Rhodopsin was crystallised using the hanging-drop vapour diffusion method to a resolution of 2.2 Å (PDB id 1U19) with the detergent heptylthioglucoside (Okada et al. 2004). Sequence comparison and molecular modeling studies support the hypothesis that most family A members are folded in the same manner as rhodopsin (Kristiansen 2004). The structure of rhodopsin is the only crystal structure of any GPCR that include the seven transmembrane segments, and bovine rhodopsin is used as a template structure in the modeling of GPCRs.

1.3.2 Molecular modeling

Molecular modeling is ‘the generation, manipulation and representation of three-dimensional molecular structures and their associated physical, chemical, biological and pharmacological properties’ (Ravna 2006). The homology modeling approach is based on the observation that structure is more conserved than sequence, such that a known protein structure can be used to construct a model of a homologous protein. The known protein structure is termed the

'template' structure, whereas the protein with unknown structure is the 'target' protein. Homology modeling of a membrane protein is done by aligning the target protein with the amino acid sequence of a homologue membrane structure with known folding pattern, having the start and endpoints of transmembrane domains predicted, and building a model based on this alignment.

Molecular mechanical calculations, which are calculations on atomic nuclei, and quantum mechanical calculations, which are calculations on the electronic systems of molecules, are both included as techniques in molecular modeling. The Laws of thermodynamics state that molecules spontaneously seek the lowest energy and by performing energy minimisation calculations on the model, the lowest energy conformation of the model is calculated. Molecular dynamics is the simulation of molecular motion during a short period of time and is used to refine the model further (Ravna 2006).

1.3.3 Structure prediction using ICM and AMBER

Programs such as ICM and AMBER that are used to generate three-dimensional models and refine them, uses force fields that are determined through experimental work and high-level quantum mechanical calculations. Force fields are the functional form and parameter sets that are used to describe the potential energy of a system of atoms. The basic functional form of force fields includes bonded terms, which relate to atoms that are linked by covalent links, whereas the nonbonded terms describe the long-range electrostatic van der Waals forces. The specific composition of the terms depend on the force field, but a general form for the total energy can be written as:

$$E_{tot} = E_{covalent} + E_{noncovalent}$$

where $E_{covalent} = E_{bond} + E_{angle} + E_{dihedral}$ and $E_{noncovalent} = E_{electrostatic} + E_{vanderWaals}$.

Force fields also define parameter sets for each atom in a model. The parameter sets include variables for atomic mass and partial charge for each atom, and equilibrium bond lengths and angles for pairs, triplets, and quadruplets of bonded atoms. For instance, force fields include distinct bonding patterns for an oxygen atom in a carbonyl group and in a hydroxyl group. The parameter sets are derived from studies on small organic molecules, which are more manageable for experimental studies and quantum mechanical calculations.

The ICM software package can be used to predict the structure of a protein by homology modeling when there is no detailed three-dimensional structure of the target available, which is the case for most membrane proteins. Included in the ICM package is an all atom internal force field, which is a force field that provides parameters for all atoms in the system, hydrogen included. In addition, ICM contains the MMFF94 force field. After model building, ICM may be used in energy refinement of the generated models and for docking of small flexible ligands, see later paragraph.

The AMBER package consists of a set of molecular mechanical force fields and a package of molecular simulation programs. AMBER may be used to perform energy minimisations and molecular dynamics studies and for analysing the results. There are three main steps in AMBER simulations, namely preparation, simulation and analysis, and different programs of the AMBER package carry out the different steps. During the preparation for molecular dynamics each atom in the molecule is assigned an atom or particle type. A coordinate file (*prmcrd*), which contains the Cartesian coordinates of all the atoms, and a parameter-topology file (*prmtop*), which contains all other information (atom names and masses, force field parameters, lists of bonds, angles, and dihedrals are needed, as well as additional bookkeeping information) needed to compute energies and forces are generated during the preparation phase. The main preparation programs in AMBER are ANTECHAMBER and LEaP, which functions include assembling force fields for residues or molecules that are not defined otherwise and construction of biological molecules. ANTECHAMBER programs include antechamber, which performs molecular conversion (for example, conversion of a pdb file to a prep file or a Gaussian input file) and assigns atom type and generates charge. The parmchk program is used to assign an addition force field (in a frcmod file) if parameters are missing in the prep file.

AMBER force fields have limited parameters for organic molecules and have for that reason not been widely used in drug design and other studies of ligand-protein or ligand-DNA interactions. Another force field that covers most drugs, the GAFF (General AMBER force field) force field, is used when constructing the ligands for binding energy calculations and molecular docking studies (Wang et al. 2004).

The Gaussian 03 program uses the HF/6-31G* basis set for performing quantum mechanical calculation on the ligands for assignment of RESP charges in antechamber. RESP charges are preferred for AMBER simulations. The AM1-BCC charge method imitates the HF/6-312G* basis set for calculations of electrostatic potential of a molecule and quickly generates atomic point charges that may be used for computer simulations (Jakalian et al. 2002). The ligand charges are calculated using the AM1-BCC method for molecular docking in ICM.

1.3.4 Molecular docking

Molecular docking aims to predict the structure of the intermolecular complex that is formed between two or more molecules and has become a useful tool in structure-based drug design and discovery (Sousa et al. 2006). When performing molecular docking, protein flexibility is a critical aspect. Proteins are not rigid structures and ideally the proteins should be flexible during the docking simulations. However, introducing flexibility in the macromolecular proteins during docking is not yet possible in ICM, as the computational workload will be great. Instead docking studies in ICM are performed using a flexible ligand and a rigid receptor protein.

1.3.5 Ramachandran plot

Ramachandran plots, where the psi and phi angles of amino acid residues are plotted against each other, may be constructed to evaluate a model built by homology modeling.

Amino acids are linked together by covalent bonds at the Cα atoms and the only degrees of freedom they have are rotations around these bonds, i.e. around the Cα-C' bond (psi, ψ) and the N-Cα bond (phi, φ), as the CO-NH (the peptide bond) is rigid and planar due to double bond character. In a Ramachandran plot, the angle pairs ψ and φ are plotted against each other in a diagram called a Ramachandran plot, which shows allowed combinations of the ψ and φ angles. Most combinations of ψ and φ angles for an amino acid are not allowed because they cause steric collisions between the side chains and main chain. Each point in the Ramachadran plot represents psi and phi values for an amino acid residue (Branden 1999).

α helices in proteins are found when a stretch of consecutive residues all have the φ,ψ angle pair approximately -60° and -50°, corresponding to the allowed region in the bottom left quadrant of the Ramachandran plot. In the Ramachandran plot, left-handed α helices are seen in the upper right quadrant of Ramachandran plots, whereas β strands are seen in the upper left quadrant of the plot.

Residues that may be expected to be located outside the allowed regions of the Ramachandran plot are glycine and proline. The glycine side chain consists of only a hydrogen atom and steric collisions do not occur that often as with the other amino acid residues that have longer side chains. In the case of proline, the side chain is bonded to the main chain nitrogen atom and forms a ring structure, thereby preventing the nitrogen atom from participating in hydrogen-bonding and also providing some sterical hindrance to the α helical conformation.

2. AIM OF STUDY

Insight into ligand-receptor interactions is of pivotal importance for designing new ligands with therapeutic potential. In order to study these interactions three-dimensional structural information about the receptor structure is necessary. The detailed three-dimensional structure of the 5-HT_{1E} receptor is not known, but the x-ray crystallographic structure of bovine rhodopsin, which shares the same three-dimensional topology as the 5-HT_{1E} receptor, gives the possibility of using the homology modeling approach to construct a three-dimensional model of the 5-HT_{1E} receptor.

The specific aims of the present study were:

1. Construction of a model of the 5-HT_{1E} receptor using molecular modeling
2. Docking of a series of ligands into the putative binding site of the 5-HT_{1E} receptor model using automatic docking
3. Study structure activity relationships of the 5-HT_{1E} receptor.

3. METHODS

3.1 Construction of the 5-HT_{1E} receptor model in ICM

The Molsoft ICM program package version 3.4-4 was used to align the rhodopsin and 5-HT receptor sequences and to build the model of the 5-HT_{1E} receptor based on this sequence alignment. The reason for aligning all 5-HT receptor amino acid sequences and bovine rhodopsin was to avoid mismatching of the 5-HT_{1E} receptor amino acid sequence and rhodopsin, which is more likely to occur when only the two amino acid sequences are aligned. The crystallised structure of bovine rhodopsin B chain (PDB id: 1U19) was used as the template structure for homology modeling of the 5-HT_{1E} receptor target. The rhodopsin protein is a heterodimeric protein but as the chains do not differ much, the A chain just as well could have been used for the modeling of the 5-HT_{1E} receptor. The amino acid sequences of the other 5-HT receptors and of rhodopsin chain A were removed from the alignment before building of the model.

The $5HT_{1E}$ receptor and rhodopsin structures were aligned with ICM and the alignment was manually adjusted to avoid gaps in the helices and making sure that the highly conserved residues in each transmembrane helix were aligned together. The alignment was also adjusted so that the two cysteine residues that form the disulfide bridge, C95 and C173 in the 5-HT_{1E} receptor sequence, were aligned with the corresponding cysteine residues in the rhodopsin sequence. The model of 5-HT_{1E} receptor was constructed using the Build model function in ICM. In ICM, the most conserved domains are used as 'rigid bodies' by using the average position of Cα atoms in the domain to keep the most conserved regions fixed during the building procedure. The loops were constructed by homology modeling with existing loops in the PDB database. The loop sequences were used as input for searching the PDB database for corresponding sequences, and the loops were selected based in internal energy and its interaction energy with loop environments on the model. Prediction of the loops in ICM is very unreliable when the loops have more than three residues.

3.2 Refinement of the $5HT_{1E}$ receptor model by ICM and AMBER

The refineModel macro in ICM was used to energy refine the 5-HT_{1E} receptor model generated in the previous step. By using this macro, the side chains are globally optimised and the backbone annealed in order to find the lowest energy conformation of the model. The first step of refinement is sampling of the side chain conformational positions using 'Montecarlofast'. In the second step, the backbone molecules are randomly moved repeated times (harmonic restraints pull the atoms to static points in space represented by the corresponding atoms in the template), followed by a local energy minimisation. The complete energy for the structure is then calculated. The resulting structure is either accepted or discarded based on its energy and temperature (a high temperature indicates unfavourable structures). After the annealing of the backbone, a second side chain sampling is performed.

Following the refinement of the 5-HT_{1E} receptor model in ICM, the model was further refined by using the AMBER8 package to perform molecular dynamics on the model. Before running molecular dynamics on a structure generated by homology modeling in ICM, an energy minimisation in AMBER needed to be performed. To minimise the structure, the SANDER program of AMBER was used. The first minimisation was a short steepest descent minimisation that was performed not to reach energy minimum but to relieve possible bad steric interactions in the structure. The NCYC flag controls which minimisation algorithm SANDER uses. When NCYC is lower than MAXCYC, which is the total number of minimisation cycles, SANDER uses the steepest descent algorithm for the first NCYC steps, then switches to conjugate gradient algorithm for the remaining. The 5-HT_{1E} receptor model was minimised 250 cycles using the steepest descent algorithm, followed by 250 steps of conjugate gradient minimisation during the first energy minimisation in AMBER (MAXCYC 500; NCYC 250). The second energy minimisation of the 5-HT_{1E} receptor was a longer conjugate gradient minimisation, where the first 500 minimisation cycles were performed using the steepest descent algorithm, followed by 2000 cycles of conjugate gradient minimisation (MAXCYC 2500; NCYC 500).

After energy minimisations, the SANDER program of AMBER was used to perform a implicit solvent Generalised Born molecular dynamics simulation. During the molecular dynamics simulation, the helices were restrained in Cartesian space using a harmonic

potential of 10 kcal/mol, as it turned out that the helices moved too much if they were not restrained, causing disruption of the conserved protein core. The nonbonded cutoff radius was set at 12 Å. The initial temperature in the system was 0 Kelvin, and the temperature was increased until the reference temperature in which the system was to be kept, 300 Kelvin, was reached. 500,000 MD steps were run with a time step of 0.001 ps, resulting in a 500 ps MD simulation.

3.3 Ramachandran plot

The ICM program was used to generate Ramachandran plots of the structure both after the refinement by ICM and after the model had been refined by ICM and AMBER (500 ps MD simulation). The Ramachandran plots were generated in order to determine if there were many amino acids except glycine that had psi and phi angles outside their allowed regions. Many amino acid residues outside their allowed regions in a model may indicate that the model is not optimal. The Ramachandran plot of the 5-HT_{1E} receptor refined by ICM is presented in figure 12.

3.4 Construction of ligand models

The xLEaP program of the AMBER molecular package was used to construct models of the ligands seen in figure 5 for docking. The ligands were built in xLEaP using the GAFF force field. After constructing the ligands, the ANTECHAMBER program of AMBER was used to create prep files with BCC charges for docking.

The SANDER program was used to minimise the ligands. The first 500 minimisation cycles were performed using the steepest descent algorithm, followed by 2000 cycles of conjugate gradient minimisation. After minimisation, pdb files were generated from the topology and restart coordinate files using the AMBPDB conversion program.

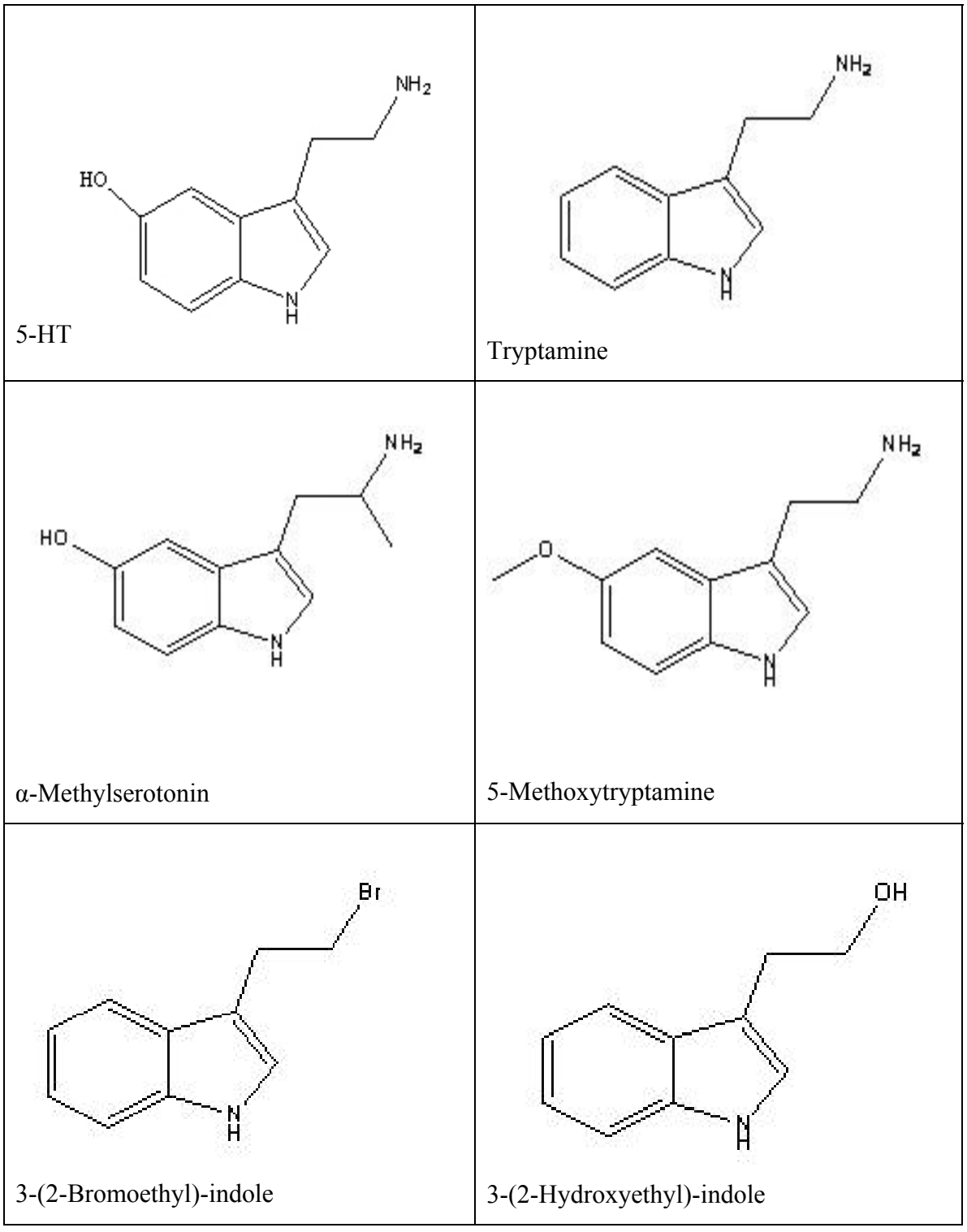


Figure 5. Two-dimensional representation of the nonprotonated forms of the ligands generated in AMBER and docked in ICM. The isomeric form of ligand 2 is not shown.

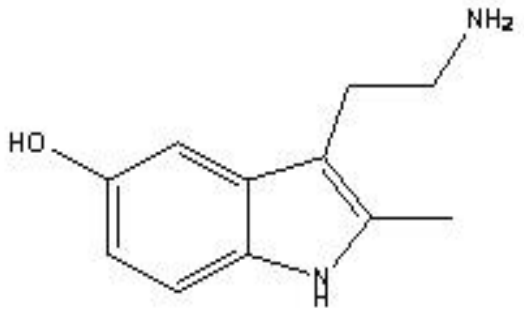
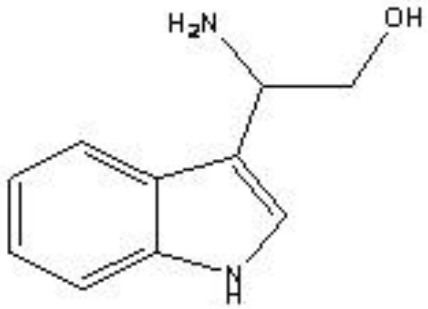
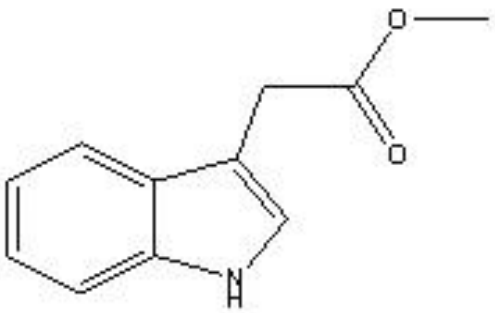
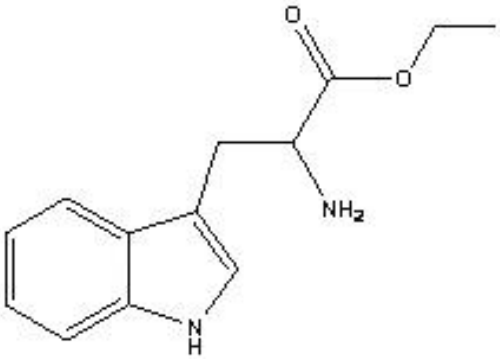
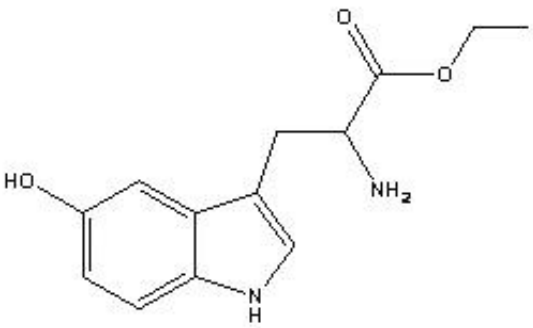
 2-Methyl-5-hydroxytryptamine	 Tryptophanol
 Methyl-3-indoylacetate	 Tryptophane ethylester
 5-hydroxy-tryptophane ethylester	

Figure 5 continued. Two-dimensional representation of the nonprotonated forms of the ligands generated in AMBER and docked in ICM. Isomeric forms of ligand 7, ligand 9 and ligand 10 are not shown.

3.5 Docking of ligands in the 5-HT_{1E} receptor

The aspartate residue in TMH3, D102, and its neighbours within 5 Å were selected as the main binding region for ligand interaction based on mutational studies of other receptors. As seen in figure 5, the ligands to be docked in this project are very similar and the batch docking method was therefore used. Batch docking is a docking procedure where the ligands are converted to mol files and contained in the same file. Instead of docking each ligand separately, all ligands are docked in one run and the results are displayed in the order the ligands are placed in the file containing all ligands.

ICM performs docking of fully flexible ligands in a rigid receptor using a Monte Carlo minimisation procedure in internal coordinates to find the global minimum of the energy function. A torsional or positional conformation change, followed by local minimisation, is performed at each step.

3.6 Calculation of binding energy

The docking results were observed and ligand orientations (poses) that were outside the defined binding area were discarded. Ligand poses that were in the binding area but where the ligands terminal amine moieties were not in the proximity of the side chain oxygen atoms of the aspartate residue D102 were also discarded, as well as the poses that were very close to or colliding with receptor residues. The poses of the ligands that do not contain a terminal amine moiety were accepted when they were located in the putative binding area of the receptor.

The calcBindingEnergy macro of ICM was used to calculate the energy of binding of the accepted ligand poses to the receptor. This macro evaluates the binding of each ligand in complex with the receptor by estimating the electrostatic, hydrophobic and entropic binding terms of the complex. The calculated binding energies are presented in table 3 and 4.

3.7 Minimisation of receptor and ligand complexes in AMBER

The results from the molecular docking in ICM showed that some of the ligand poses in the putative binding area were very close to receptor residues and some seemed to be colliding with the residues, while the ligand poses for two ligands, 5-methoxytryptamine and the R-form of tryptophanol, all had positive binding energy. In an attempt to further improve the results from the docking study in ICM, the SANDER program of AMBER was used to minimise each accepted receptor-ligand complex.

The Gaussian 03 program was used to assign the ligands RESP charges for energy minimisation in AMBER, as RESP charges are the charges suggested for AMBER calculations. After the energy minimisation of the receptor-ligand complexes, the SANDER program of AMBER was used to convert the restart coordinate files from the energy minimisation to pdb files that were loaded in ICM. The new binding energies of the complexes after the energy minimisation in AMBER were calculated in ICM using the calcBindingEnergy macro.

4. RESULTS

4.1 Homology modeling

4.1.1 Hydropathy plot of the 5-HT_{1E} receptor sequence

Figure 4 shows a hydropathy plot of the 5-HT_{1E} receptor sequence using a window size of 9.

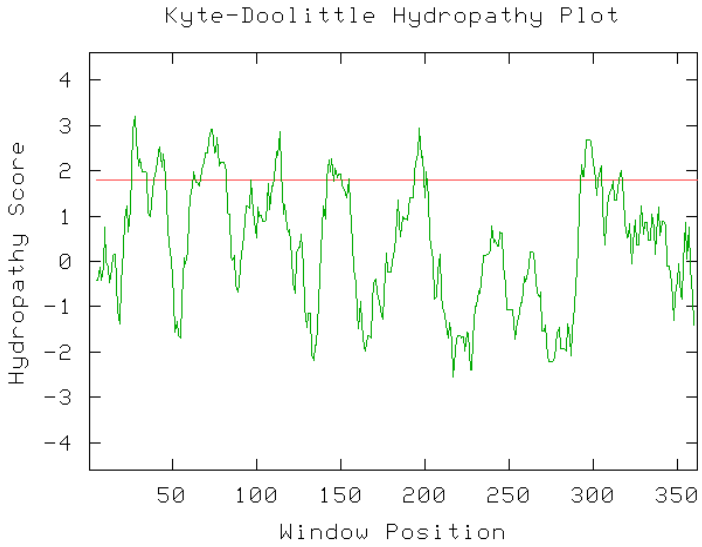


Figure 6 Kyte-Doolittle Hydropathy plot of the $5HT_{1E}$ receptor sequence. The window size is 9 was used to generate the hydropathy scores. Peaks that have scores greater than 1.8 (indicated by red line) show indicate transmembrane regions. The plot was generated using the server at http://occawlonline.pearsoned.com/bookbind/pubbooks/bc_mcampbell_genomics_1/medialib/activities/kd/kyte-doolittle.htm.

4.1.2 Alignment of the 5-HT_{1E} receptor and rhodopsin amino acid sequences

Figure 7 shows the alignment of the 5-HT_{1E} receptor sequence and the rhodopsin B chain sequence and also the putative localisation of the transmembrane helices.

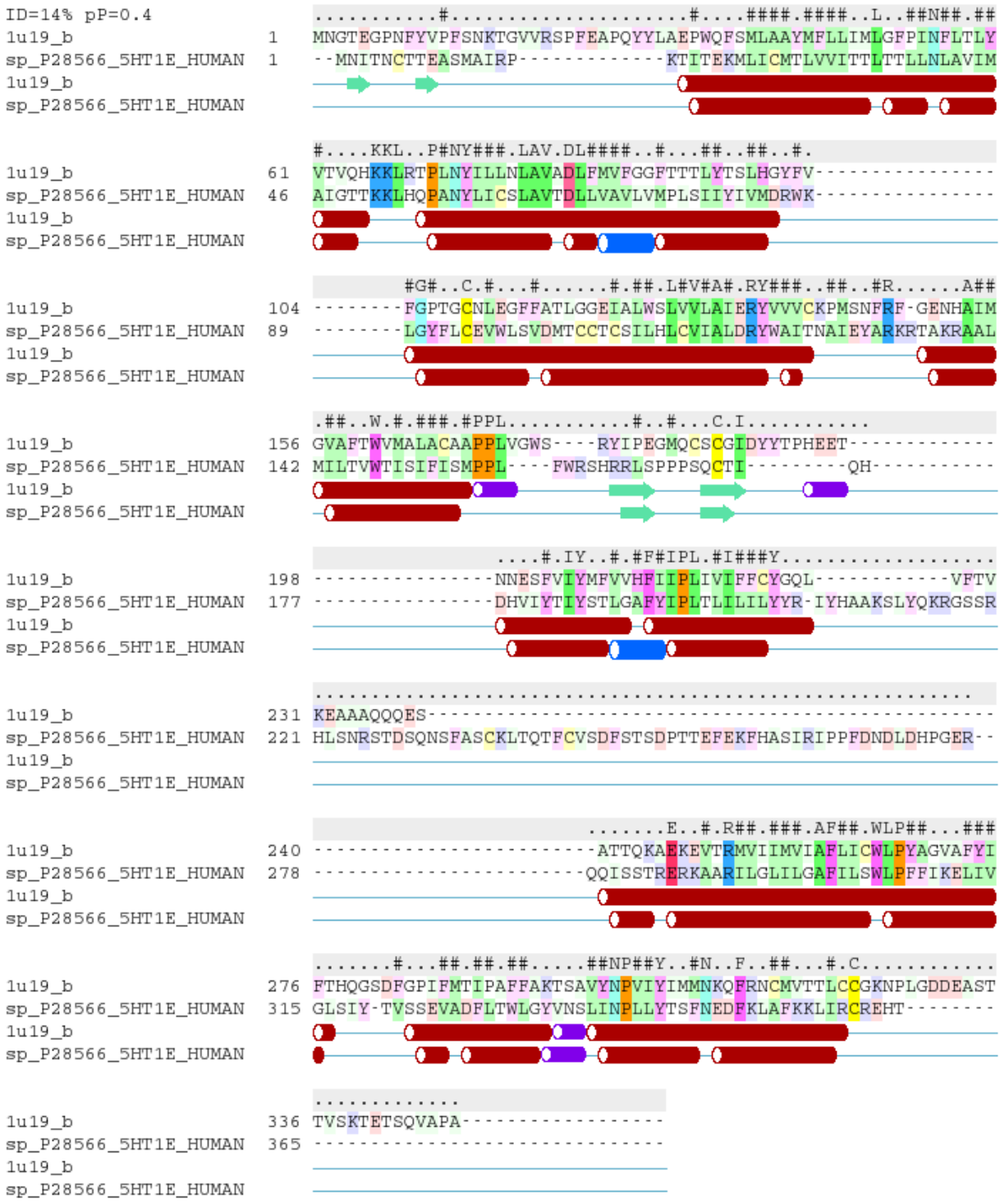


Figure 7. Alignment of the 5-HT_{1E} receptor sequence and the rhodopsin B chain sequence (1u19_b) by ICM. Barrels indicate the putative transmembrane helices of the two receptors while green arrows indicate β strands.

4.1.3 5-HT_{1E} receptor structure

The 5-HT_{1E} receptor model that had only been refined in ICM was used to interpret the results in this thesis. Figure 6 shows the model of the 5-HT_{1E} receptor constructed in ICM with the putative binding area displayed. The binding area is located toward the extracellular side, beneath the beta strand motif of the second extracellular loop. Figure 8 shows that the putative binding area of the 5-HT_{1E} receptor is buried between the helices toward the extracellular side of the membrane beneath two beta strands of the second extracellular loop. The figure shows, in addition to the seven transmembrane helices, the presence of two α helices in the longer third intracellular loop and one α helix located precisely after the end of transmembrane helix 7 in the C-terminal end of the receptor that runs along the membrane.

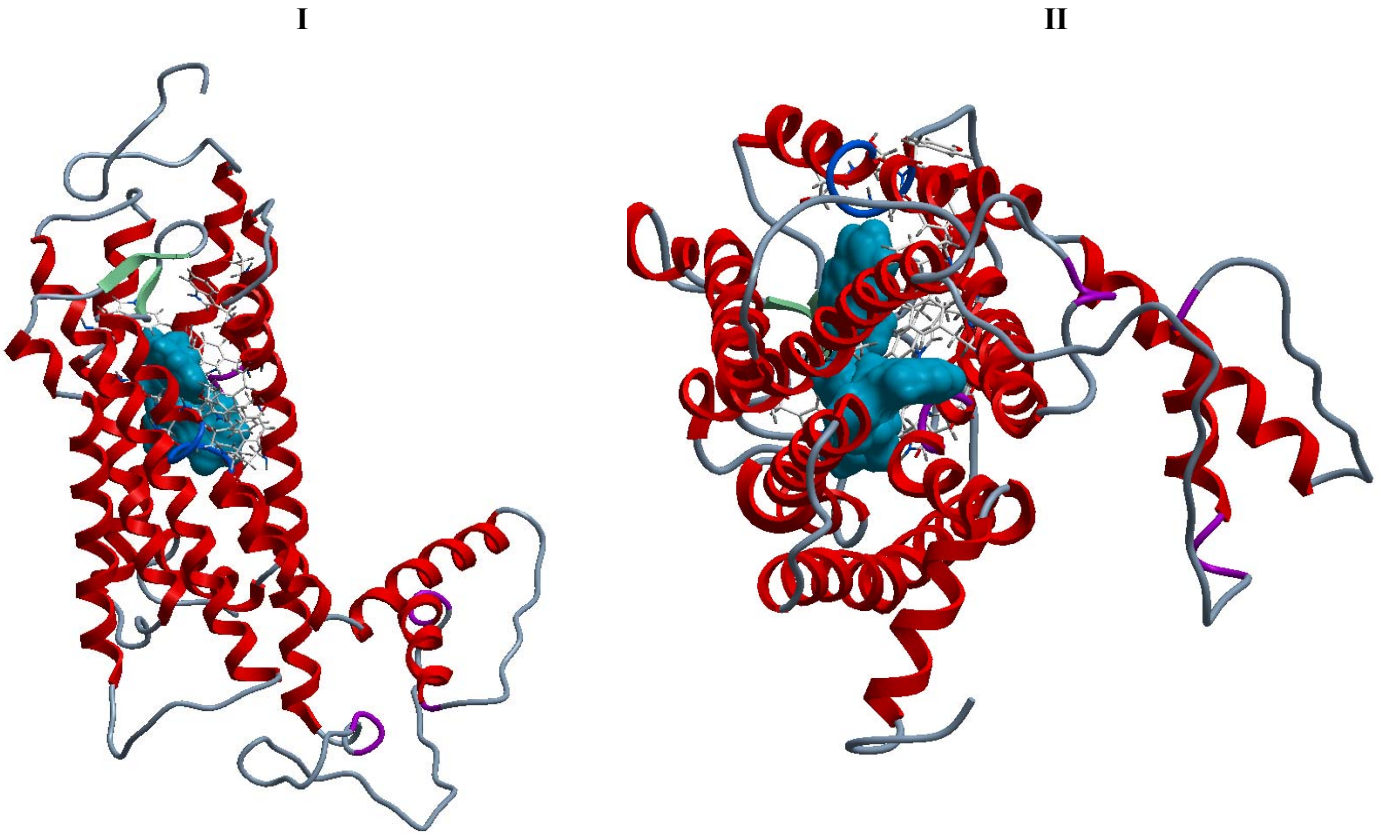


Figure 8. 5-HT ligand pose A, tryptamine poses B and D, and S-5-hydroxy-tryptophane ethylester ligand pose A displayed using skin representation. The receptor viewed in the membrane plane (I) and from the cytoplasmic side (II). The residues that are displayed are the same as those displayed in the other figures. Alpha helices are displayed in red, beta strands in green and pi and 3_{10} helices in blue and purple, respectively.

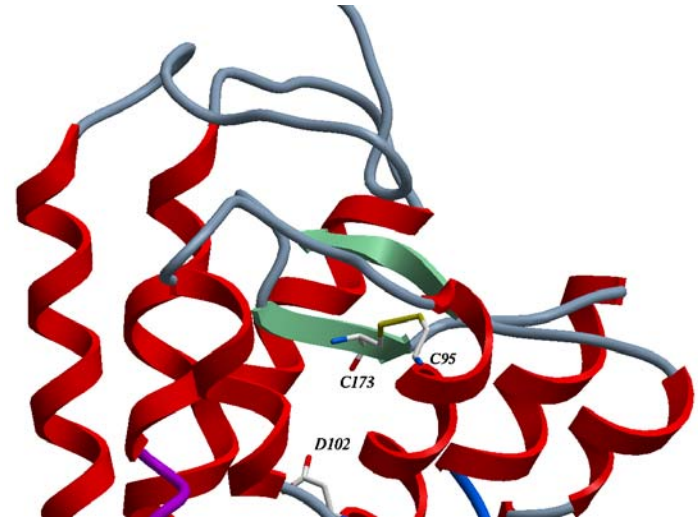
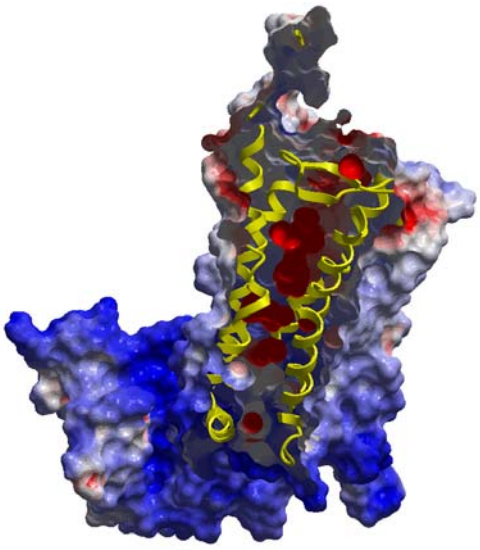


Figure 9. The 5-HT_{1E} receptor. (I) The parts of the transmembrane helices of the 5-HT_{1E} receptor toward the extracellular side shown in membrane plane and localisation of the D102 in transmembrane helix 3 relative to the second extracellular loop containing two β strands (represented by green arrows). The disulfide bridge between C95 and C173 is also shown. The transmembrane regular α helices are coloured in red and pi and 3_{10} helices in blue and purple, respectively. (II) Localisation of residues R120 (R3.50) and E286 (E6.30) at the intracellular part of the 5-HT_{1E} receptor.

Figure 9 shows the localisation of the aspartate of transmembrane helix 3 relative to the second extracellular loop and the localisation of the two residues R120 and E286 suggested to be important in receptor activation. The major difference seen when comparing the model of the 5-HT_{1E} receptor refined in ICM with that refined in both

ICM and AMBER, are in the loop areas. Representations of the electrostatic potential of the 5-HT_{1E} receptor model with the binding area displayed are presented in figure 10.

I



II

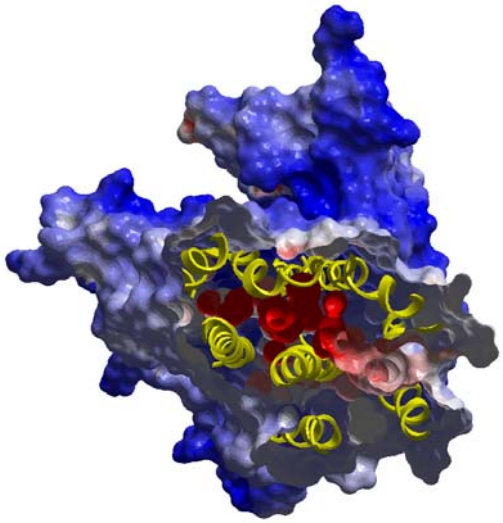


Figure 10. 5-HT_{1E} receptor model viewed in the membrane plane (I) and from the extracellular side (II). Clipping plane has been used to expose the ligand binding area of the receptor. Red colour indicates negatively charged areas, blue colour indicates positively charged areas, white colour indicates neutral areas. The helices are displayed in yellow.

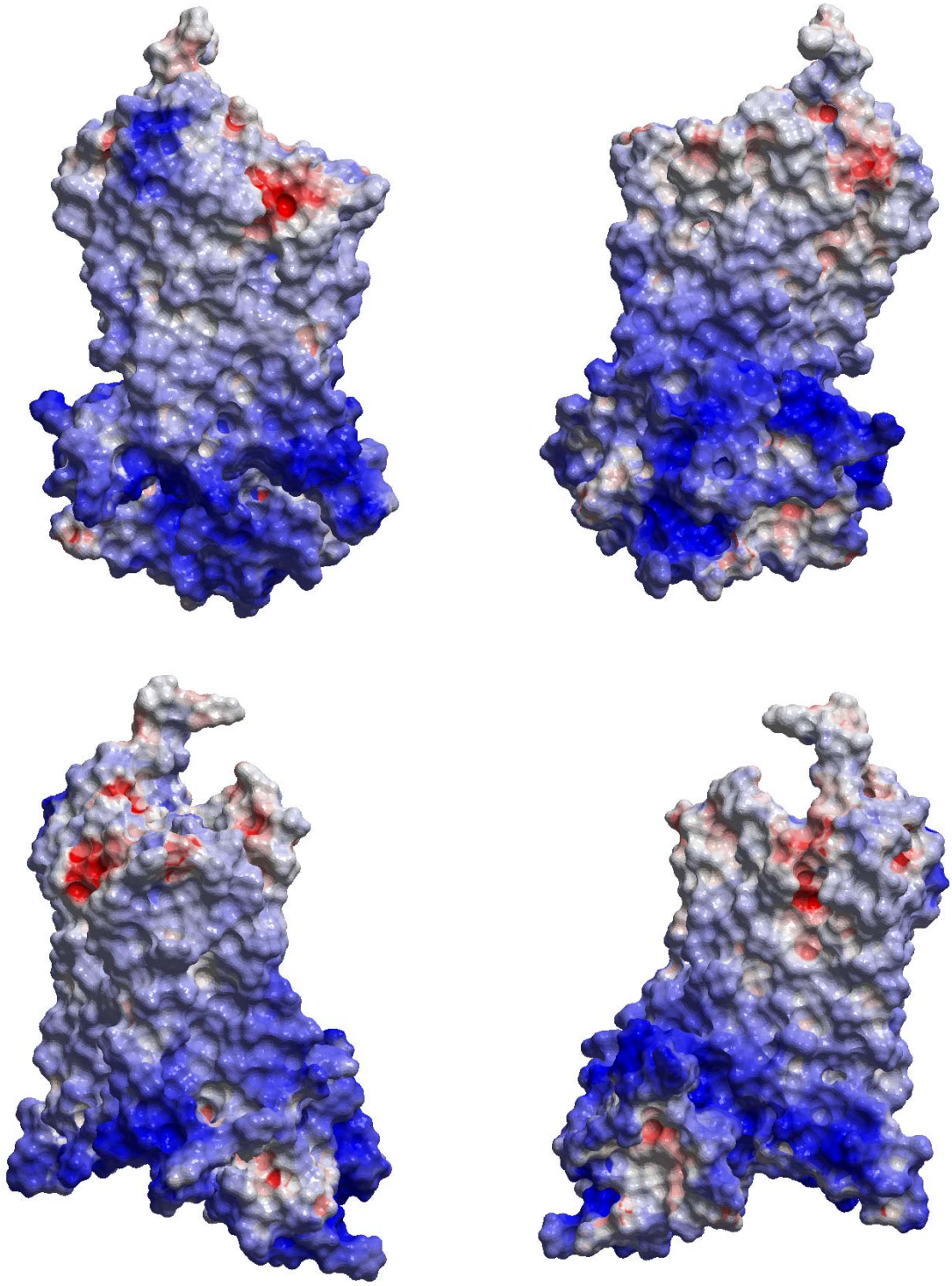


Figure 11. The 5-HT_{1E} receptor model coloured by EPS viewed in the membrane plane. Red colour indicates negatively charged areas, whereas blue colour indicates positively charged areas.

4.1.4 Ramachandran plot

Figure 9 shows the Ramachandran plot of the 5-HT_{1E} receptor refined in ICM. The Ramachandran plot of the receptor after refinement in both ICM and AMBER showed that the exact same amino acid residues had psi and phi angles outside the allowed areas and is not shown.

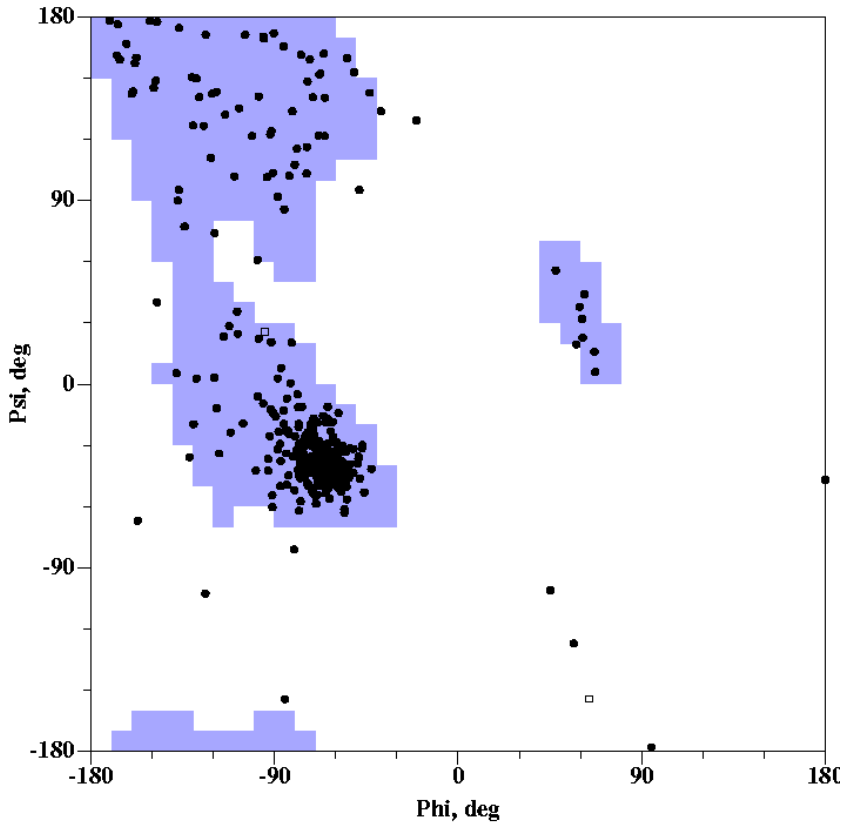


Figure 12. Ramachandran plot of the 5-HT_{1E} receptor model refined in ICM showing allowed combinations of the conformational angles phi and psi in blue areas. Each point in the plot represents psi and phi values for an amino acid residue.

There are 21 non-glycine amino acid residues that are outside their allowed regions in the Ramachandran plot. Five are located in the N-terminal part of the structure (M1, T7, M12, A13, R15), one in the first intracellular loop close to transmembrane helix 2 (H54), one is located at the boundary of transmembrane helix 2 and the first extracellular loop (T125) and one is located at the boundary of transmembrane helix 4 and the second extracellular loop (P156). There are three amino acid residues that are outside their allowed regions in the second extracellular loop: L166, which is located in the first β

strand of this loop, and two consecutive proline residues (P168 and P169), which form a motif known to end secondary structure elements, which are located at the end of the first β strand and causes a kink in the structure of the loop at this location. Another residue outside the allowed region is found at the boundary of the second extracellular loop and the fifth transmembrane helix (D178), while one residue is located in the middle of transmembrane helix 5, amino acid residue F191. This residue is found in the pi-helix of this transmembrane helix. Four other residues outside their allowed regions are located in the third intracellular loop, namely S232, F233, S235 and C236. The last non-glycine residues outside the allowed regions of the Ramachandran plot are found at the boundary of transmembrane helix 6 and the third extracellular loop (S317), and precisely after the helix that runs along with the plasma membrane in the cytoplasm (helix 8), the residues C361, R362, and E363.

The Ramachandran plot shows that most of the residues are located in right-handed α helices but it also shows the presence of a short left-handed α helix. The plot further shows the presence of β strands.

4.2 Docking results

As the results after docking of the ligands in the 5-HT_{1E} receptor model refined in both ICM and AMBER were disappointing (see calculated binding energies in table 4), the results generated by docking in the 5-HT_{1E} receptor model refined in ICM were chosen to be interpreted here. The docking results from the docking in the ICM refined 5-HT_{1E} receptor model are presented in figure 13 and in appendix 2.

4.2.1 Docking results, 5-HT

The results from docking of 5-HT in the 5-HT_{1E} receptor are presented in figure 10. Figure 5 shows the two-dimensional representation of the ligand docked.

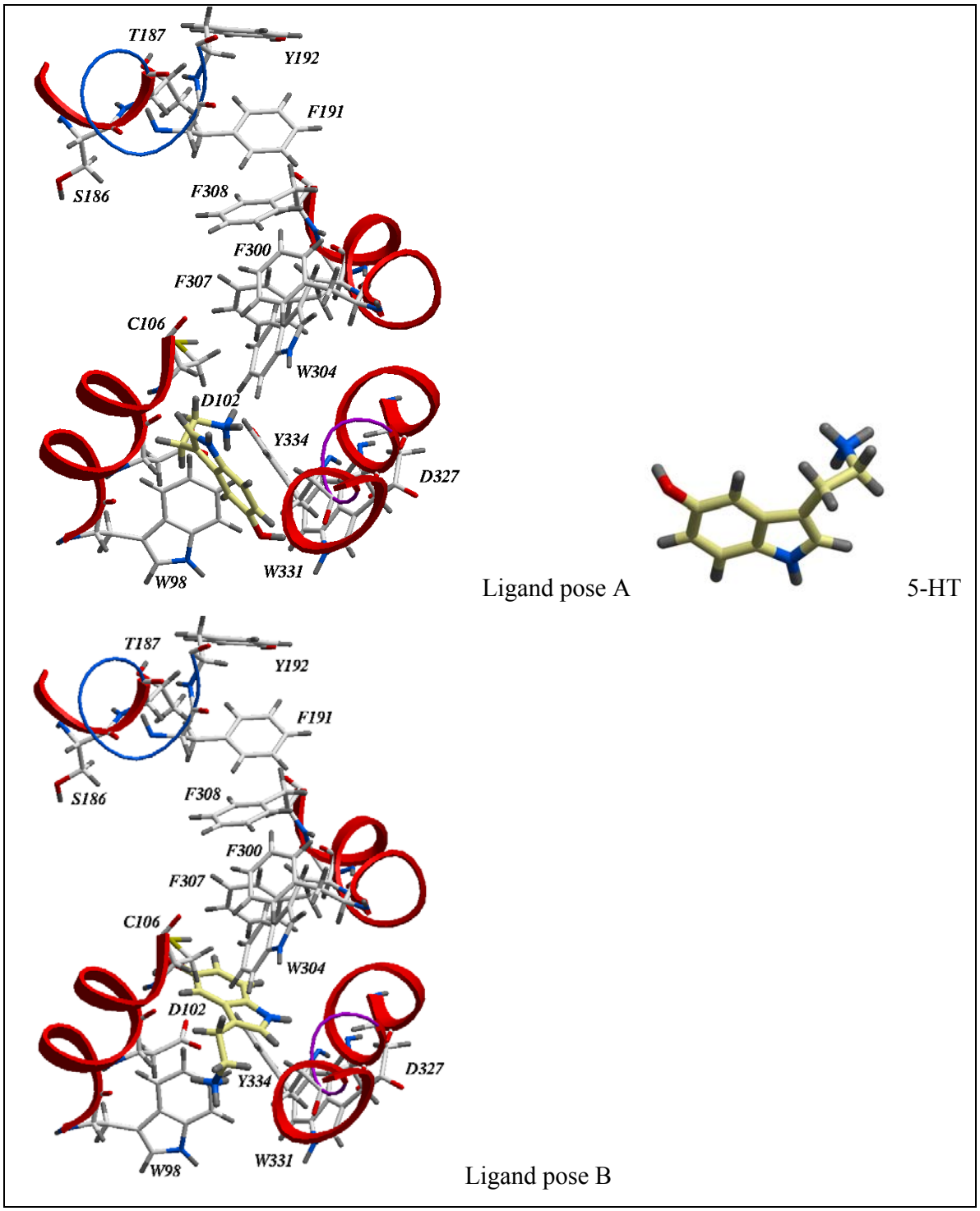


Figure 13. 5-HT in the putative binding site of the 5-HT_{1E} receptor model. Alpha helices are displayed in red, 310/phi in blue/purple and the receptor amino acid side chains and the ligand are displayed by atom type, in which blue colour indicates nitrogen atoms and red colour oxygen atoms.

Ligand pose A (ΔG: -0.94 kcal/mol) forms five hydrogen bonds to receptor residues. Four of the hydrogen bonds are between the ligand protonated amine terminal and the carboxyl side chain group of D102 and one between the ligand 5-hydroxyl group to the receptor residue Y334 main chain oxygen atom.

Ligand pose B (ΔG: -0.38 kcal/mol) forms four hydrogen bonds with the receptor residues. One hydrogen bond is formed between the ligand indole amine group and the backbone oxygen atom of receptor residue T330 and three hydrogen bonds are formed between the ligand protonated amine terminal and the side chain carboxyl group of residue D102.

4.2.2 Docking results, Tryptamine

The results from docking tryptamine in the 5-HT_{1E} receptor are presented in appendix A2.1. Figure 5 shows the two-dimensional representation of the ligand docked.

Ligand pose A (ΔG: -3.90 kcal/mol) forms three hydrogen bonds with the receptor. Two hydrogen bonds are formed between the ligand amine terminal moiety and the side chain carboxyl group of residue D102, and one hydrogen bond is formed between the ligand indole amine group and the main chain oxygen atom of D102.

Ligand pose B (ΔG: -1.97 kcal/mol) forms five hydrogen bonds with receptor residues. Two hydrogen bonds are formed between the ligand terminal amine moiety and the D102 side chain carboxyl group and three hydrogen bonds are formed to the D102 side chain oxygen atom.

Ligand pose C (ΔG: -1.94 kcal/mol) is positioned in almost exactly in the same position as ligand pose B and forms similar hydrogen bonds.

Ligand pose D (ΔG: -0.44 kcal/mol) forms two hydrogen bonds with receptor residues. One hydrogen bond is formed between the ligand terminal amine moiety and the D102 side chain carboxyl group and one is formed to the Y334 side chain hydroxyl group.

4.2.3 Docking results, α-Methylserotonin, R-form

The result of docking of R-form of α-Methylserotonin in the 5-HT_{1E} receptor model is shown in appendix A2.2a. Figure 5 shows the two-dimensional representation of the ligand docked.

Ligand pose A (ΔG: -1.05 kcal/mol) forms three hydrogen bonds to the receptor: two hydrogen bonds are formed from the ligand terminal amine moiety to the receptor residue D102 side chain carboxyl oxygen atoms, and one hydrogen bond from its indole amine group to the backbone oxygen atom of residue T330. Figure 5 shows the two-dimensional representation of the ligand docked.

4.2.4 Docking results, α-Methylserotonin, S-form

The results from docking of the S-form of α-Methylserotonin in the 5-HT_{1E} receptor model are presented in appendix A2.2b. Figure 5 shows the two-dimensional representation of the ligand docked.

Ligand pose A (ΔG: -4.07 kcal/mol) forms seven hydrogen bonds to the receptor residues. Four hydrogen bonds are formed between the ligand protonated amine terminal and the D102 side chain carboxyl group, and one hydrogen bond from the ligand 5-hydroxyl group to the C106 backbone nitrogen atom and to the D102 backbone oxygen atom. There is also one hydrogen bond between the ligand indole amine group and the side chain hydroxyl group of residue S337.

Ligand pose B (ΔG: -3.89 kcal/mol) forms one hydrogen bond from the terminal amine moiety and the D102 side chain oxygen atom od2, and three hydrogen bonds are formed

between the amine moiety and the main chain D102 oxygen atom. In addition, one hydrogen bond is formed between the ligand 5-hydroxyl group and the D102 hydroxyl group and two hydrogen bonds to the Y334 hydroxyl group.

Ligand pose C (ΔG: -1.73 kcal/mol) forms two hydrogen bonds to the receptor residue D102 side chain carboxyl oxygen atoms from the ligand amine terminal moiety and one hydrogen bond to the Y334 hydroxyl group. In addition there are hydrogen bonds between the ligand indole amine group and the backbone oxygen atom of residue T330, and between the ligand 5-hydroxyl group and the D102 backbone oxygen atom.

4.2.5 Docking results, 5-Methoxytryptamine

The result from docking of 5-Methoxytryptamine in the 5-HT_{1E} receptor is presented in appendix A2.3. Figure 5 shows the two-dimensional representation of the ligand docked.

The two poses of this ligand that were located in the putative binding area of the 5-HT_{1E} receptor had positive binding energy. The ligand pose with the best binding energy (ΔG 4.44 kcal/mol) is presented here. This ligand pose formed one hydrogen bond from its terminal amine to the side chain oxygen atom of residue D102. The receptor-ligand complex was energy minimised using the SANDER program of AMBER, which resulted in negative binding energy for the complex (ΔG = -5.53 kcal/mol).

4.2.6 Docking results, 3-(2-Bromoethyl)-indole

The results from docking of 3-(2-Bromoethyl)-indole in the 5-HT_{1E} receptor are shown in appendix A2.4. This ligand does not contain a terminal amine moiety and ligand poses were accepted when they were positioned in the putative binding area of the receptor. Figure 5 shows the two-dimensional representation of the ligand docked.

Ligand pose A (ΔG: -2.36 kcal/mol) forms one hydrogen bond to the receptor, namely between the ligand 5-hydroxyl group and the carboxyl side chain group of residue D102.

Ligand pose B (ΔG: -2.00 kcal/mol) forms one hydrogen bond to the carboxyl group of residue D102 from the ligand 5-hydroxyl group.

Ligand pose C (ΔG: -1.68 kcal/mol) forms one hydrogen bond to the carboxyl group of residue D102 from the ligand 5-hydroxyl group.

4.2.7 Docking results, 3-(2-Hydroxyethyl)-indole

The results from docking of 3-(2-Hydroxyethyl)-indole in the 5-HT_{1E} receptor are presented in appendix A.2.5. This ligand does not contain a terminal amine moiety and the ligand poses accepted were positioned in the putative binding area of the receptor. Figure 5 shows the two-dimensional representation of the ligand docked.

Ligand pose A (ΔG: -3.24 kcal/mol) forms one hydrogen bond between the ligand side chain hydroxyl group and the side chain amine group of residue N336.

Ligand pose B (ΔG: -2.31 kcal/mol) forms three hydrogen bonds to receptor residues. Two hydrogen bonds are formed between the ligand side hydroxyl group and the backbone nitrogen atoms of N336 and S337, and one hydrogen bond between the ligand indole ring amine group and the main chain oxygen of residue D102.

Ligand pose C (ΔG: -1.98 kcal/mol) forms four hydrogen bonds. One hydrogen bond is formed between the ligand side chain hydroxyl group and the backbone oxygen atom of residue D102 and the C106 side chain, one between the ligand indole amine group and the main chain oxygen atom of residue C173 and one to the side chain oxygen of T174.

Ligand pose D (ΔG: -0.21 kcal/mol) forms hydrogen bonds to the carboxyl group of the D102 side chain from the indole amine group, as well as from the ligand hydroxyl group to the backbone nitrogen atom of S337 and to the main chain oxygen atom of residue Y334.

4.2.8 Docking results, 2-Methyl-5-hydroxytryptamine

The result from docking 2-Methyl-5-hydroxytryptamine in the 5-HT_{1E} receptor is shown in appendix A2.6. Figure 5 shows the two-dimensional representation of the ligand docked.

Ligand pose A (ΔG: -2.99 kcal/mol) is the only ligand pose of this ligand in the putative binding area of the 5-HT_{1E} receptor. The ligand forms seven hydrogen bonds to residue D102 and one hydrogen bond to residue Y334: There are three hydrogen bonds from the ligand terminal amine to the main chain oxygen atom of D102, two bonds from the terminal amine to the carboxyl group of the D102 side chain, two hydrogen bonds between the D102 carboxyl side chain group and the ligand 5-hydroxyl group, and one hydrogen bond between the ligand 5-hydroxyl group and the side chain hydroxyl group of residue Y334.

4.2.9 Docking results, Tryptophanol, R-form

The results from docking of the tryptophanol R-form in the 5-HT_{1E} receptor is shown in appendix A2.7a. Figure 5 shows the two-dimensional representation of the ligand docked.

All five ligand poses of this ligand that were located in the putative binding area of the 5-HT_{1E} receptor had positive binding energy. The ligand pose with the best binding energy (ΔG 4.80 kcal/mol) is presented here. This ligand forms five hydrogen bonds to residue D102; two from the ligand terminal amine to the carboxyl group of residue D102 and two hydrogen bonds to the backbone oxygen atom of the residue. There is also a hydrogen bond between the ligand indole amine group and the side chain hydroxyl group of residue S337. The receptor-ligand complex was energy minimised using the SANDER program of AMBER, which resulted in negative binding energy for the complex (ΔG = -2.58 kcal/mol).

4.2.10 Docking results, Tryptophanol, S-form

The results from docking of the tryptophanol S-form in the 5-HT_{1E} receptor is shown in appendix A2.7b. Figure 5 shows the two-dimensional representation of the ligand docked.

Ligand pose A (ΔG: -4.16 kcal/mol) is the only conformation of the S-form of tryptophanol. Six hydrogen bonds are formed between the ligand and residue D102: five from the ligand terminal amine to the side chain carboxyl group of residue D102 and one between the ligand side hydroxyl group and the main chain oxygen atom of residue D102. From the ligand hydroxyl group there is also one hydrogen bond to the backbone nitrogen atom of residue C106 and one to the backbone oxygen atom of residue V101.

4.2.11 Docking results, Methyl-3-indoylacetate

The results from docking of methyl-3-indoylacetate in the 5-HT_{1E} receptor are presented in appendix A2.8. This ligand does not contain any terminal amine moiety and the ligand poses accepted were positioned in the putative binding site of the 5-HT_{1E} receptor. Figure 5 shows the two-dimensional representation of the ligand docked.

Ligand pose A (ΔG: -2.22 kcal/mol) forms nine hydrogen bonds with the receptor. There are two hydrogen bonds between the ligand side chain methylacetate oxygen atoms and residue T174 side chain oxygen and one hydrogen bond to the Q172 side chain amine group. There are also three hydrogen bonds between the ligand methylacetate oxygen atoms and the side chain amine of residue K310, and one hydrogen bond to the side chain carboxyl group of D102 and one to the hydroxyl group of Y334.

Ligand pose B (ΔG: -0.77 kcal/mol) forms eight hydrogen bonds with the receptor. There is one bond between the ligand indole amine group and the carboxyl group of residue D102, one hydrogen bond between the ligand side chain carboxyl group and the side chain amine group of residue Q172, two hydrogen bonds from the ligand methylacetate

oxygen atoms to T174 side chain hydroxyl group, and four hydrogen bonds between the ligand side chain methylacetate oxygen atoms and the terminal amine group of residue K310.

4.2.12 Docking results, Tryptophane ethylester , R-form

The results from docking of the R-form of tryptophane ethylester in the 5-HT_{1E} receptor are shown in appendix A2.9a. Figure 5 shows the two-dimensional representation of the ligand docked.

Ligand pose A (ΔG: -6.91 kcal/mol) forms three hydrogen bonds with the $5HT_{1E}$ receptor. One hydrogen bond is formed between the ligand terminal amine moiety and the carboxyl group of residue D102. In addition, the ligand side chain oxygen atoms form one hydrogen bond to the S337 backbone nitrogen atom, and one hydrogen bond to the side chain amine group of residue N336.

Ligand pose B (ΔG: -1.76 kcal/mol) forms six hydrogen bonds with the $5HT_{1E}$ receptor. There are three hydrogen bonds between the ligand amine terminal and the backbone oxygen atom of residue D102 and one hydrogen bond to the carboxyl group of residue D102. Further, two hydrogen bonds from the ligand side chain oxygen group (o1) to the backbone nitrogen atom of residue S337 and the side chain amine group of residue N336, respectively, are formed.

4.2.13 Docking results, Tryptophane ethylester, S-form

The results of docking of the S-form of tryptophane ethylester in the 5-HT_{1E} receptor are shown in appendix A2.9b. Figure 5 shows the two-dimensional representation of the ligand docked.

Ligand pose A (ΔG: -35.88 kcal/mol) forms four hydrogen bonds to the $5HT_{1E}$ receptor. There are three bonds between the ligand terminal amine moiety and residue D102 side

chain carboxyl group. In addition, there is one hydrogen bond between the terminal amine moiety and the hydroxyl group of the Y334 side chain.

Ligand pose B (ΔG: -11.21 kcal/mol) forms seven hydrogen bonds with the receptor. Four hydrogen bonds are formed between the amine terminal and the side chain carboxyl group of residue D102. In addition, there are two hydrogen bonds between the ligand side chain oxygen atoms and the Y334 side chain hydroxyl group and one hydrogen bond to the side chain amine group of residue Q172.

Ligand pose C (ΔG: -10.83 kcal/mol) forms six hydrogen bonds with the receptor. There are two hydrogen bonds between the ligand terminal amine moiety and the carboxyl group and one to the backbone oxygen of residue D102, one hydrogen bond between the ligand indole amine group and the Y334 backbone oxygen atom, as well as two hydrogen bonds from the ligand side chain oxygen atoms to the N336 amine side chain group.

Ligand pose D (ΔG: -9.30 kcal/mol) forms seven hydrogen bonds with the $5HT_{1E}$ receptor. There are two hydrogen bonds between the ligand terminal amine moiety and the carboxyl group of D102, as well as three hydrogen bonds from the amine terminal to the backbone oxygen atom of D102. There are also two hydrogen bonds from the ligand side chain oxygen groups to the side chain nitrogen atom of residue N336.

Ligand pose E (ΔG: -7.03 kcal/mol) forms five hydrogen bonds with the $5HT_{1E}$ receptor. One hydrogen bond is formed between the ligand terminal amine moiety and the carboxyl side chain group of residue D102, whereas two hydrogen bonds are formed between the ligand amine terminal and the backbone oxygen group of residue D102. A hydrogen bond is also formed between the ligand indole amine group and the Y334 main chain oxygen atom and between the ligand side chain double bonded oxygen atom and the side chain amine group of residue N336.

Ligand pose F (ΔG: -6.80 kcal/mol) forms four hydrogen bonds with the receptor; three of the hydrogen bonds are between the ligand terminal amine moiety and the side chain

carboxyl group of residue D102. The last hydrogen bond is formed between the ligand indole amine group and the side chain hydroxyl group of residue S337.

Ligand pose G (ΔG: -4.85 kcal/mol) forms seven hydrogen bonds with the $5HT_{1E}$ receptor. Three hydrogen bonds are formed between the ligand terminal amine and the D102 backbone oxygen atom, whereas one hydrogen bond is formed to the carboxyl side chain group of residue D102. There are two hydrogen bonds between the o1 side chain oxygen of the ligand and the side chain amine group of the N336. In addition, one hydrogen bond between the ligand indole amine group and the backbone oxygen atom of residue V71 is formed.

Ligand pose H (ΔG: -3.18 kcal/mol) forms two hydrogen bonds with the receptor. Both bonds are between the ligand terminal amine moiety and the side chain carboxyl group of residue D102.

Ligand pose I (ΔG: -0.10 kcal/mol) forms six hydrogen bonds with the receptor. Three hydrogen bonds are formed between the ligand terminal amine and the side chain carboxyl group of residue D102 and two hydrogen bonds are formed to the backbone D102 oxygen atom. There is also one hydrogen bond between the o2 side chain oxygen atom of the ligand and the C106 side chain sulphur group.

4.2.14 Docking results, 5-Hydroxy-tryptophane ethylester, R-form

The results from docking of the R-form of 5-hydroxy-tryptophane ethylester in the $5HT_{1E}$ receptor are presented in appendix A2.10a.

Ligand pose A (ΔG: -3.96 kcal/mol) forms eight hydrogen bonds with the receptor. One hydrogen bond is formed between the D102 side chain carboxyl group and three hydrogen bonds between the D102 backbone oxygen atom and the ligand amine terminal, three hydrogen bonds between the ligand side chain oxygen groups and the side chain

nitrogen group N336, and one hydrogen bond between the side chain hydroxyl group of residue S337 and the indole amine group of the ligand.

Ligand pose B (ΔG: -1.37 kcal/mol) forms six hydrogen bonds with the receptor. There are formed two hydrogen bonds between the ligand amine terminal and the D102 side chain carboxyl group and one hydrogen bond to the D102 main chain oxygen atom, as well as one hydrogen bond from the ligand side chain double bonded oxygen atom to the C106 side chain sulphur group, one hydrogen bond to the main chain oxygen atom of M103 from the ligand indole amine group, and one hydrogen bond between the ligand amine terminal and the hydroxyl group of residue Y334.

4.2.15 Docking results, 5-Hydroxy-tryptophane ethylester, S-form

The results from docking of the S-form of 5-hydroxy-tryptophane ethylester in the 5-HT_{1E} receptor are presented in appendix A2.10b.

Ligand pose A (ΔG: -5.65 kcal/mol) forms ten hydrogen bonds with the receptor. There are four bonds between the ligand and residue D102; two are between the ligand terminal amine and the side chain carboxyl group and two are between the ligand terminal amine and the backbone oxygen atom. There is one hydrogen bond between the indole amine group of the ligand and the main chain oxygen group of residue G333, one hydrogen bond between the ligand 5-hydroxyl group and side chain oxygen atom of residue N340, one hydrogen bond between the ligand 5-hydroxyl group and the side chain amine group of residue N336 and one between the ligand side chain oxygen groups and the side chain amine group of residue N336, the main chain nitrogen of residue S337, as well as one to the S109 side chain hydroxyl group.

Ligand pose B (ΔG: -2.76 kcal/mol) forms nine hydrogen bonds with the $5HT_{1E}$ receptor. There are five hydrogen bonds between the ligand amine terminal and the carboxyl group of residue D102, two hydrogen bonds between the ligand 5-hydroxyl group and S109 hydroxyl side chain group and one hydrogen bond to the C105 backbone oxygen group.

Finally, there is one hydrogen bond between the ligand double bonded oxygen side chain atom and the side chain sulphur group of residue C106.

4.2.16 Calculation of binding energy

Tables 3 and 4 show the binding energies that were calculated using the calcBindingEnergy macro in ICM after docking of the ligands in the model of 5-HT_{1E} receptor refined by ICM and in the model of the 5-HT_{1E} receptor refined by both ICM and AMBER. The results show that more ligands are positioned in the putative binding site of the ICM refined 5-HT_{1E} receptor model than in the ICM and AMBER refined model of the 5-HT_{1E} receptor. The accepted ligand pose of each ligand with the best binding energy is named ligand pose A, the second best ligand pose B and so forth.

4.2.17 Minimisation of receptor-ligand complexes

The results from the minimisation of the receptor-ligand complexes in AMBER are not shown, except for the results of the minimisation of the ligand poses with best binding energy of the two ligands where all ligand poses had positive binding energy after docking in ICM. These results are presented in appendix A2.3 and A2.7b. The reason for not showing all results is that problems were encountered during the minimisations in AMBER. For many of the receptor-ligand complexes the minimisation stopped early because of reported problems in the receptor structure far from the putative binding site of the receptor, and there was no time to investigate what was causing these problems.

Table 3. Binding energy, ICM model, ΔG (kcal/mol). A-K denote different ligand poses of each ligand.

	5-HT	Tryptamine	α-Methylserotonin, R	α-Methylserotonin, S
A	-0.94	-3.9	-1.05	-4.07
B	-0.38	-1.97		-3.89
C		-1.94		-1.73
D		-0.44		
	5-Methoxytryptamine	3-(2-Bromoethyl)-indole	3-(2-Hydroxyethyl)-indole	2-Methyl-5-hydroxytryptamine
A	4.44	-2.36	-3.24	-2.99
B		-2	-2.31	
C		-1.68	-1.98	
D			-0.21	
	Tryptophanol, R	Tryptophanol, S	Methyl-3-indoylacetate	Tryptophane ethylester, R
A	4.8	-4.16	-2.22	-6.91
B			-0.77	-1.76
	Tryptophane ethylester, S	5-Hydroxy-tryoptophane ethylester, R	5-Hydroxy-tryoptophane ethylester, S	
A	-35.88	-3.96	-5.65	
B	-11.21	-1.37	-2.76	
C	-10.83			
D	-9.3			
E	-7.03			
F	-6.8			
G	-4.85			
H	-3.18			
I	-0.1			

Table 4. Binding energy, ICM and AMBER refined model, ΔG (kcal/mol). A-E denote different ligand poses of each ligand.

	5-HT	Tryptamine	α-Methylserotonin, R	α-Methylserotonin, S
A	-0.49	-0.56	-1.69	-9.95
B		-0.38	-0.24	-8.63
	5-Methoxytryptamine	3-(2-Bromoethyl)-indole	3-(2-Hydroxyethyl)-indole	2-Methyl-5-hydroxytryptamine
A	-8.14	None	-0.99	-4.64
B	-2.87			
	Tryptophanol, R	Tryptophanol, S	Methyl-3-indoylacetate	Tryptophane ethylester, R
	None	None	None	None
	Tryptophane ethylester, S	5-Hydroxy-tryoptophane ethylester, R	5-Hydroxy-tryoptophane ethylester, S	
A	-21.49	-4.54	-7.31	
B	-8.33		-6.32	
C	-7.33		-4.43	
D	-1.99		-2.38	
E			-0.65	

5. DISCUSSION

5.1 Sequence analysis and receptor structure

The hydropathy plot of the 5-HT_{1E} receptor sequence in figure 6 predicts that there may be at least six hydrophobic stretches corresponding to transmembrane helices in the 5-HT_{1E} receptor protein, as six peaks in the plot have hydropathy scores over 1.8. The plot also indicates two less hydrophobic regions (hydropathy scores of approximately 1) following the sixth peak. These peaks correspond to the seventh transmembrane helix and the intracellular C-terminal helix seen in the structure of the model generated by ICM. The hydropathy plot also predicts that there are two additional hydrophobic stretches in the region between peak 5 and 6, that corresponding to the third, longer intracellular loop. Comparison of the hydropathy plot and the model of the 5-HT_{1E} receptor generated by homology modeling using bovine rhodopsin as a template, shows that ICM predicts two α helices in the third intracellular loop of the 5-HT_{1E} receptor – one ICL3 α helix toward the N-terminal part of the loop consisting of amino acid residues 206-224 and one ICL3 α helix toward the C-terminal end of the loop consisting of amino acid residues 235-245. In figures 10 and 11, where the electrostatic potential of the 5-HT_{1E} receptor model is shown, these two helices may at least account for some of the electronegatively charged areas seen in this otherwise electropositive area.

The results of the hydropathy plot are in accordance with the results obtained from the alignment of the 5-HT_{1E} receptor sequence with that of bovine rhodopsin, shown in figure 7, and the model built based on this alignment, shown in figures 9-11. The alignment of rhodopsin and the 5-HT_{1E} receptor in ICM predicts that helix 1 is comprised of residues 19-49, helix 2 of residues 56-85, helix 3 of residues 91-124, helix 4 of residues 139-157, helix 5 of residues 179-203, helix 6 of residues 281-315, and helix 7 of residues 323-346, followed by a helix consisting of amino acid residues 348-359. In this last helix, residue F350, which corresponds to residue F313 in rhodopsin, is highly conserved within family A of GPCRs.

The alignment of the 5-HT_{1E} receptor and the sequence of bovine rhodopsin in ICM reveals that the two proteins share 14 percent sequence homology. Normally, such low sequence homology between the template and target structure would be detrimental for the prediction of the target structure by homology modeling, as it is believed that the template and target sequences should have at least 50 percent of their residues in common in order to obtain accurate models. Membrane proteins, however, despite low sequence homology, fold into similar three-dimensional structures. Functional important residues, which are important for proper folding of the structure or involved in the signalling mechanisms of the receptor, are highly conserved among such membrane proteins and may be used to classify the different proteins. The highly conserved residues among family A GPCRs that place the 5-HT_{1E} receptor in this family are residues N40 in TMH1, D68 in TMH2, R120 in TMH3, W147 in TMH4 and residues P194, P306, and P341 of TMH5-7, respectively. Further, the presence of the biogenic receptor fingerprint residues D3.32 (D102), W7.40 (W331) and Y7.43 (Y334) in the sequence shows that this sequence indeed encodes a biogenic amine receptor, as no other family A GPCR contains these three residues. In addition, residues F6.51 (F307) and F6.52 (F308) are only found in mammalian 5-HT receptors and show that the sequence encodes a G-protein coupled 5-HT receptor.

As the alignment of the 5-HT_{1E} receptor and rhodopsin amino acid sequences in figure 7 shows, there is no sequence homology between the two sequences in the N-terminal, extracellular parts of the proteins. The 5-HT_{1E} receptor and bovine rhodopsin also share low sequence homology in the first transmembrane helix, though many residues are similar in that they are hydrophobic. Transmembrane helix 2 shows a somewhat more sequence homology between the sequences than the first transmembrane helix; however, rhodopsin is not a biogenic amine receptor and does not have an aspartate residue in position 3.32 but an alanine residue. As the negatively charged side chain carboxyl group of the D3.32 residue is thought to form an ionic interaction with the positively charged biogenic amine ligands, it is not optimal that these residues differ in the 5-HT receptors and rhodopsin.

The 5-HT_{1E} receptor forms a disulfide bond between cysteine residue 95, which is located close to the extracellular side in transmembrane helix 3, and cysteine residue 173, found in the β strand that is located toward the C-terminal end of the second extracellular loop. This disulfide bridge is highly conserved and occurs in rhodopsin as well as in all the 5-HT receptors. Disulfide bonds in general are though to stabilise protein three-dimensional structures and in family A GPCRs, this disulfide bridge covalently attaches the second extracellular loop in the physical proximity of the binding site. Figure 9 shows the location of this loop relative to D102 (D3.32) in the 5-HT_{1E} receptor. The 5-HT_{1E} receptor has three consecutive proline residues (P168, P169, and P170) in the loop between the β strands, which none of the other 5-HT receptors nor rhodopsin have. Rhodopsin and the 5-HT_{2B} receptor have a proline residue at the first position (residue P180 in the rhodopsin structure and P202 in the 5-HT_{2B} receptor). Two consecutive proline residues is known to disrupt secondary structure elements.

Residues D119, R120, and Y121 of the 5-HT_{1E} receptor constitute the conserved DRY motif which is located toward the intracellular end of the third transmembrane helix. R120 is the reference residue of transmembrane helix 3. This motif is highly conserved among family A members (Kroeze et al. 2002) and is found in all 5-HT receptors. In rhodopsin, the aspartate residue is a glutamic acid. Studies on rat 5-HT_{2A} receptors suggest that residue R3.50 may be involved in a strong ionic interaction with residue E6.30, which is residue E286 in the 5-HT_{1E} receptor structure, when the receptor is in its inactive state and that disturbance of this interaction may cause transmembrane helices 3 and 6 to move apart, thereby activating the receptor. In the 5-HT_{1E} receptor structure, the side chains of residues R120 and E286 are pointing toward each other, suggesting that this receptor activation theory may apply for the 5-HT_{1E} receptor as well. Figure 9 shows the location of R120 and E286 in the 5-HT_{1E} receptor.

Transmembrane helix 6 contains a highly conserved region of amino acids FXXXW, where residues X vary among the receptors. In the 5-HT_{1E} receptor sequence, this region corresponds to residues F300 (6.44), I301 (6.45), L302 (6.46), S303 (6.47), and W304 (6.48); the 5-HT_{1E} receptor is the only 5-HT receptor that has a serine residue in position

6.47. Rhodopsin has residues L, I, and C in position 6.45-6.47. The x-ray structure of rhodopsin reveals that W6.48 (which is residue 265 in rhodopsin sequence) is in close proximity of the retinal chromophore and is involved in the activation and release of the chromophore (Okada et al. 2004). In the modelled structure of the 5-HT_{1E} receptor generated in this project, the tryptophan in this position is in close proximity of the suggested ligand binding site and may contribute to stacking of the ligands in the binding site

Transmembrane helix 7 contains a conserved NPXXY region. In the 5-HT_{1E} receptor this region corresponds to residues N340 (7.49), P341 (7.50), L 342 (7.51), L343 (7.52), and Y344 (7.53). Residues X vary between being leucine, isoleucine, and phenylalanine in the 5-HT and rhodopsin receptor sequences. It has been suggested that the aspargine in this motif forms hydrogen bonds with the reference residue of transmembrane helix two, D68, and that the residues may be important for receptor activation. The two residues are located close to each other in the structure of the 5-HT_{1E} receptor (not shown).

The extracellular and intracellular loops of the 5-HT receptors are the least conserved parts of the receptor. It has been hypothesised that the ligands must pass through a channel made up of residues in the first and third extracellular loops and the extracellular parts of the transmembrane helices in order to reach their binding site buried between the helices (Kristiansen 2004). The residue corresponding to W88 in the 5-HT_{1E} receptor sequence in the first extracellular loop is conserved among all the 5-HT receptors. The residue in this position in rhodopsin is another aromatic residue, namely phenylalanine. Except for this conserved residue, the sequence homology in the first extracellular loop is low. The second extracellular loop contains the conserved disulfide bridge that contributes in placing the second extracellular loop in the proximity of the receptor ligand binding site and makes it possible for loop residues to contact the ligands. The 5-HT_{1E} receptor contains a stretch of residues in this loop that is not found in any of the other 5-HT receptor sequences, namely residues S162, H163, R164, R165, L166 and S167. Residues R164-S167 constitute the N-terminal β strand of this loop.

Of the intracellular loops, the third intracellular loop is the longest. In the 5-HT_{1E} receptor structure this loop consists of residues 204-280. This loops most likely is the site of G-protein coupling.

The crystal structure of rhodopsin shows that rhodopsin has a helix 8 in the C-terminal intracellular part of the sequence (Okada et al. 2004). This helix corresponds to residues 348-361 in the 5-HT_{1E} receptor.

Figure 12 shows the Ramachandran plot of the 5-HT_{1E} receptor refined by ICM. The Ramachandran plot of the both 5-HT_{1E} receptor models showed that the same residues were outside their allowed regions in both plots, showing that the molecular dynamics simulation did not introduce additional unfortunate angles in the structure. The Ramachandran plot shows that there are few residue that have psi and phi angles outside their allowed region in the receptor, and that the residues that are have angles outside the allowed regions not located in the conserved heptahelical structure of the receptors, except for F191, which is located in the pi helix of transmembrane helix.

Figures 10 and 11 shows the 5-HT_{1E} receptor electrostatic potential. The electrostatic potentials of the model indicate that the receptor has a dipolar charge distribution with the extracellular side mainly negative and the intracellular side mainly positive. The plasma membrane of cells separates positive and negative charges; the extracellular side has a relative excess of positive charge and the intracellular side has a relative excess of negative charge. The charge distribution over the membrane will therefore stabilise the structure of the 5-HT receptor in the membrane. This finding is in accordance with the ‘positive inside rule’ that states that all membrane proteins are positively charged toward the intracellular side and negatively charged regions toward the extracellular side.

5.2 Results from docking of ligands in the 5-HT_{1E} receptor

5.2.1 Docking in the 5-HT_{1E} receptor model refined by ICM and AMBER

The ligands shown in figure 5 were docked in the model of the 5-HT_{1E} receptor that had been refined by both ICM and AMBER. The results from this docking are shown in table 4 in form of binding energies, while the binding energies from docking in the 5-HT_{1E} ICM refined model are shown in table 3.

The results from docking in the 5-HT_{1E} receptor model refined by ICM and AMBER show that fewer ligands fitted into the putative binding site of the receptor compared to in the 5-HT_{1E} receptor model refined only by ICM. Comparison of the results from the two dockings shows that some of the binding energies for the ligands are in the same range in both receptor models but some differ quite much. Some ligands, for example 3-(2-Bromoethyl)-indole and both forms of tryptophanol, show no binding in the putative binding area of the model refined by both ICM and AMBER, whereas they are positioned in the binding area of the 5-HT_{1E} receptor model refined only by ICM.

During MD simulation in AMBER the helices were restrained in Cartesian space. Comparison of the two models of the 5-HT_{1E} receptor (not shown) showed that the helices had not moved much during the MD simulation, as expected, indicating a stable receptor structure. However, the loops of the model have moved relatively much and this, together with some minor movement of the residue side chains, may explain the poor docking results obtained when performing docking in the 5-HT_{1E} receptor model refined by both ICM and AMBER. As discussed earlier, the extracellular loops, especially the second extracellular loop, are important for ligand recognition and this may therefore explain the poorer results from docking of the ligands in this model than in the model refined only by ICM.

As the refineModel macro in ICM is considered to be enough refinement of a model built using the homology modeling approach in order to obtain good molecular docking

results, and the results after docking in the ICM and AMBER refined model were poorer than the results from docking in the ICM refined model, the 5-HT_{1E} receptor model refined by ICM was chosen to be used for interpretation of the results obtained in this project.

5.2.2 Docking in the 5-HT_{1E} receptor model refined by ICM

The results from docking of 5-HT in the putative binding site of the 5-HT_{1E} receptor are shown in figure 13, while figures showing the results from docking of the other ligands are shown in appendix 2. The calculated binding energies from the docking are contained in table 3. The figures and table only show the results where the ligands were positioned in the putative binding area and had negative binding energy, except for two ligands, α-methylserotonin and the R-form of tryptophanol, for which no ligand poses positioned in the binding area having negative binding energy were found. Instead, the ligand pose with best positive binding energy of each of these two ligands are presented in appendix 2.

The results show that no ligand poses were located in close proximity of residues S186, T187, F191 or Y192, which are suggested to be potential hydrogen bonding residues (S186 and T187) and part of a hydrophobic ligand binding pocket (F191 and Y192). In the model, T187 and Y192 are pointing away from the ligand binding pocket toward the membrane. It may be that F191 is pointing in the wrong direction and might actually be more important than shown here. The Ramachandran plot showed that this residue had psi and phi angles outside the allowed region so it is possible that F191 in reality is more involved in ligand binding than shown in this docking study. Also, residue D327, which is located in position 7.36, does not either seem to be directly involved in ligand binding.

Figure 8 shows the putative ligand binding area of the 5-HT_{1E} receptor created by displaying the molecular surfaces of four ligand poses, 5-HT pose A, tryptamine pose B and D, and pose A of the S-form of 5-hydroxy-tryptophane ethylester. These ligand poses were chosen as they represent the areas where most ligands were positioned in this study.

5-HT ligand pose A is positioned such that the aromatic tyrosine reside 334 side chain stacks against the ligand indole ring in a sandwich fashion and the ligand terminal amine moiety is in close proximity of the carboxyl oxygen atoms of the D102 side chain, see figure 13. The indole ring of the tryptamine ligand pose B is positioned such that it is almost perpendicular to the aromatic side chain of residue W98 and not far from Y334, whereas ligand pose D is positioned in a completely other orientation, wher the indole ring is positioned in the proximity of the aromatic side chains of residues F300, F307, and F308. The results show that the side chains of these aromatic residues may be involved in stacking of the ligands. The terminal amine moieties of both ligand poses are in close proximity of the carboxyl group of residue D102 and form hydrogen bonds to this residue. The last ligand pose representing an area where ligands can interact with the receptor putative binding site of the 5-HT_{1E} receptor, is ligand pose A of the 5-hydroxyl-tryptophane ethylester ligand. The indole ring of this ligand pose is positioned in such a way that the aromatic side chain of residue W304 may be involved in stacking of this ligand pose.

Most of the ligand poses are positioned in the same area of the putative binding site as ligand poses A of 5-HT and ligand pose B of tryptamine. Fewer ligands poses are found in the area represented by ligand pose D of tryptamine and only ligand pose A of the 5-hydroxyl-tryptophane ethylester ligand is positioned in that area.

The two ligands where all ligand poses in the binding site have positive binding energy, 5-methyltryptamine and the R-form of tryptophanol, are not located in the same area of the binding site. Appendix A2.3 shows that the ligand pose of 5-methyltryptamine is located in the same area as ligand pose B of tryptamine and that the indole rings of these two ligand poses are positioned approximately in the same orientation. The receptor-5-methoxytryptamine complex was minimised using the SANDER program of AMBER and comparison of the receptor-ligand complex before AMBER minimisation and after AMBER minimisation, see appendix A2.3, shows that the major difference before and after the minimisation are found in the ligand structure. The receptor residues show minor differences in their positions before and after the energy minimisation. The second ligand

where all ligand poses had positive binding energy was the R-form of tryptophanol, see appendix A2.7a. The indole ring of this ligand is located in the approximate same position as tryptamine ligand pose B. Minimisation of the receptor-R-tryptophanol complex resulted in negative binding energy of the complex. Only one true ligand pose of the S-form of tryptophanol was observed. This ligand pose show relatively good binding energy when compared to the binding energies of the poses of the other ligands docked (ΔG of -4.16 kcal/mol).

Of the receptor residues that the ligands form hydrogen bonds to, three are located in the second extracellular loop, namely residues Q172, C173, and T174. Residue C173 forms the disulfide bridge (to residue C95) and is conserved but the other two residues are not. Only rhodopsin and the 5-HT_{1E} and 5-HT_6 receptors have a glutamine residue in the position corresponding to Q172 in the 5-HT_{1E} receptor. The side chain of this residue is involved in forming hydrogen bonds to some of the ligands. The 5-HT_{1E} and the 5-HT_{1A} receptor are the only receptors in the alignment that have threonine in the position corresponding to T174 in the 5-HT_{1E} receptor. The side chain of this residue is involved in forming hydrogen bonds to the ligands. Further, the 5-HT_{1E} receptor is the only receptor in the alignment that has a methionine residue in position 103. However, only one ligand forms hydrogen bonds to this residue and the bond is formed to the main chain not the side chain of the residue suggesting that it does not play an important role in ligand binding. Only one ligand forms hydrogen bonds to residue K310, which has been suggested may account for the pharmacological differences observed between the 5-HT_{1E} and 5-HT_{1B} receptors, namely methyl-3-indoylacetate. This ligand also forms hydrogen bonds to Q172 and T174.

Another observation is that many of the ligand poses form many hydrogen bonds to the receptor residues. Lipinski's rule-of-5 states that drugs that have the possibility of participating in many hydrogen bonds may not be easily absorbed, distributed, metabolised, or excreted in the body. The results from this docking study show that many of the ligands tested participate in many hydrogen bonds and may for that reason not be suitable drug compounds.

The calculation of binding energy shows that 5-HT, except for the two ligands having positive binding energy, is the ligand that has the lowest binding energy, the best ligand pose of 5-HT having a binding energy of -0.94 kcal/mol. 5-HT and the R-form of α-methylserotonin are the only ligands having binding energies higher than -2 kcal/mol. The best ligand poses of ligands 3-(2-bromoethyl)-indole, 3-(2-hydroxyethyl)-indole, 2-methyl-5-hyroxytryptamine and methyl-3-indoylacetate have binding energies in the range of -2 to -3 kcal/mol, whereas the best poses of ligands tryptamine, the S-form of α-methylserotonin, the S-form of tryptophanol, and the 5-hydroxy-tryptophane ethylester R-form all have binding affinities of approximately -3 to -4 kcal/mol. Two ligands, the S-form of 5-hydroxy-tryptophane ethylester and the R-form of tryptophane ethylester, have binding energies between approximately -5 and –7 kcal/mol, whereas the best ligand pose of the S-form of tryptophane ethylester has a binding energy of -35.88 kcal/mol. However, since all other binding energies are much higher than this value, the result is probably not correct.

5.2.3 Experimental ligand binding study

Appendix 3 contains the results from a radio ligand binding study performed by Lise Roman Moltzau at the Department of Pharmacology at the University of Oslo. The radio ligand binding study performed was a competition binding assay with the ligand and 5-HT_{1E} receptor expressed in HEK293 cells. Using unlabeled ligands that competed with [^{3}H]5-HT the affinities for the ligands were determined by displacement of [^{3}H]5-HT. The ligands were not synthesised at the University of Oslo but bought commercially. None of the ligands are very good – they neither have high binding affinity nor high specificity for the 5-HT_{1E} receptor.

The results from this radio ligand binding study do not fit the results from the docking study performed here. The results from the ligand binding study show that 5-HT is the ligand that has best affinity for the 5-HT_{1E} receptor, whereas the results obtained here show that many of the ligands have better binding energy than 5-HT. It further shows that

four ligands, 3-(2-bromoethyl)-indole, 3-(2-hydroxyethyl)-indole, methyl-3-indoylacetate, D,L-tryptophan ethylester hydrochloride, and L-tryptophan ethylester hydrochloride have negative pKi values and most likely do not bind to the 5-HT_{1E} receptor, whereas the results obtained in this docking study give relatively good binding energies for these ligands, in fact, the two forms of tryptophan ethylester are the ligands that have the best binding energies in the docking study. 5-methoxytryptamine and the R-form of tryptophanol, the ligands where no ligand poses with negative binding energies were found in the docking study, are shown to bind to the 5-HT_{1E} receptor in the radio ligand binding study.

The ligands were not synthesised in Oslo but bought commercially. None of the ligands are very good – they neither have high binding affinity nor high specificity for the 5-HT_{1E} receptor.

The model of the 5-HT_{1E} receptor may not be optimal. The alignment of the 5-HT_{1E} receptor and rhodopsin amino acid sequences may not be completely correct and as residues that are important for ligand binding in the 5-HT_{1E} receptor, such as D3.32 and Y7.43 are not conserved in rhodopsin, the side chains of these residues may be pointing in the wrong direction in the modelled structure. Rhodopsin has an alanine residue in position 3.32 and a lysine residue in position 7.43. Another important reason may be that the receptor molecule is rigid during docking in ICM. The protein flexibility is considered to be one of the major challenges in molecular docking; ideally protein flexibility should be taken into account but of today, this is not possible in ICM.

6. LITTERATURE

Bai F., Yin T., Johnstone E. M., Su C., Varga G., Little S. P. and Nelson D. L. (2004) Molecular cloning and pharmacological characterization of the guinea pig 5-HT1E receptor. *Eur J Pharmacol* **484,** 127-139.

Berman H. M., Westbrook J., Feng Z., Gilliland G., Bhat T. N., Weissig H., Shindyalov I. N. and Bourne P. E. (2000) The Protein Data Bank. *Nucleic Acids Res* **28,** 235-242.

Bockaert J., Claeysen S., Becamel C., Dumuis A. and Marin P. (2006) Neuronal 5-HT metabotropic receptors: fine-tuning of their structure, signaling, and roles in synaptic modulation. *Cell Tissue Res* **326,** 553-572.

Branden C. T., J. (1999) *Introduction to Protein Structure*. Garland Publishing, Inc., New York.

Cabrera-Vera T. M., Vanhauwe J., Thomas T. O., Medkova M., Preininger A., Mazzoni M. R. and Hamm H. E. (2003) Insights into G protein structure, function, and regulation. *Endocr Rev* **24,** 765-781.

Jakalian A., Jack D. B. and Bayly C. I. (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem* **23,** 1623-1641.

Kristiansen K. (2004) Molecular mechanisms 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.

Kroeze W. K., Kristiansen K. and Roth B. L. (2002) Molecular biology of serotonin receptors structure and function at the molecular level. *Curr Top Med Chem* **2,** 507-528.

Levy F. O., Gudermann T., Birnbaumer M., Kaumann A. J. and Birnbaumer L. (1992) Molecular cloning of a human gene (S31) encoding a novel serotonin receptor mediating inhibition of adenylyl cyclase. *FEBS Lett* **296,** 201-206.

McAllister G., Charlesworth A., Snodin C., Beer M. S., Noble A. J., Middlemiss D. N., Iversen L. L. and Whiting P. (1992) Molecular cloning of a serotonin receptor from human brain (5HT1E): a fifth 5HT1-like subtype. *Proc Natl Acad Sci U S A* **89,** 5517-5521.

Mirzadegan T., Benko G., Filipek S. and Palczewski K. (2003) Sequence analyses of G-protein-coupled receptors: similarities to rhodopsin. *Biochemistry* **42,** 2759-2767.

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

Rang H. P. D., M.M.; Ritter, J.M.; Moore, P.K. (2003) *Pharmacology*. Churchill Livingstone, Edinburgh.

Ravna A. W. (2006) Three-dimensional models of neurotransmitter transporters and their interactions with cocaine and S-citalopram. *World J Biol Psychiatry* **7,** 99-109.

Raymond J. R., Mukhin Y. V., Gelasco A., Turner J., Collinsworth G., Gettys T. W., Grewal J. S. and Garnovskaya M. N. (2001) Multiplicity of mechanisms of serotonin receptor signal transduction. *Pharmacol Ther* **92,** 179-212.

Santarelli L., Saxe M., Gross C., Surget A., Battaglia F., Dulawa S., Weisstaub N., Lee J., Duman R., Arancio O., Belzung C. and Hen R. (2003) Requirement of hippocampal neurogenesis for the behavioral effects of antidepressants. *Science* **301,** 805-809.

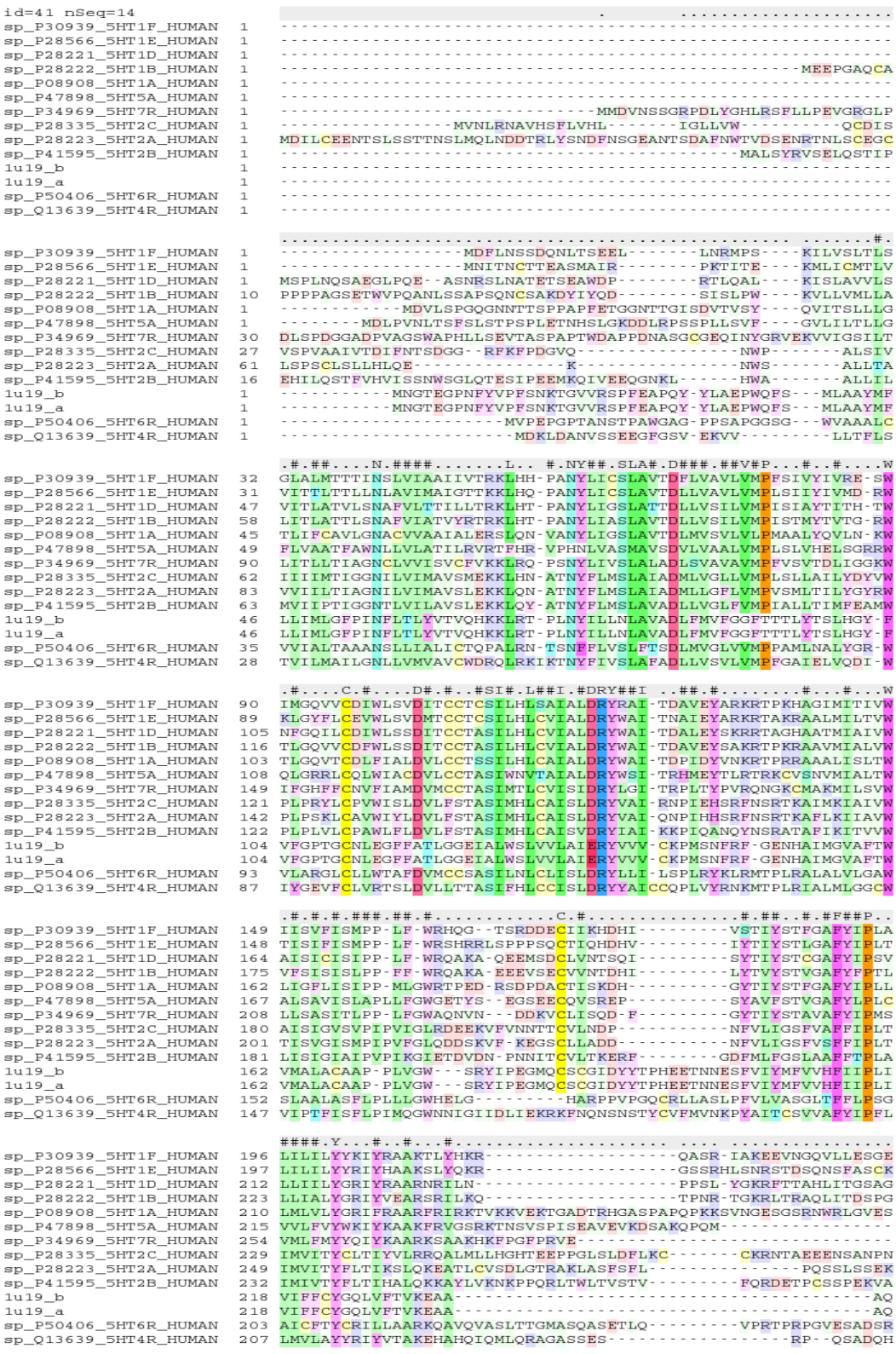
Sousa S. F., Fernandes P. A. and Ramos M. J. (2006) Protein-ligand docking: current status and future challenges. *Proteins* **65,** 15-26.

Wang J., Wolf R. M., Caldwell J. W., Kollman P. A. and Case D. A. (2004) Development and testing of a general amber force field. *J Comput Chem* **25,** 1157-1174.

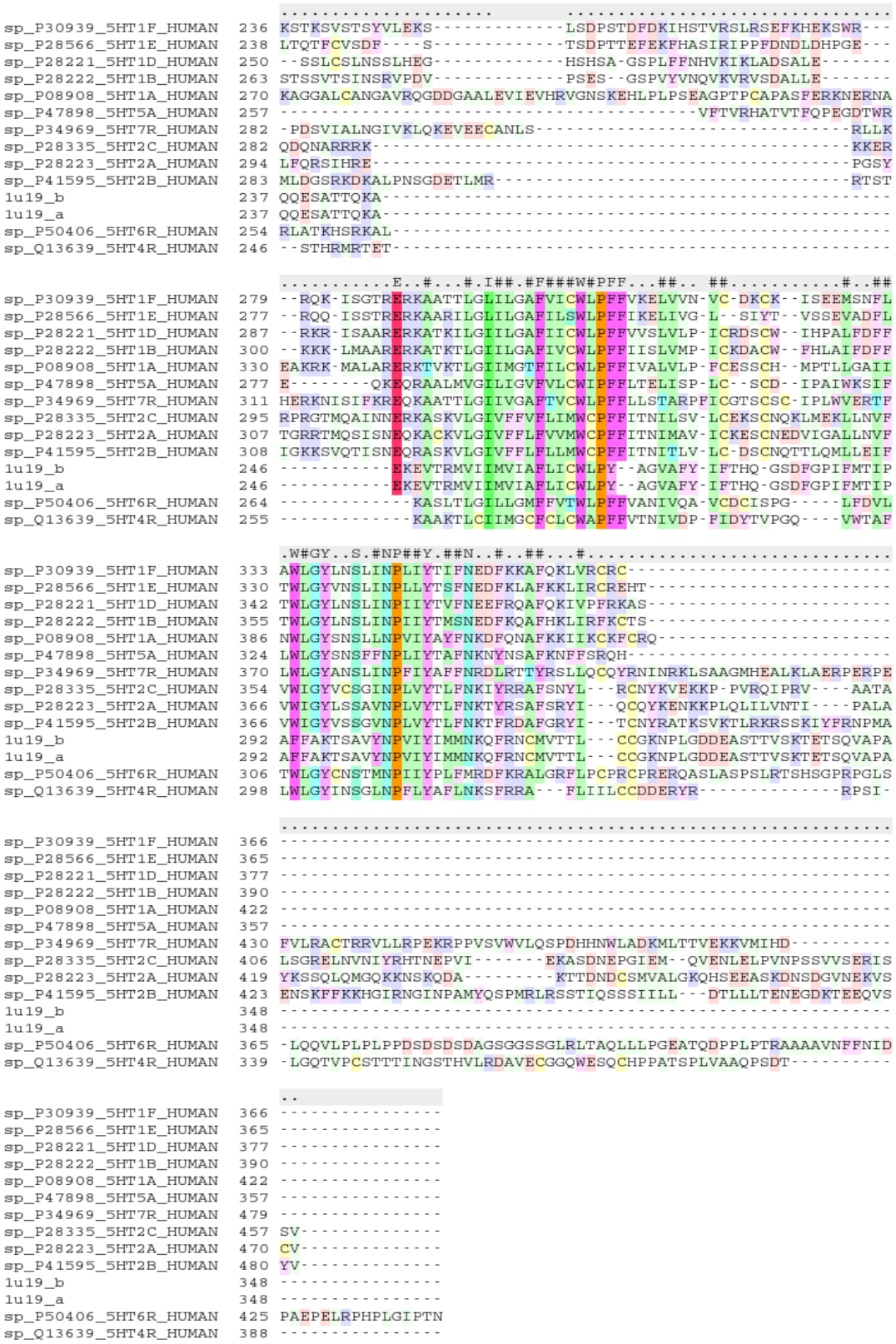
Zgombick J. M., Schechter L. E., Macchi M., Hartig P. R., Branchek T. A. and Weinshank R. L. (1992) Human gene S31 encodes the pharmacologically defined serotonin 5-hydroxytryptamine1E receptor. *Mol Pharmacol* **42,** 180-185.

APPENDIX A1

Alignment of rhodopsin and 5-HT receptor amino acid sequences



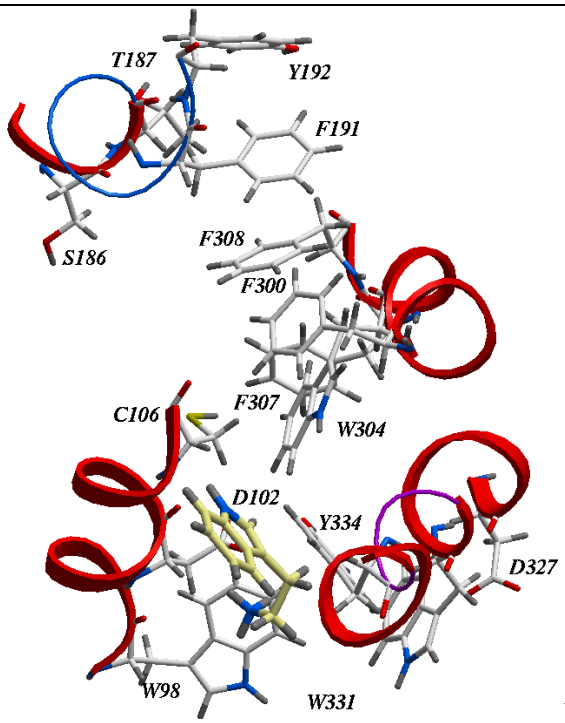
Appendix A1. Alignment of the human 5-HT receptor and bovine rhodopsin amino acid sequences generated by ICM.



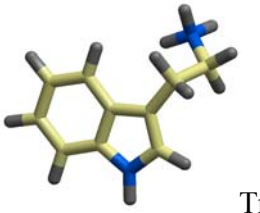
Appendix A1 continued. Alignment of the human 5-HT receptor and bovine rhodopsin amino acid sequences generated by ICM.

APPENDIX A2

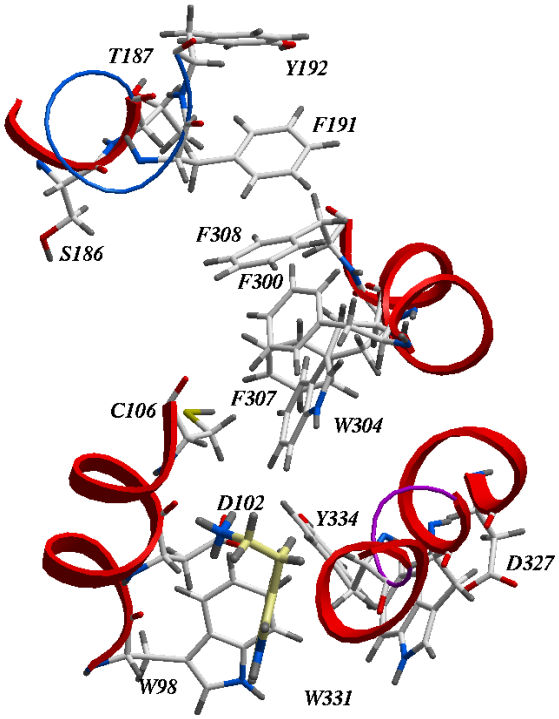
A2.1 Tryptamine



Ligand pose A

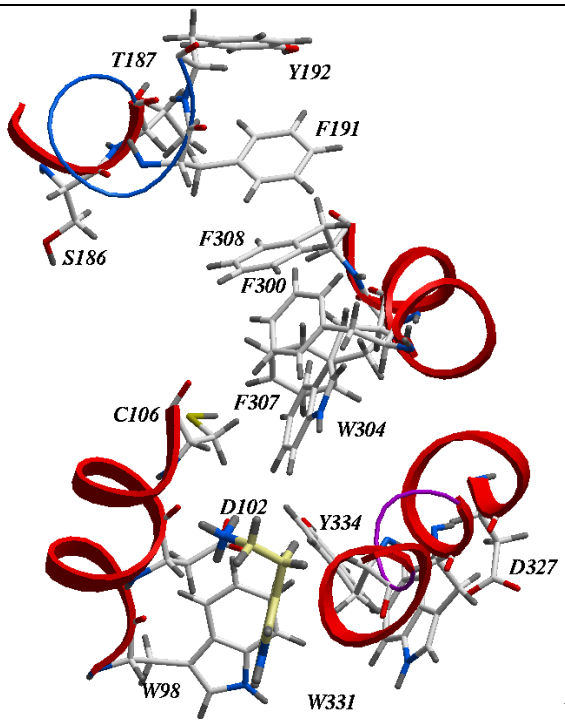


Tryptamine

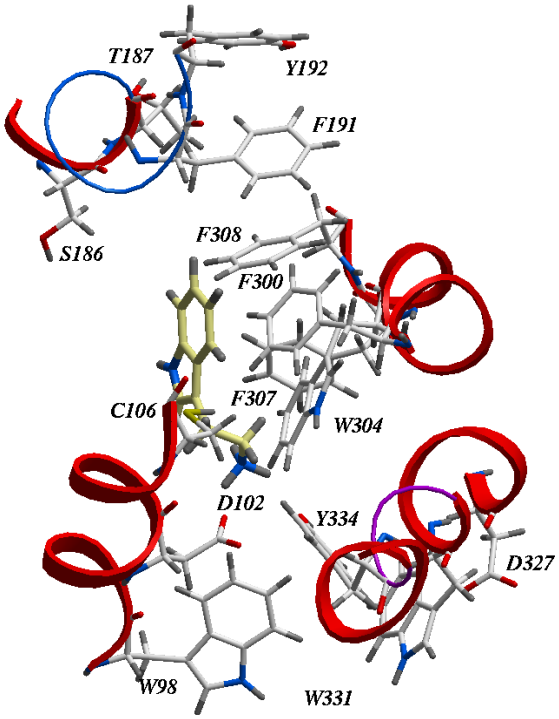


Ligand pose B

Appendix A2.1. Ligand poses of tryptamine in the putative binding site of the 5-HT_{1E} receptor.



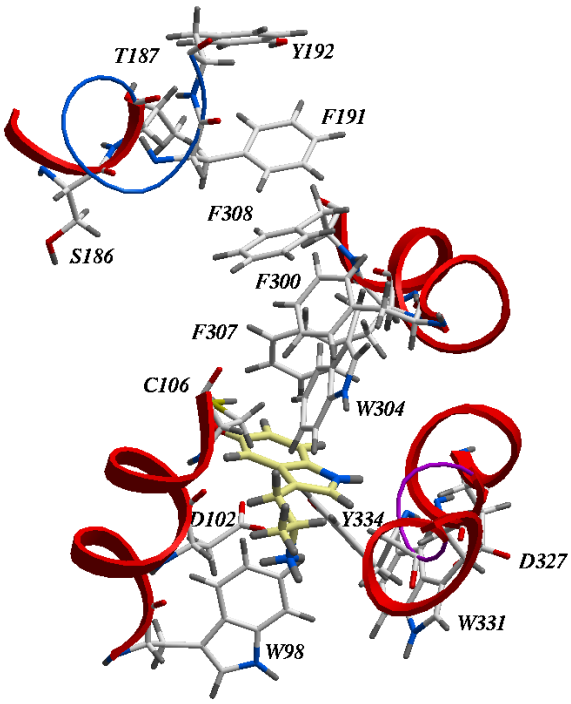
Ligand pose C

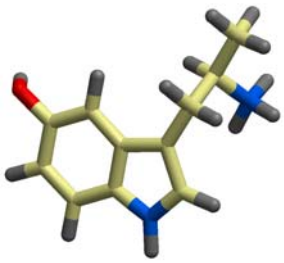


Ligand pose D

Appendix A2.1 continued. Ligand poses of tryptamine in the putative binding site of the 5-HT_{1E} receptor.

A2.2.α-Methylserotonin

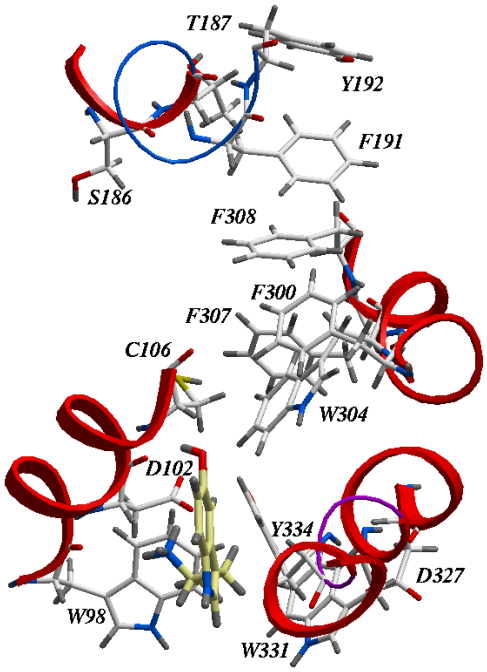




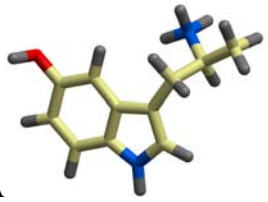
Ligand pose A

α-Methylserotonin, R-form

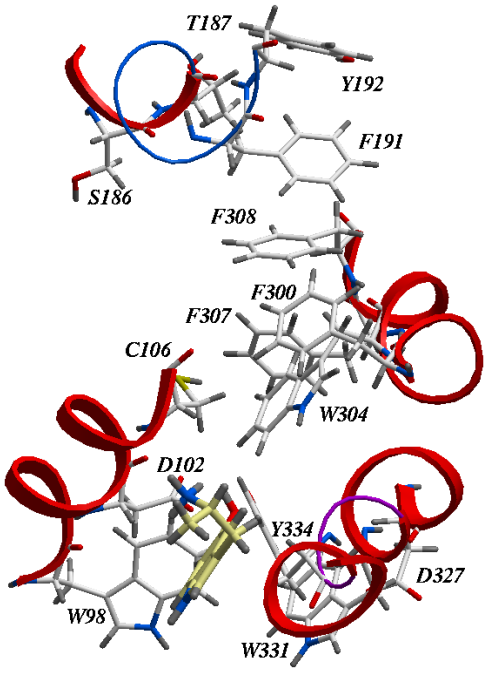
Appendix A2.2a Ligand pose A of α-Methylserotonin, R-form, in the putative binding site of the 5-HT_{1E} receptor.



Ligand pose A

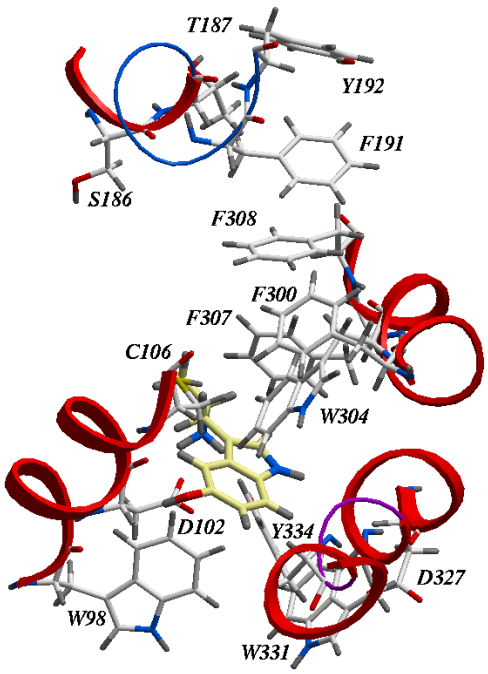


α-Methylserotonin, S-form



Ligand pose B

Appendix A2.2b Ligand poses of α-Methylserotonin, S-form, in the putative binding site of the 5-HT_{1E} receptor.

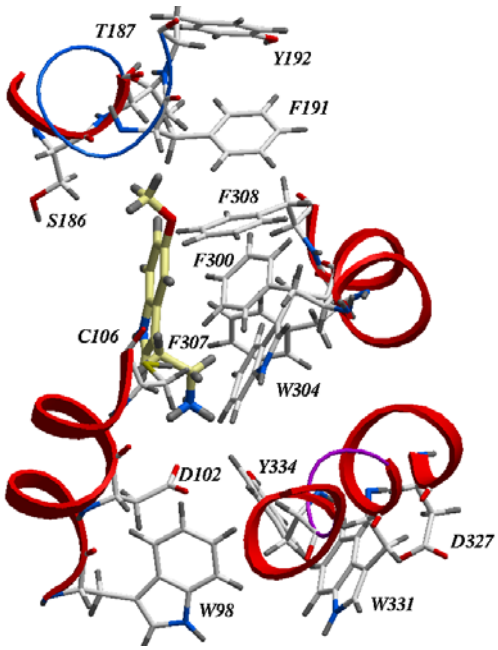


Ligand pose C

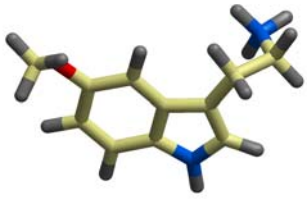
Appendix A2.2b continued. Ligand poses of α-Methylserotonin, S-form, in the putative binding site of the 5-HT_{1E} receptor.

A2.3 5-Methoxytryptamine

I

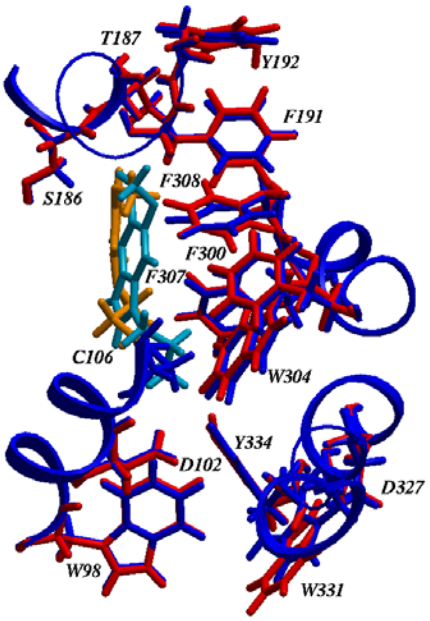


Ligand pose A



5-Methoxytryptamine

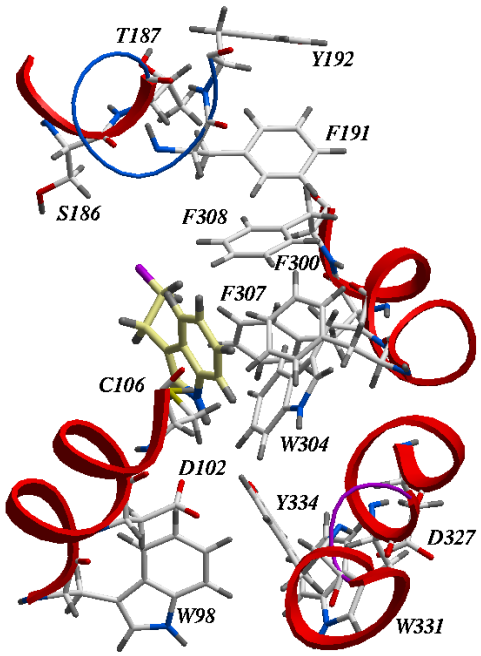
II



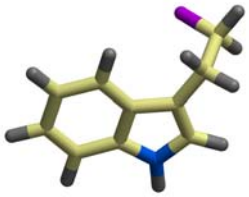
Ligand pose A; ICM: blue and AMBER: red

Appendix A2.3. 5-Methoxytryptamine in the putative binding site of the 5-HT_{1E} receptor. Figure I shows the ligand pose yielding positive binding energy (ΔG: 4.44 kcal/mol). Figure II shows the superimposition of the receptor-ligand complex before (coloured blue/light blue) and after (coloured red/orange) energy minimisation of the complex in AMBER.

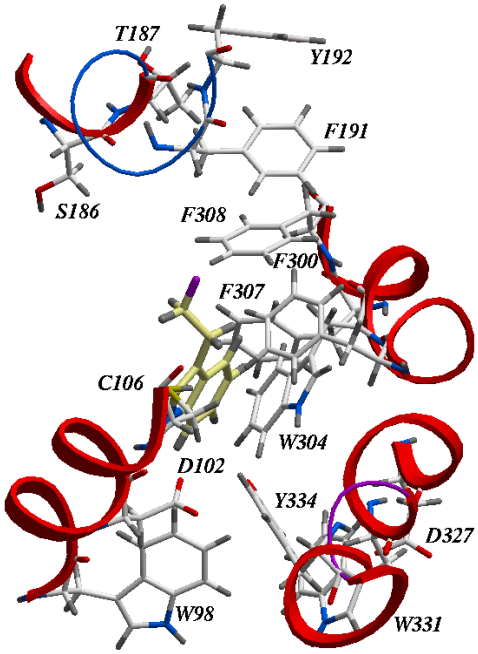
A2.4 3-(2-Bromoethyl)-indole



Ligand pose A

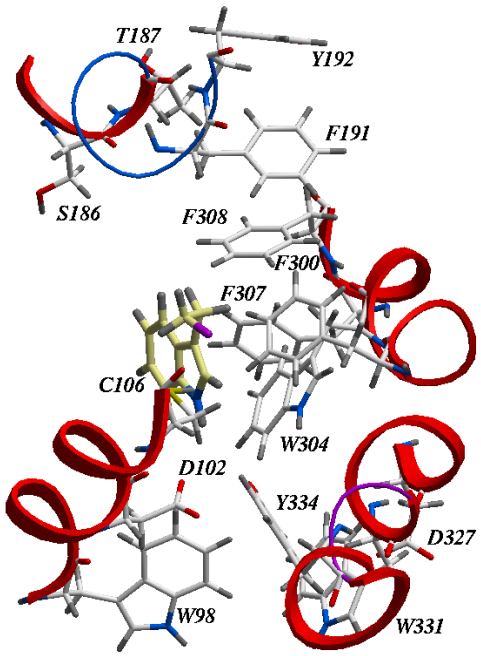


3-(2-Bromoethyl)-indole



Ligand pose B

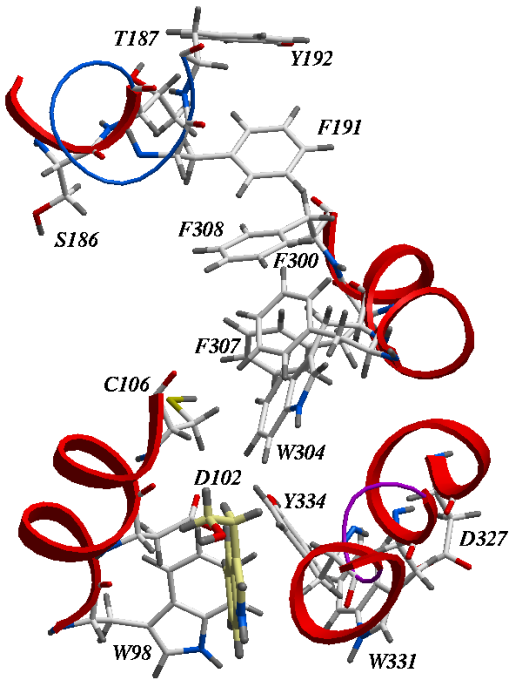
Appendix A2.4. Ligand poses of 3-(2-Bromoethyl)-indole in the putative binding site of the 5-HT_{1E} receptor.

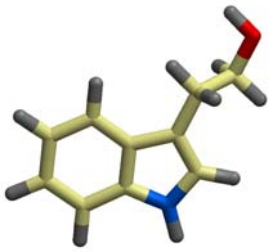


Ligand pose C

Appendix A2.4 continued. Ligand poses of 3-(2-Bromoethyl)-indole in the putative binding site of the 5-HT_{1E} receptor.

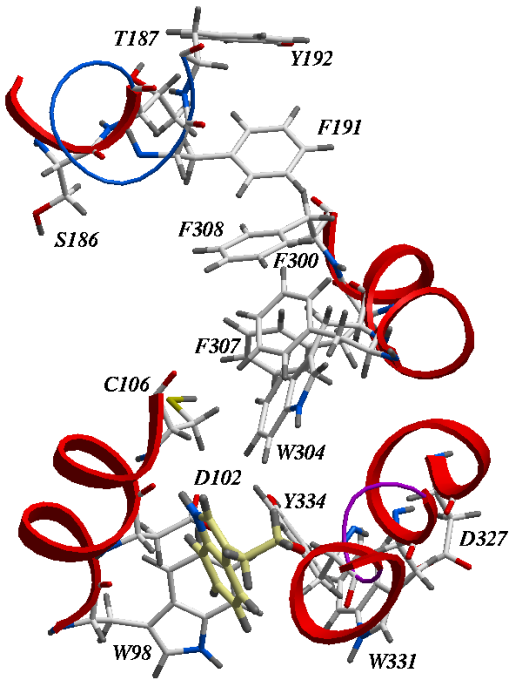
A2.5 3-(2-Hydroxyethyl)-indole





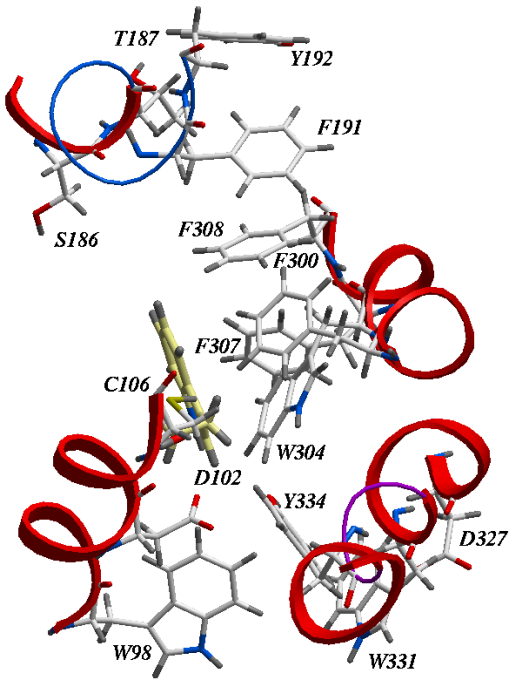
Ligand pose A

3-(2-Hydroxyethyl)-indole

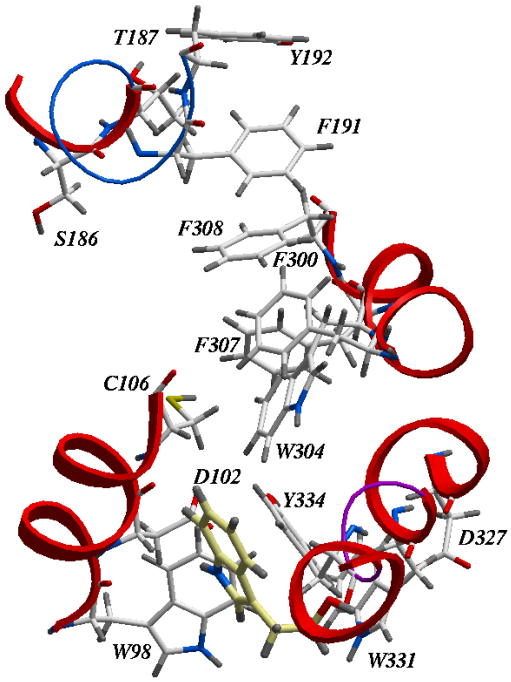


Ligand pose B

Appendix A2.5. Ligand poses of 3-(2-Hydroxyethyl)-indole in the putative binding site of the 5-HT_{1E} receptor.



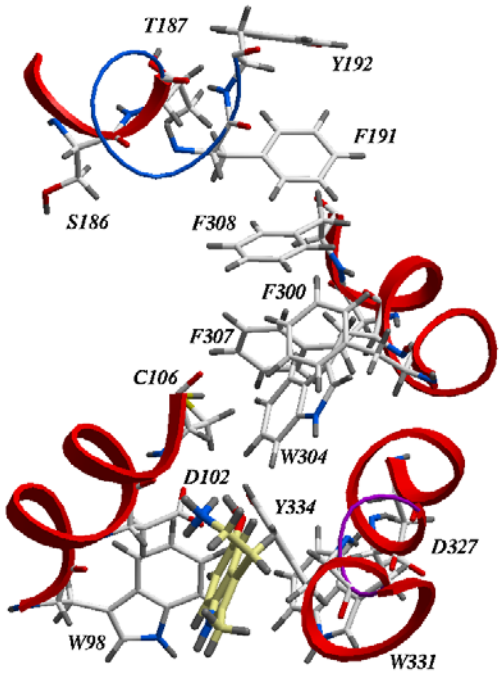
Ligand pose C



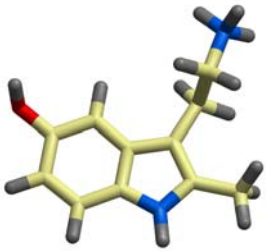
Ligand pose D

Appendix A2.5 continued. Ligand poses of 3-(2-hydroxyethyl)-indole in the putative binding site of the 5-HT_{1E} receptor.

A2.6 2-Methyl-5-hydroxytryptamine



Ligand pose A

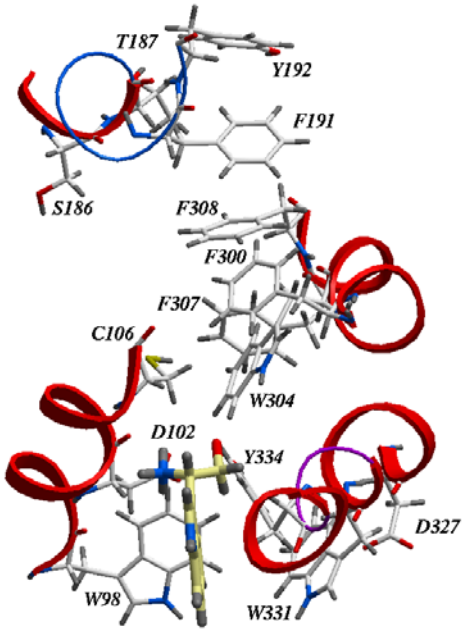


2-Methyl-5-hydroxytryptamine

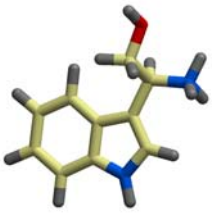
Appendix A2.6. Ligand pose of 2-Methyl-5-hydroxytryptamine in the putative binding site of the 5-HT_{1E} receptor.

A2.7 Tryptophanol

I

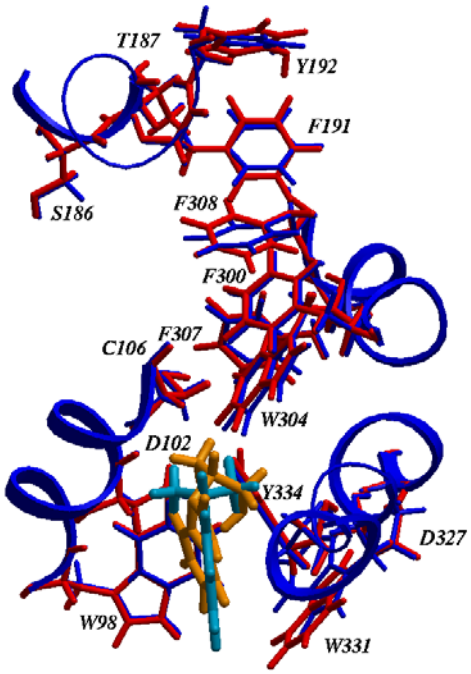


Ligand pose A



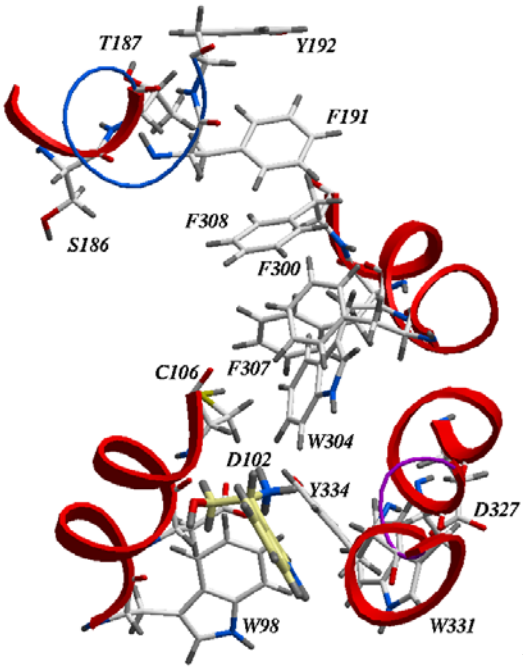
Tryptophanol, R-form

II

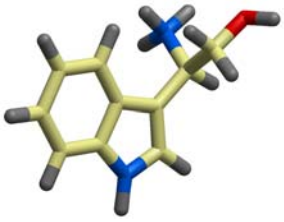


Ligand pose A, ICM: blue; AMBER: red

Appendix A2.7a. Ligand pose of tryptophanol, R-form, in the putative binding site of the 5-HT_{1E} receptor. Figure I shows the ligand pose yielding positive binding energy (ΔG: 4.80 kcal/mol). Figure II shows the superimposition of the receptor-ligand complex before (coloured blue/light blue) and after (coloured red/orange) energy minimisation of the complex in AMBER.



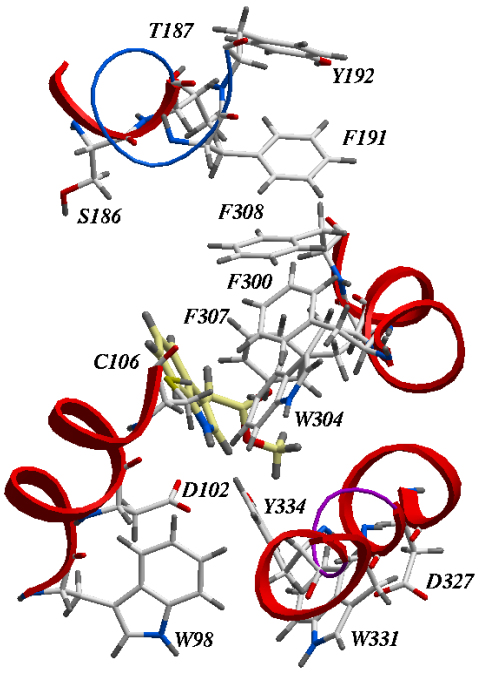
Ligand pose A



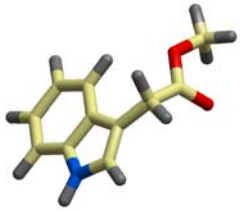
Tryptophanol, S-form

Appendix A2.7b. Ligand pose of tryptophanol, S-form, in the putative binding site of the 5-HT_{1E} receptor.

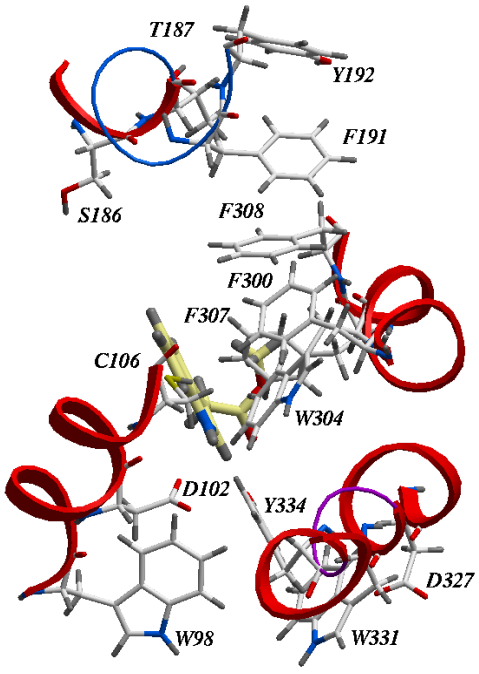
A2.8 Methyl-3-indoylacetate



Ligand pose A



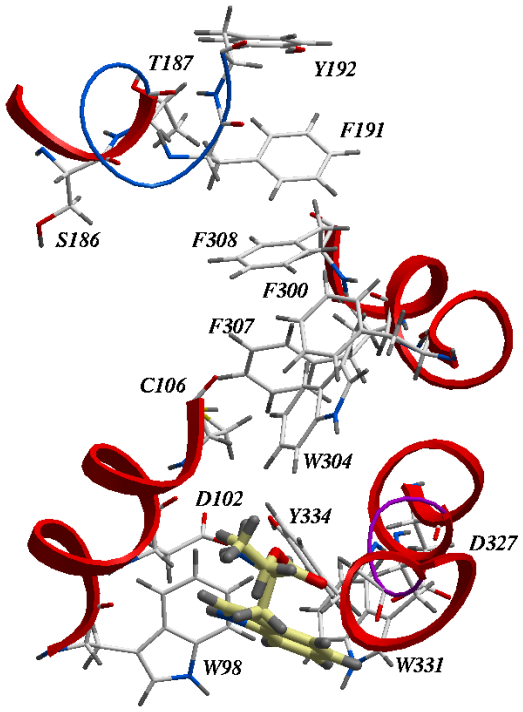
Methyl-3-indoylacetate



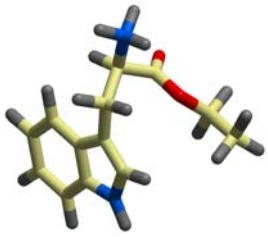
Ligand pose B

Appendix A2.8. Ligand poses of methyl-3-indoylacetate in the putative binding site of the 5-HT_{1E} receptor.

A2.9 Tryptophane ethylester

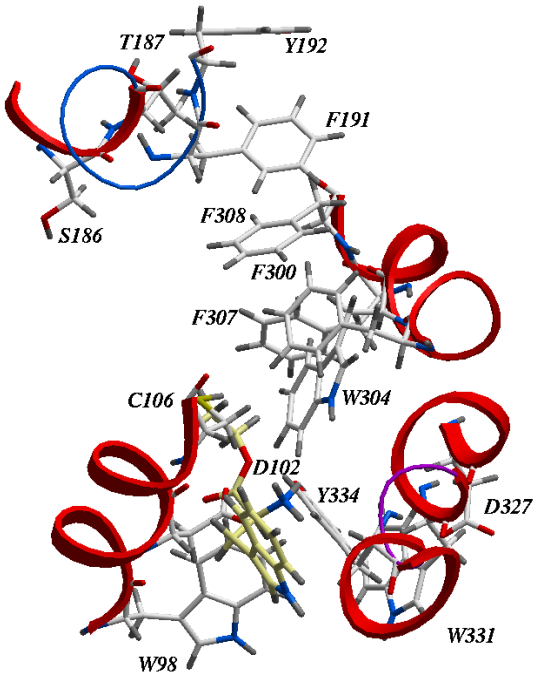


Ligand pose A

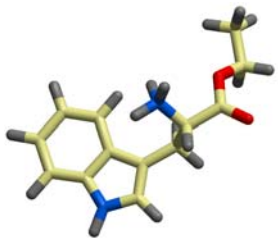


Tryptophane ethylester, R-form

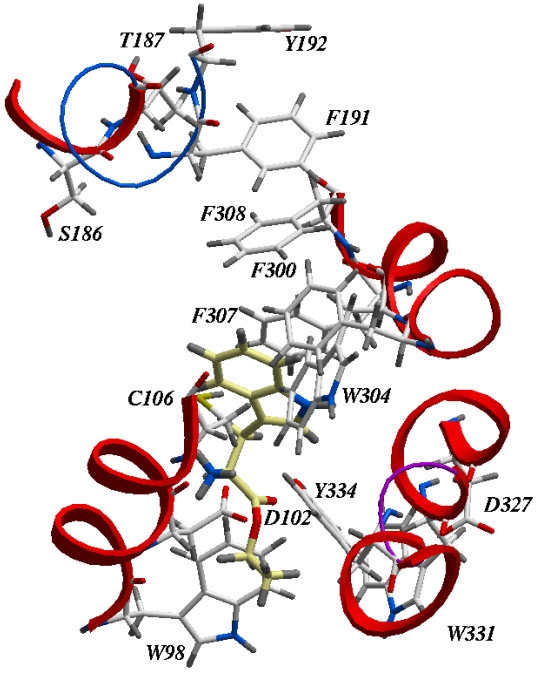
Appendix A2.9a. Ligand poses of tryptophane ethylester, R-form, in the putative binding site of the 5-HT_{1E} receptor.



Ligand pose A

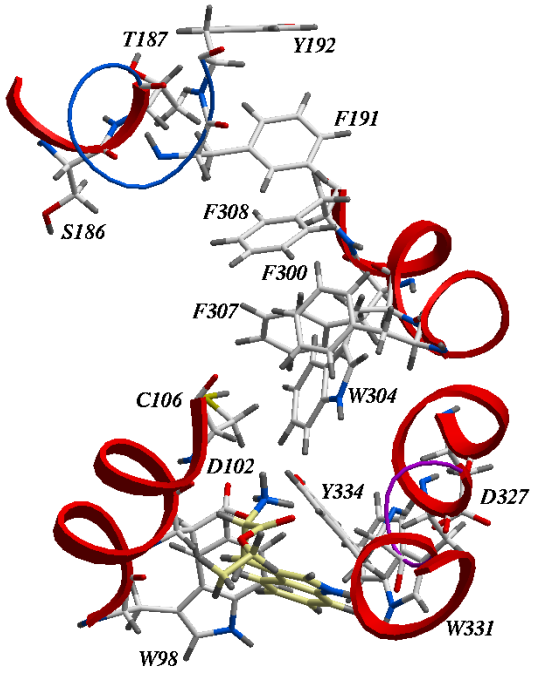


Tryptophane ethylester, S-form

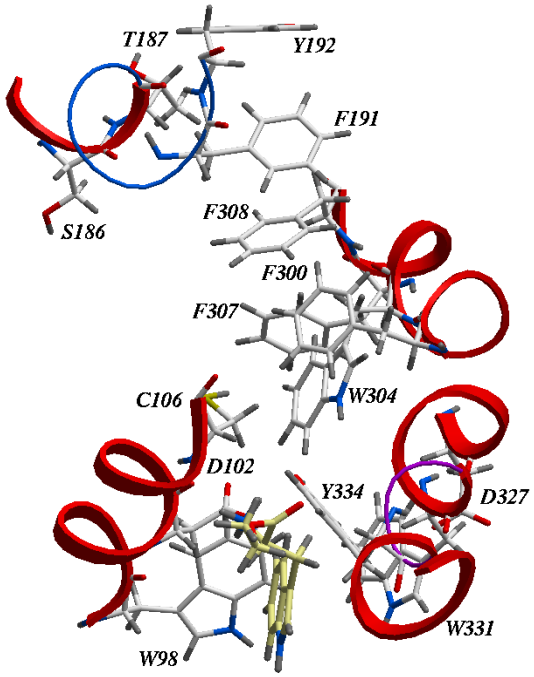


Ligand pose B

Appendix A2.9b. Ligand poses of tryptophane ethylester in the putative binding site of the 5-HT_{1E} receptor.

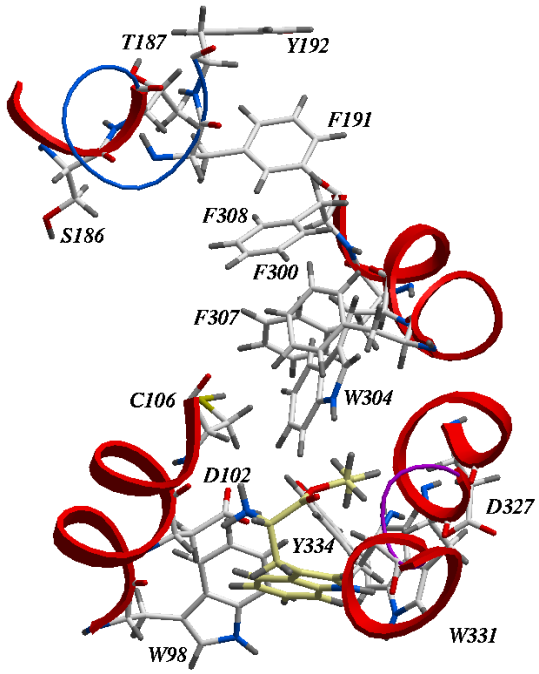


Ligand pose C

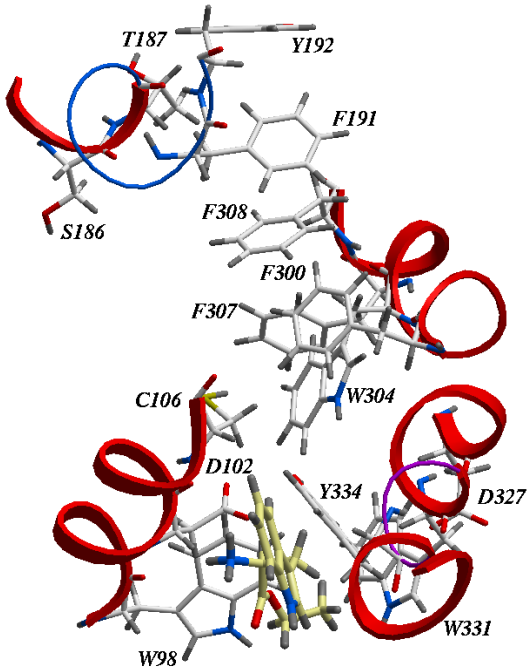


Ligand pose D

Appendix A2.9b continued. Ligand poses of tryptophane ethylester in the putative binding site of the 5-HT_{1E} receptor.

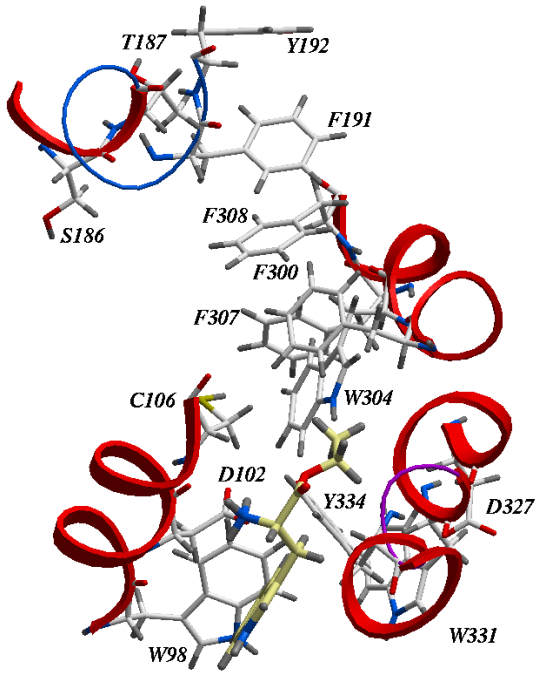


Ligand pose E

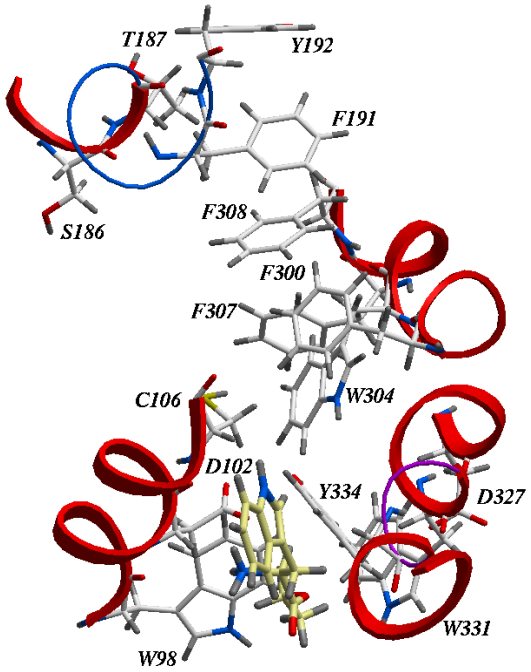


Ligand pose F

Appendix A2.9b continued. Ligand poses of tryptophane ethylester in the putative binding site of the 5-HT_{1E} receptor.

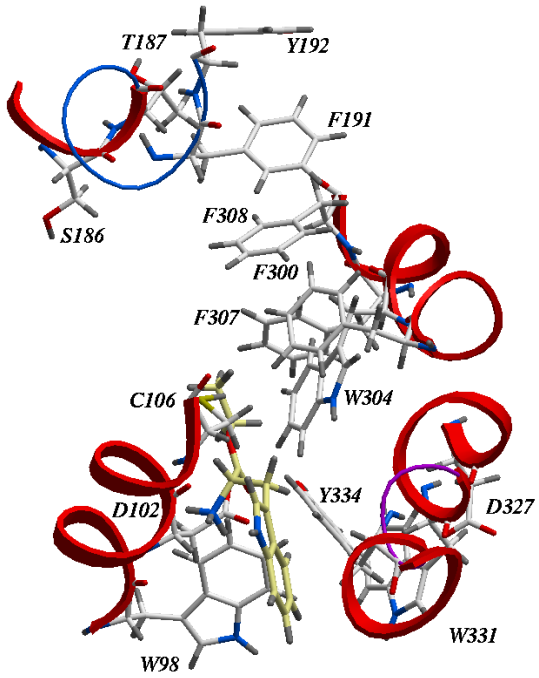


Ligand pose G



Ligand pose H

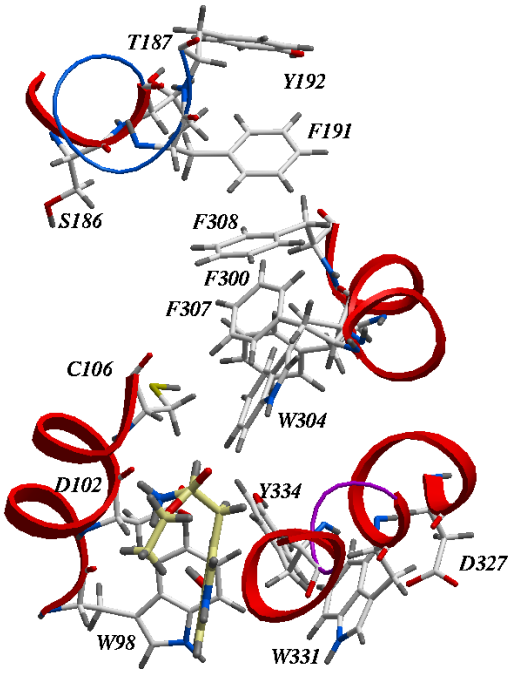
Appendix A2.9b continued. Ligand poses of tryptophane ethylester in the putative binding site of the 5-HT_{1E} receptor.

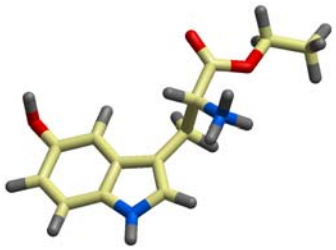


Ligand pose I

Appendix A2.9b continued. Ligand poses of tryptophane ethylester in the putative binding site of the 5-HT_{1E} receptor.

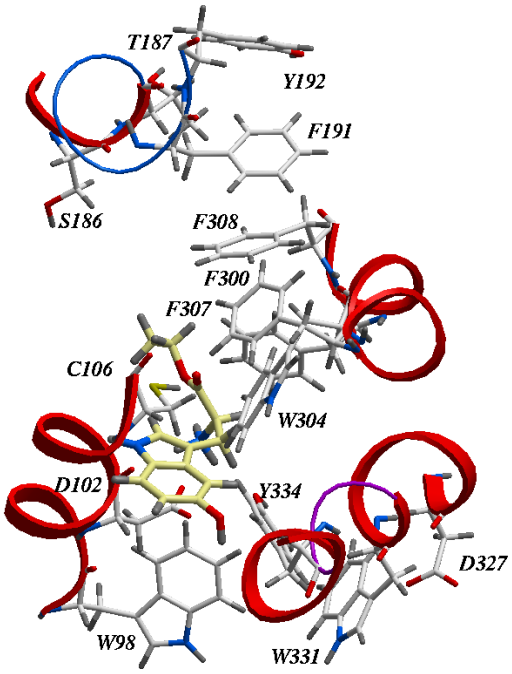
A2.10 5-Hydroxy-tryptophane ethylester





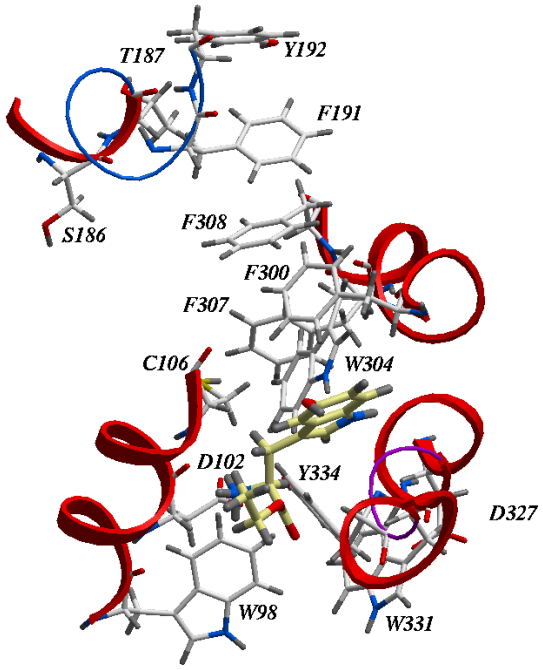
Ligand pose A

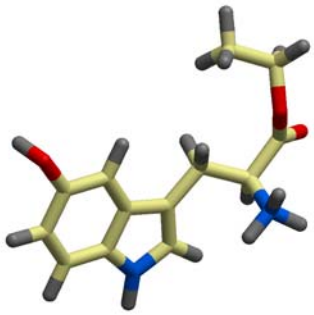
5-hydroxy-tryptophane ethylester, R-form



Ligand pose B

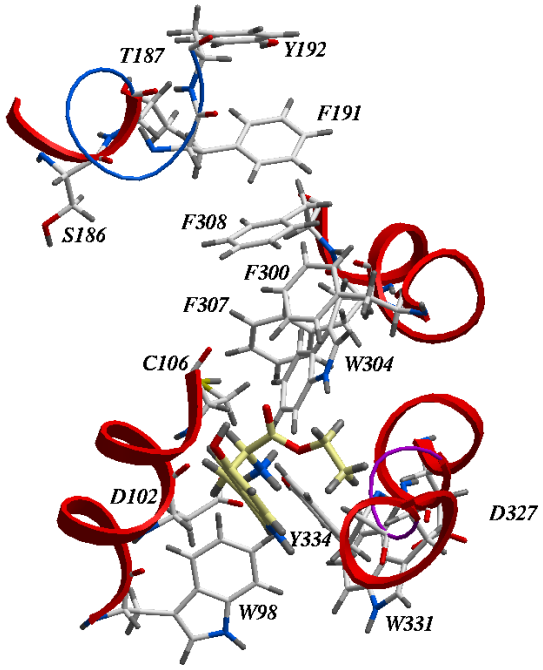
Appendix A2.10a. Ligand poses of 5-hydroxy-tryptophane ethylester, R-form, in the putative binding site of the 5-HT_{1E} receptor.





Ligand pose A

5-hydroxy-tryptophane ethylester, S-form



Ligand pose B

Appendix A2.10b. Ligand poses of 5-hydroxy-tryptophane ethylester, S-form, in the putative binding site of the 5-HT_{1E} receptor.

APPENDIX A3

Experimental results by Lise Roman Moltzau

Appendix A3. pK_i-values of the different ligands tested on the human 5-HT_{1E} receptor by Lise Roman Moltzau.

Compound	pK_i±SEM	n
5-HT	7.86±0.09	3
Tryptamine	5.85±0.09	3
α-Methylserotonin	6.73±0.09	3
5-Methoxytryptamine	5.49±0.18	3
3-(2-Bromoethyl)-indole	-3.94±2.45	3
3-(2-Hydroxyethyl)-indole	-0.54±1.46	3
2-Methyl-5-hydroxytryptamine	5.62±0.02	3
D-Tryptophanol	3.89±0.03	3
L-Tryptophanol	4.36±0.10	3
Methyl-3-indoylacetate	-1.60±1.43	3
D,L-Tryptophan ethylester hydrochloride	-4.58±2.05	3
L-Tryptophan ethylester hydrochloride	-3.56±2.05	3
5-Hydroxy-DL-tryptophan ethylester	4.28±0.03	3