UiT
THE ARCTIC
UNIVERSITY
OF NORWAY

FACULTY OF HEALTH SCIENCES
DEPARTMENT OF MEDICAL BIOLOGY
MEDICAL PHARMACOLOGY AND TOXICOLOGY RESEARCH GROUP

Binding mode of novel multimodal serotonin transporter compounds in 5-hydroxytryptamine receptors

Isak Bøgwald

Master thesis in molecular biotechnology (MBI-3941)
June 2016

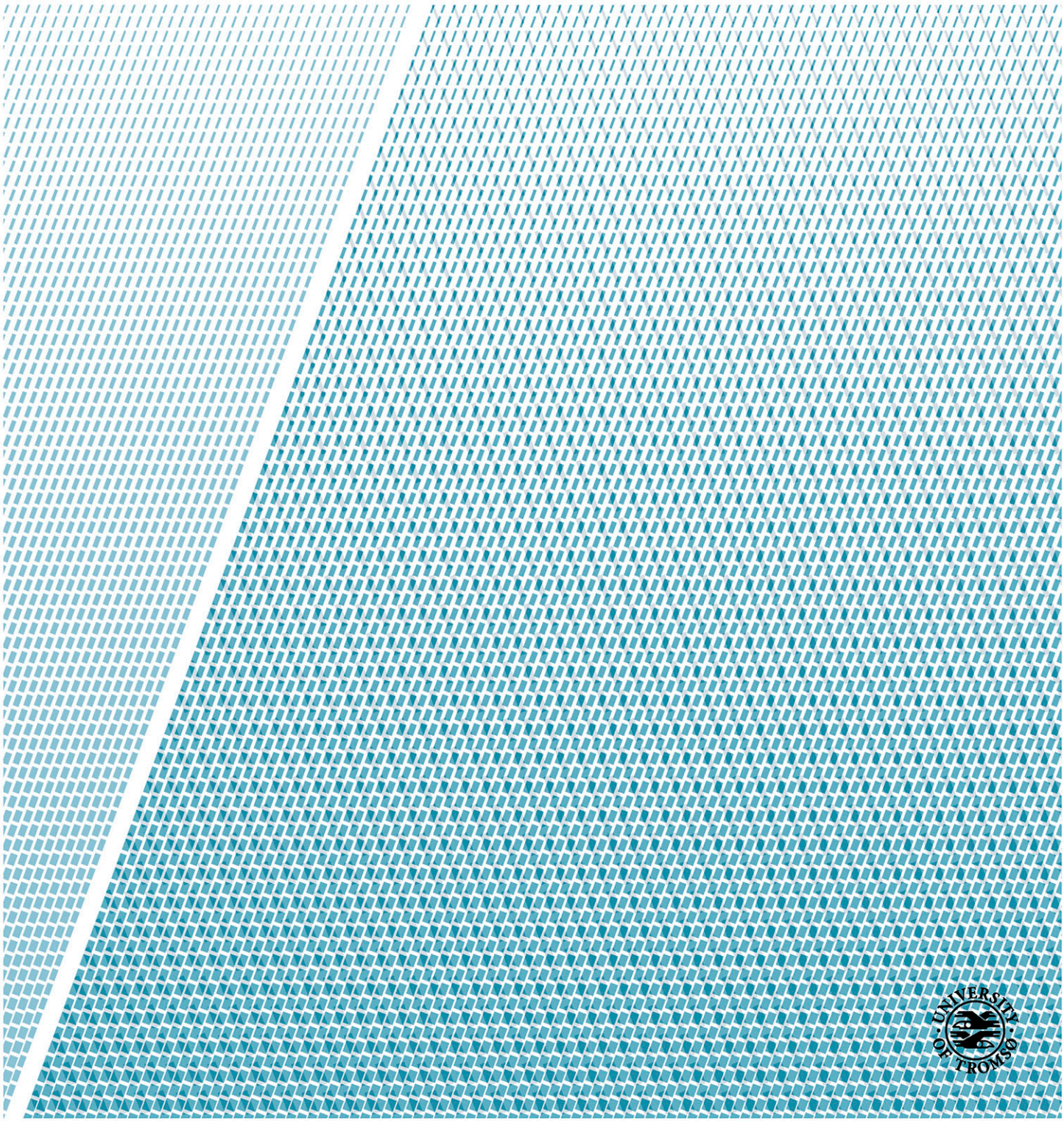


Table of contents

Acknowledgements

The master thesis was written at the Medical Pharmacology and Toxicology research group, Department of Medical Biology, Faculty of Health Science, UiT from September 2015 to May 2016.

I would first like to thank my supervisors Ingebrigt Sylte and Kurt Kristiansen for giving me the opportunity to write my thesis in the research group, allowing me to work with interesting fields like pharmacology, computational chemistry and structural biology. Their door was always open when I had questions, and their encouragement helped me to finally complete my master thesis.
Of the other people working at the research group, I would like to thank Thibaud Freyd for help with the methods used, Mari Gabrielsen for valuable input to the writing of the thesis, and Linn Evenseth for good advice and friendship.

I would also like to thank my good friends André and Espen, for all the good times over the years.

Finally, I wish to thank my mother Mette, father Jarl, sister Kamilla, brother Benjamin, godfather James, and the rest of my family for their love and support. This thesis would not have been possible without them.

Tromsø, May 2016

Isak Bøgwald

Abstract

Antidepressants are the most common treatment of depression, one of the leading causes of suicide and disability worldwide. Currently marketed antidepressants have certain limitations; they have a delayed response time, only about 1/3 of the patients respond to the first agent prescribed, and many of them produce side effects that reduce the quality of life. The need for more efficacious and faster-acting antidepressants with fewer side effects is thus apparent.

Studies have shown that 5-HT receptors (5-HTRs) are involved in many of the adverse effects of antidepressants, and may be responsible for efficacy issues and the delayed onset of therapeutic action. Some novel multimodal (two or more pharmacological actions) antidepressants combine inhibition of the serotonin transporter (SERT) with agonist or antagonist activity at 5-HTRs, to counteract the activity responsible for the aforementioned problems with the present antidepressants.

This study continues a previous virtual screening study, where we identified new compounds for SERT. Several of the compounds also showed affinity for one or more 5-HTRs. Although affinities are known, their ligand – 5-HTRs binding modes and their mode of action (agonist or antagonist action) for the target 5-HTRs have not been established. The aim of this study was to predict their mode of action, and to identify binding modes important for high affinity, by the use of computational methods. Homology modeling was used to construct models of 5-HT_{1A}R, 5-HT_{2A}R, 5-HT_6R and 5-HT_7R. The models were used for molecular docking and calculations of structural interaction fingerprints.

Several residues important for affinity to the target receptors were identified, and preferable binding modes were determined. The mode of action of the compounds was predicted based on their preferences for agonist/antagonist-selective models, and on previous studies of agonists and antagonists showing that agonists form strong polar interactions transmembrane helix 5 (TM5). The results indicated that several of the compounds might have potential to be developed into new antidepressant drugs.

Abbreviations

3D	Three-dimensional
5-HT	5-hydroxytryptamine
5-HTR	5-hydroxytryptamine receptor
5-HT_XR	5-hydroxytryptamine receptor subtype
AR	Adrenergic receptor
ATP	Adenosine triphosphate
BEDROC	Boltzmann-Enhanced Discrimination of Receiver Operating Characteristic
BDNF	Brain-derived neurotrophic factor
cAMP	Cyclic adenosine monophosphate
CNS	Central nervous system
D_3R	Dopamine D_3 receptor
DAG	Diacylglycerol
DAT	Dopamine transporter
DOPE	Discrete optimized protein energy
ECL	Extracellular loop
GPCR	G protein-coupled receptor
ICL	Intracellular loop
ICM	Internal Coordinates Mechanics
IFD	Induced fit docking
IP_3	Inositol triphosphate
K_i	Inhibitory constant
MAO	Monoamine oxidase
MAO-A	Monoamine oxidase isoform A
MAOI	Monoamine oxidase inhibitor
MDD	Major depressive disorder
NET	Norepinephrine transporter
NDRI	Norepinephrine-dopamine reuptake inhibitor
nM	Nanomolar
NRI	Norepinephrine reuptake inhibitor
NSS	Neurotransmitter sodium transporter
PDB	Protein Data Bank
PIP_2	Phosphatidylinositol 4,5-bisphosphate

PKC	Protein kinase C
PLC	Phospholipase C
PNS	Peripheral nervous system
RMSD	Root-mean-square deviation
SARI	Serotonin antagonist and reuptake inhibitor
SERT	Serotonin transporter
SIFt	Structural interaction fingerprint
SNRI	Serotonin-norepinephrine reuptake inhibitor
SPARI	Serotonin partial agonist reuptake inhibitor
SSRI	Selective serotonin reuptake inhibitor
STAR*D	Sequenced Treatment Alternatives to Relieve Depression
TCA	Tricyclic antidepressant
TM	Transmembrane α-helix
VSW	Virtual screening workflow
WHO	World Health Organization
Å	Ångström

1. Introduction

1.1. Depression and major depressive disorder

The diagnosis of major depressive disorder (MDD) requires a distinct change of mood, characterized by sadness or irritability and accompanied by at least several psychophysiological changes. Examples of these changes can be disturbances in sleep, appetite and sexual desire, physical manifestations such as constipation and slowing of speech and action, and emotional symptoms like loss of the ability to experience pleasure, crying and suicidal thoughts. The changes must last a minimum of two weeks and interfere considerably with work and family relations.[1]

A Norwegian psychiatric epidemiological study of adults aged 18-65 in a random sample of Oslo residents showed that the lifetime prevalence of major depression was found in 17.8 % of the respondents, where the prevalence for women (24.0 %) was more than two times higher than that for men (9.9 %).[2]

Depression is a complex disorder, with a course that varies greatly for the affected. The underlying causes for major depression are difficult to determine with absolute certainty, and there are indications that biological, psychological, and social factors all play a role in causing depression.[3] Current leading theories on the biological mechanisms of depression include genetics, monoamine neurotransmitter-deficiency, and neurotrophy.

The World Health Organization (WHO) ranked depression as the fourth leading cause of disability worldwide, and is projected to be the second leading cause by the year 2020.[4] According to the WHO, more than 800,000 deaths from suicide are recorded every year, and psychological autopsy studies (analysis of cause of death by psychological factors) from many countries show that more than 90% of suicide victims have one or more psychiatric illnesses, where the most common diagnosis is a major depressive episode.[5]

The most frequently prescribed therapies of depression are antidepressant medication and psychotherapy, either alone or in combination. Controversial treatments such as electroconvulsive therapy and stimulation of specific brain regions are less common. Other treatments, more associated with lifestyle changes, include bright light therapy, exercise, meditation, improved sleep routines and smoking cessation.

1.1.1. Biological basis of depression

Depression is a mental illness where the precise etiology is not well understood, though it is well known that biological, psychological and social factors are all part of this mood disorder.[6] Biological mechanisms of depression include genetics[7], the monoamine deficiency hypothesis, stress and the hypothalamic-pituitary-adrenal axis[8], neurotrophic factors[9], among others.

The discovery that monoamine deficiency is linked to depression was first predicted by observations of side effects of drugs that were developed for other diseases. An antihypertensive agent induced depression in a proportion of patients, by the depletion of presynaptic levels of serotonin and norepinephrine in the brain. Administration of monoamine precursors proved to reverse some of the symptoms.[10]

An antimycobacterial agent improved mood in tubercular patients with depression, and this agent was found to inhibit monoamine oxidase (MAO), the enzyme that degrades monoamines in the presynaptic nerve terminal. Inhibiting the enzyme prevents the degradation of serotonin and norepinephrine, and the observation that also depressed nontubercular patients experienced improved mood lead to the development of MAO inhibitors (MAOIs), the first antidepressants.[11] These agents produced increased levels of serotonin and norepinephrine in the brain; supporting the hypothesis that depression correlates with monoamine levels.[10]

Further evidence supporting the monoamine deficiency hypothesis comes from tryptophan depletion experiments. Tryptophan is the precursor amino acid for serotonin, and studies show that a depletion of tryptophan causes depressed patients in remission on antidepressants to experience a relapse of depressive symptoms, thus showing that decreasing levels of monoamines correlate with depression.[12]

Brain imaging studies of depressed patients demonstrate a reduction of the total volume of neurons in the limbic brain regions, notably the hippocampus and the prefrontal cortex.[13] Preclinical studies showed that depression and repeated exposure to stress result in atrophy of hippocampus neurons and loss of glia cells. These alterations could be reversed with antidepressant treatment, and this ability of the brain to remodel itself is called neuronal plasticity.[14]

Depression or stress decreases the expression of certain factors in the limbic brain regions, such as the brain-derived neurotrophic factor (BDNF). Neurotrophic factors are critical

regulators of the formation and plasticity of neuronal networks, and antidepressant treatment has been found to increase BDNF expression in hippocampus neurons of the brain. This is the basis of the proposal that depression is associated with reduced BDNF levels, and may explain the delay in onset of therapeutic effects of antidepressants. However, there is less evidence that depleting BDNF causes depression, suggesting neurotrophic factor levels are not solely responsible for the cause of depression.[9,13,14]

The antidepressants most commonly used today work by increasing the levels of the neurotransmitter serotonin (5-hydroxytryptamine, 5-HT) by inhibiting its reuptake by the serotonin transporter (SERT). In addition to reducing symptoms of depression, an increase of serotonin in the synaptic cleft is correlated with adverse effects. Present antidepressants have other limitations such as delay in onset of therapeutic action, and only about 1/3 of patients respond to the first agent prescribed. These limitations encourage the development of new and better antidepressant drugs to treat depression.

The increase of serotonin in the synaptic cleft leads to stimulation of serotonin receptors (5-hydroxytryptamine receptors, 5-HTRs), and this is theorized to account for many of the issues with the present antidepressants. A novel idea for developing better antidepressants is to combine SERT inhibition with effects on 5-HTRs.

1.2. Monoamine neurotransmitters

The nervous system coordinates our actions and is responsible for transmitting, receiving, and interpreting information from all parts of the body. It is divided into two parts; the central nervous system (CNS), which consists of the brain and the spinal cord, and the peripheral nervous system (PNS), which connects limbs and organs to CNS through nerves.

Synapses are structures in the nervous system that permit neurons to transfer signals to other neurons and effector organs, and neurotransmitters are the chemicals that move across the synaptic cleft, allowing this signaling between neurons and between neurons and effector organs to occur (figure 1).[15]

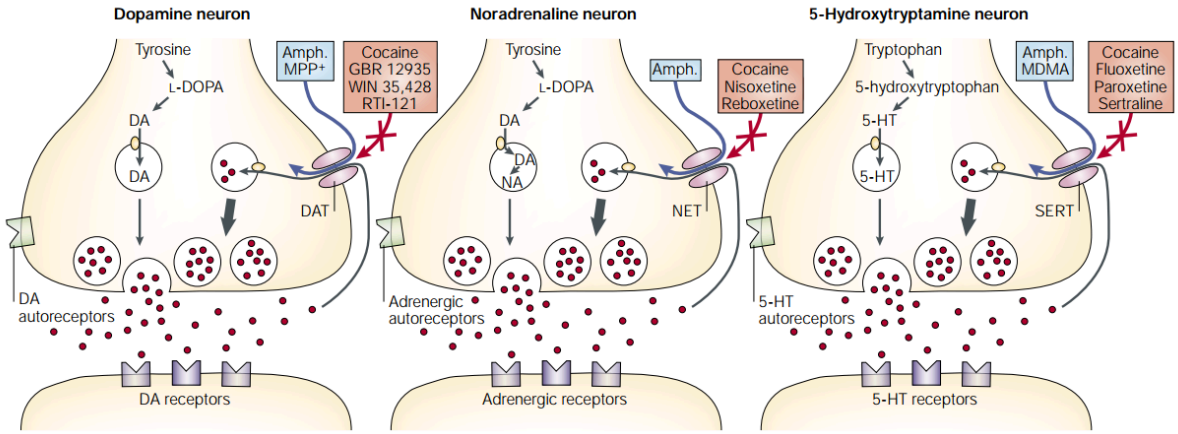


Figure 1. Schematic representation of dopamine, noradrenaline and 5-HT synaptic terminals.[15] Compounds in blue boxes are agents that act as monoamine transporter substrates; compounds in red boxes are agents that block the monoamine transporter protein. DA, dopamine; DAT, dopamine transporter; L-DOPA, L-3,4-dihydroxyphenylalanine; MPP+, 1-methyl-4-phenylpyridinium; Amph, amphetamine; NA, noradrenaline; NET, noradrenaline transporter; 5-HT, 5-hydroxytryptamine; SERT, serotonin transporter; MDMA, 3,4-methylenedioxy-methamphetamine.

Monoamine neurotransmitters are derived from amino acids like tyrosine and tryptophan. They contain one amine group connected to an aromatic ring structure via a two-carbon chain (figure 2). Monoamine neurotransmitters include tryptamines like serotonin (5-hydroxytryptamine, 5-HT), and catecholamines such as noradrenaline, adrenaline, and dopamine.

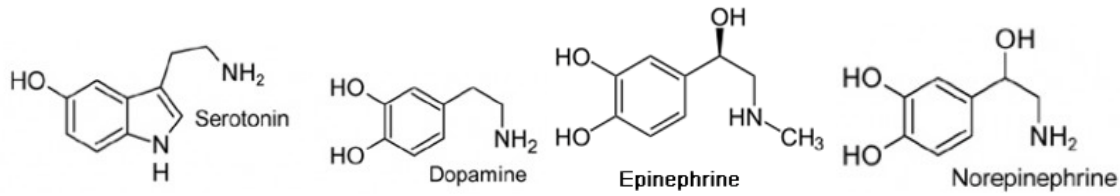


Figure 2. Monoamine neurotransmitter structures, showing the amine group connected to an aromatic ring structure via a two-carbon chain.

1.3. G protein-coupled receptors

Guanine nucleotide-binding protein-coupled receptors (GPCRs) comprise the largest integral membrane protein family in the human genome, with over one thousand members.[16] They are also known as seven-transmembrane receptors because the protein chain crosses the plasma membrane seven times, or heptahelical receptors as the seven transmembrane regions are secondary structure α-helices.

The superfamily of GPCRs shares the common topology of an extracellular N-terminal, seven transmembrane helices connected via loops, and an intracellular C-terminal (figure 3). The transmembrane helices (TMs) have a high degree of amino acid conservation, while the terminals and the intracellular loops (ICLs) and extracellular loops (ECLs) are more variable in length and sequence.

Ligands that bind to GPCRs include sensory signal mediators, biogenic amines (biogenic substances with one or more amine groups, such as hormones, neurotransmitters, etc.), chemokines, and peptide hormones. GPCRs have a major role in many physiological functions and in multiple diseases, and represent an estimated 30-45 % of current drug targets.[17]

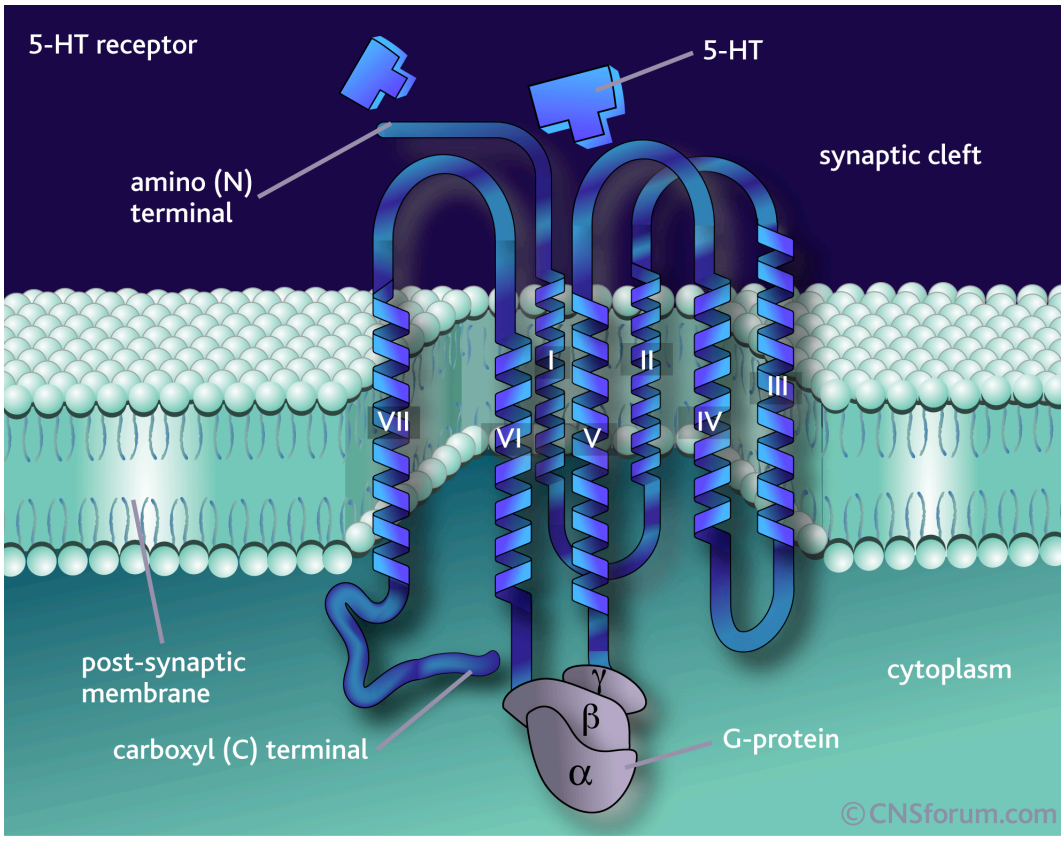


Figure 3. A schematic representation of a 5-HT GPCR, showing the seven transmembrane helices, the extracellular N-terminal and the extracellular loops at the synaptic cleft, the intracellular C-terminal and the intracellular loops at the cytoplasmic side, and the G protein subunits (CNSforum.com).

Activation of GPCRs occurs when they recognize an endogenous agonist, which causes activation of the heterotrimeric G protein. The active form of the G protein stimulates and inhibits specific effector proteins, such as enzymes and ion channels, resulting in rapid changes in the concentration of intracellular signaling molecules.[18,19]

GPCRs are divided into five families (A, B, C, Frizzled and Adhesion), and are classified by their ligands, amino acid sequences, clustering of genes, and globular domains and motifs.

1.3.1. 5-HT receptors

5-HT receptors (5-HTRs) belong to family A of GPCRs, the largest family by far, containing ~85% of the total number of GPCRs. The ligands that bind to family A GPCRs are diverse and include biogenic amines like hormones and neurotransmitters, and a light sensitive compound, among others. GPCRs within family A are known to have a high degree of sequence conservation in the transmembrane helices.

The most conserved amino acid for each helix within family A is given as a reference in the numbering scheme of Ballesteros & Weinstein.[20] The reference amino acid is given the number of the helix where it is located and the number 50, thus designating the most conserved amino acid in helix 1 as Asn1.50 (in one-letter notation as N1.50). Amino acids in the helices are numbered relative to the reference amino acid. The Ballesteros & Weinstein numbering scheme is widely used in articles describing structure-activity studies on GPCRs, and this makes it easier to compare results between studies of different receptors without being dependent on where the helices start and end, as this may vary between GPCRs.[18] Among amine receptors belonging to family A GPCRs, residue D3.32 is essential for ligand binding, as an ionic interaction to a protonated amine anchors the ligand to the receptor.

5-HTRs are mediators of both inhibitory and excitatory neurotransmission. They are located in both CNS and PNS, and are activated by the neurotransmitter serotonin, their natural ligand. 5-HTRs play important roles in biological and neurological functions such as aggression, anxiety, appetite, cognition, learning, memory, mood, nausea, sleep, and thermoregulation.[21] Because of their involvement in these processes, 5-HTRs are targets for a wide range of pharmaceutical drugs like antidepressants, antipsychotics, anorectics, antiemetics, gastroprokinetics agents, antimigraine agents, hallucinogens, and entactogens.[22]

The subtypes of 5-HTRs are coupled to different G protein pathways (Figure 4), producing either an inhibitory or excitatory response. The 5-HT_1Rs and 5-HT_5Rs are coupled to the G_i/G_0 pathway, which upon activation inhibits the formation of cyclic adenosine monophosphate (cAMP) from adenosine triphosphate (ATP) by inhibiting the enzyme adenylyl cyclase. 5-HT_2Rs are coupled to the G_q/G_{11} pathway, where activation stimulates phospholipase C (PLC) activity. This stimulation promotes the formation of diacylglycerol

(DAG) and inositol triphosphate (IP_3) from phosphatidylinositol 4,5-bisphosphate (PIP_2). The secondary messengers IP_3 and DAG activate downstream signaling pathways leading to an increase of protein kinase C (PKC) activity and Ca^{2+} release. Receptor subtypes 5-HT_4, 5-HT_6, and 5-HT_7 are coupled to the G_s pathway, which activates adenylyl cyclase leading to an increase in formation of cAMP.[22,23,24]

Unlike all the other 5-HTRs, which are GPCRs, the 5-HT_3R is a ligand-gated ion channel that differs both functionally and structurally from GPCRs.

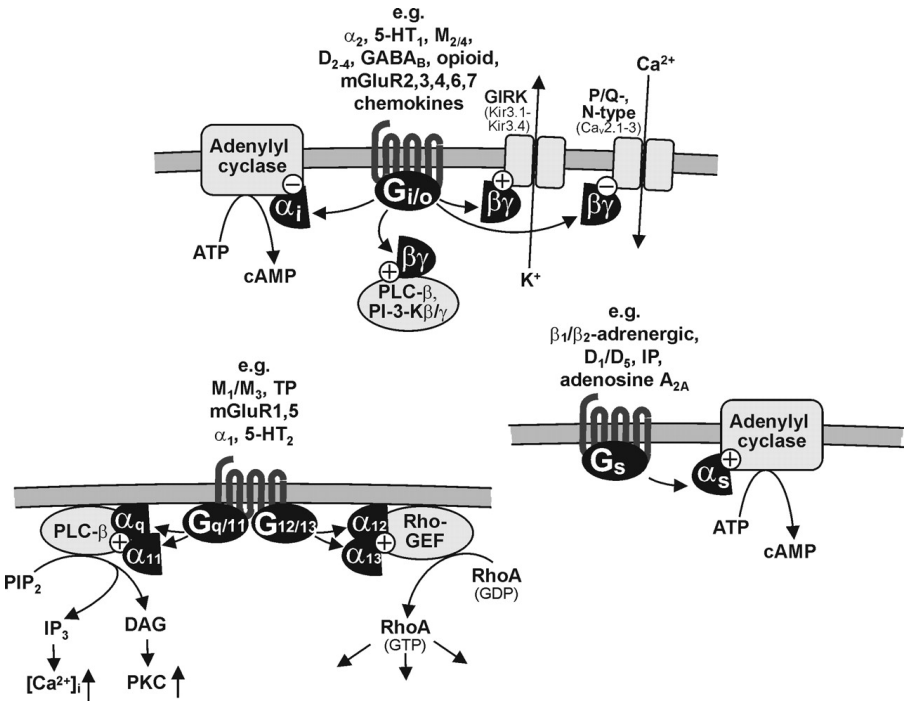


Figure 4. Typical patterns of receptor/G protein-coupling.[24] The subtypes of 5-HTRs are coupled to different pathways, either inhibitory or excitatory. Receptor subtypes 5-HT_1 and 5-HT_5 are coupled to G_i/G_0, 5-HT_2 is to coupled G_q/G_{11}, and 5-HT_4, 5-HT_6 and 5-HT_7 are coupled to G_S. α2, α2-adrenergic receptor; $D_{1–5}$, dopamine receptor subtypes 1 to 5; GIRK, G protein-regulated inward rectifier potassium channel; 5-$HT_{1,2}$, serotonin receptor subtypes 1 and 2; M1–5, muscarinic acetylcholine receptor subtypes 1 to 5; mGluR1–7, metabotropic glutamate receptor subtypes 1 to 7; PLC-β, phospholipase C-β; PI-3-K, phosphoinositide-3-kinase; PIP2, phosphatidylinositol 4,5-bisphosphate; IP3, inositol 1,4,5-trisphosphate; DAG, diacylglycerol; PKC, protein kinase C; Rho-GEF, Rho-guanine nucleotide exchange factor; TP, thromboxane A2 receptor; IP, prostacyclin receptor.

1.3.2. Activation of 5-HTRs by agonists

The activity of a receptor is induced by agonists, and inhibited by antagonists. Partial agonists activate receptors, but have only partial efficacy compared to a full agonist, while inverse agonists bind like agonists, but induce a response opposite to that of an agonist.

Detailed information is limited about exactly how an agonist induces activation of the G protein pathway in 5-HTRs, but studies of the family A GPCR β_1-adrenergic receptor (β_1-AR) have shown that interaction of agonists with polar residues (e.g. serine and threonine) of TM5 is important.[25] The hypothesis of about agonists induced receptor activation is based on the ability of the agonists to form strong polar interactions to residues in TM5, thus creating an inward shift of the helix. The resulting shift of TM5 is transmitted down the helix to the intracellular side, where rearrangements of TM5 and TM6 open the cleft for G protein binding.[26] A virtual screening study of 5-HT_{2A}R to identify agonist and antagonist molecules supports the hypothesis[27], thus making it possible to extrapolate the hypothesis to 5-HTRs.

1.4. Serotonin transporter and the serotonergic system

Neurotransmitter transporters are, like GPCRs, transmembrane proteins. They are involved in the movement of ions, small molecules, and macromolecules across the cell membrane. The serotonin transporter (SERT) is, along with the dopamine transporter (DAT) and the norepinephrine transporter (NET), a part of the monoamine subfamily of neurotransmitter sodium symporters (NSS). NSSs control the termination of signaling of biogenic amines by regulating the sodium- and chloride-dependent reuptake of neurotransmitters.

SERT removes the neurotransmitter serotonin from the synaptic cleft and simultaneously enables its reuse by the presynaptic neuron, thus regulating the concentration of serotonin by reuptake in the synaptic cleft.[28]

Along with the 5-HTRs, SERT is part of the serotonergic system (figure 5). This neurotransmitter system, and its signaling pathways, influence neurological processes including mood, sleep, cognition, pain, hunger and aggression.[29]

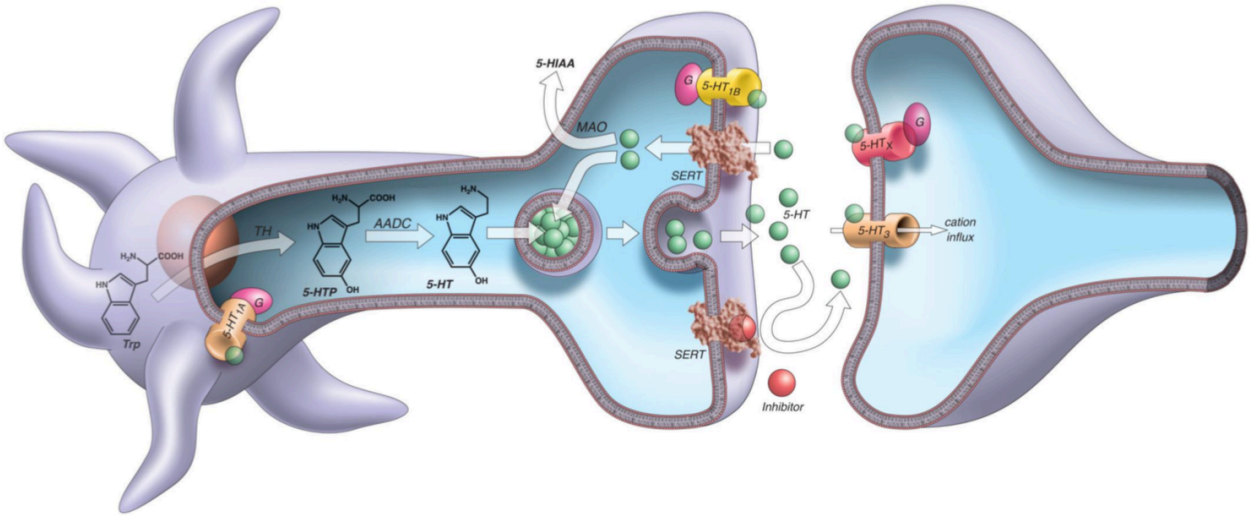


Figure 5. The serotonergic system.[30] Trp, tryptophan; TH, tryptophan hydroxylase; 5-HTP, 5-hydroxytryptophan; AADC, aromatic L-amino acid decarboxylase; 5-HT, 5-hydroxytryptamine; 5-HIAA, 5-hydroxyindoleacetic acid; MAO, monoamine oxidase; 5-HT_{1B}, 5-hydroxytryptamine receptor 1B; SERT, serotonin transporter; 5-HT_{x}, G protein-coupled receptor 5-hydroxytryptamine receptor; 5-HT_{3}, 5-hydroxytryptamine receptor 3.

1.5. Antidepressants

Antidepressants are drugs used for the treatment of depression and other conditions such as anxiety disorders, obsessive-compulsive disorder, eating disorders, chronic pain, and neuropathic pain. The goal when treating depression with antidepressants is to have complete remission of symptoms, without relapses or recurrent episodes in the future.[31] The complexity of the biology of depression, and the multiple targets involved in the disorder, often require more than one mechanism of action for the antidepressants to have therapeutic effect.[32]

1.5.1. Classes of antidepressants, their mechanism and side effects

Antidepressants have the monoamine deficiency hypothesis as their theoretical basis, employing mechanisms to increase monoamine neurotransmitter levels in the brain. They act by inhibiting the reuptake of monoamines to the presynaptic neuron mediated by the monoamine transporters (SERT, NET, DAT), or by inhibition of enzymes that degrade monoamines (e.g. MAO inhibition). The side effects (table 1) of antidepressants arise from their activity on a wide range of other receptors, including histaminic H_1, muscarinic M_1, α_1- and α_2-adrenergic receptors. The increase of serotonin levels in CNS and PNS causes stimulation of other 5-HTRs as well, and this correlates to certain side effects.

Antidepressant drugs are agonists, antagonists, or partial agonists at receptors they elicit activity on.

Table 1. Overview of antidepressant classes, their mechanism and side effects

Antidepressant class	Target protein	Mechanism	Side effects
MAOI	MAO-A	Enzyme inhibition	Weight gain, nausea, headache, drowsiness, insomnia, etc.
TCA	SERT, NET	Reuptake inhibition	Weight gain, sexual dysfunction, nausea, dry mouth, drowsiness, constipation, etc.
SSRI	SERT	Selective reuptake inhibition	Sexual dysfunction, weight gain, sleep pattern alterations, etc.
SNRI	SERT, NET	Reuptake inhibition	Nausea, weight gain, sleep pattern alterations, etc.
SPARI	SERT, 5-HT_{1A}R	Reuptake inhibition, partial agonism at 5-HT_{1A}R	Diarrhea, nausea, headache, etc.
SARI	SERT, 5-HT_{2A}R, 5-HT_{2C}R	Reuptake inhibitor, antagonism at 5-HT_{2A}R and 5-HT_{2C}R	Nausea, headache, dry mouth, blurred vision, fatigue, etc.
NRI	NET	Selective reuptake inhibition	Nausea, insomnia, dry mouth, etc.
NDRI	NET, DAT	Reuptake inhibition	Agitation, anxiety, headache, itching, etc.

MAOIs inhibit the activity of MAO-A; an enzyme that metabolizes monoamine neurotransmitters. By inhibiting MAO-A, less monoamine neurotransmitters are metabolized, resulting in an increase of monoamine concentration.

Tricyclic antidepressants (TCAs) and serotonin-norepinephrine reuptake inhibitors (SNRIs) both get their pharmacological effect by inhibiting the reuptake of norepinephrine and serotonin by their transporters SERT and NET. SNRIs have significantly fewer side effects than TCAs; this is because they, as contrary to TCAs, have little activity on histaminic, muscarinic and α_1 adrenergic receptors.

Selective inhibitor antidepressants have their main activity on a specific target, having no or minimal activity on other targets. Selective serotonin reuptake inhibitors (SSRIs) increase the concentration of serotonin in the synaptic cleft by inhibiting reuptake of serotonin by SERT. Inhibiting the other target NET, the norepinephrine reuptake inhibitor (NRI) increases norepinephrine concentration.

Norepinephrine-dopamine reuptake inhibitors (NDRIs) combine inhibition of reuptake by NET with inhibition of reuptake by DAT. Inhibition of DAT is associated with pleasure and the reward system, suggesting an additional target to relieve depressive symptoms. DAT inhibition is the main mechanism of action of different drugs of abuse, and may result in addiction.

Several new antidepressants have been developed in recent years, notably antidepressants combining reuptake inhibition of SERT with agonist or antagonist action on 5-HTRs.

Serotonin partial agonist reuptake inhibitor (SPARI) combines SERT reuptake inhibition with partial agonism of the 5-HT_{1A}R. This produces a pharmacologic synergy upon the serotonergic system, and it has been hypothesized that this causes a more immediate and lasting elevation of serotonin in the brain.[33]

Serotonin antagonist and reuptake inhibitors (SARIs) act by antagonizing 5-HT_{2A} and 5-HT_{2C} receptors, combined with reuptake inhibition of SERT.

Table 1 lists the different antidepressant classes, their mechanisms and often-occurring side effects.

1.5.2. Delayed therapeutic effects and efficacy issues

One of the biggest problems with the antidepressants available today is the delay in onset of their therapeutic effect, antidepressive response can take 2-6 weeks. When inhibiting a monoamine transporter with an antidepressant, the blockade can be detected immediately, but the therapeutic effect takes weeks to become clinically important.[34]

Presynaptic 5-HT_{1A}Rs are hypothesized to be at least partially responsible for the delay, because of their inhibitory effect on serotonergic activity when stimulated. By inhibiting serotonin reuptake, negative feedback systems cause 5-HT_{1A}Rs to downregulate and desensitize. Once 5-HT_{1A}Rs are desensitized, serotonin can no longer effectively turn off its own release, and the resulting disinhibition of the serotonin neuron causes a flurry of serotonin release. The adaption of the 5-HT_{1A}Rs takes time, and the time course of the desensitization correlates with the delay in therapeutic effect.[31]

The Sequenced Treatment Alternatives to Relieve Depression trial (STAR*D)[35] is a research study on remission rates in depressed patients treated with antidepressants. Its goal was to study the efficacy and effectiveness of antidepressants. The study consisted of four steps lasting 12 weeks each. Patients not in remission, or intolerant to the antidepressant, after each step would move on to the next. The first step was treatment with an SSRI, and only about 30% met the criteria for remission. Step two and three involved either a switch to another antidepressant, or an augmentation of the SSRI with an additional antidepressant, increasing cumulative remission rates to 50% after step two and 60% after step three. The fourth and last step was a switch to another antidepressant, further increasing the total remission rate to about 70%, for patients who remained in the study. By the end of the 12-month follow up care, only a minority of the patients had not relapsed or dropped out of the study.
The STAR*D study got criticism for its overestimation of reported remissions and questionable use of statistics, suggesting an actual lower percentage of patients in remission.[36]
In any case, the study and its subsequent criticism both make arguments for the inefficacy of antidepressant drugs, and ineffectiveness of current antidepressant treatment.

1.5.3. Multimodal activity and the search for novel antidepressants

As described in this chapter, most of the antidepressants available today come with a wide range of side effects. They have a delayed therapeutic effect and are not especially efficacious. All this proves that research and development of new drugs are needed to treat depression.
Previously mentioned SARIs, which are found to produce significantly fewer and less debilitating side effects, and SPARIs, showing increased efficacy and reduced delay of therapeutic effect, are promising future directions in the development of novel antidepressants. These antidepressants that elicit activity on 5-HTRs combined with inhibition

of SERT are known as multimodal, a term used to describe compounds with at least two different pharmacological modes of action.[37]

SSRIs produce a wide range of side effects, while SARIs (e.g. trazodone) lack debilitating side effects including sexual dysfunction, insomnia and anxiety.[31] SSRIs raise the serotonin levels to act on all 5-HTRs, causing a stimulation of 5-HT_{2A} and 5-HT_{2C} receptors, while SARIs inhibit these two 5-HTRs. It is thus plausible that 5-HT_{2A} and 5-HT_{2C} receptors are involved in adverse effects of SSRI treatment.

SPARIs (e.g. vilazodone and vortioxetine) have effects on various 5-HTRs as well as SERT. The partial agonism they have on the 5-HT_{1A}R is associated with accelerated clinical effects of antidepressants, and even seems to possess other potentially therapeutic actions such as antianxiety[38], and antiaggressive[39] properties. Vortioxetine is a novel antidepressant with effects on multiple 5-HTRs, including 5-HT_{1B}R, 5-HT_3R, and 5-HT_7R, in addition to the inhibition of SERT and partial agonism at 5-HT_{1A}R.[40] The antagonism of 5-HT_3 and 5-HT_7 receptors is suggested to increase the efficacy of vortioxetine, by inducing the release of extracellular serotonin in regions of the brain.

In addition to the 5-HTRs involved in the mechanisms of SARIs and SPARIs described above, the 5-HT_6R is considered a new target for antidepressant drugs. 5-HT_6R antagonists exert an antidepressant effect, and are speculated to accelerate onset of therapeutic action and minimize side effects.[41]

1.6. Molecular modeling

Molecular modeling describes the behavior of molecules and molecular systems by applying methods such as theoretical and computational chemistry to a target of interest. Theoretical modeling can give insight to processes and mechanisms that might be impossible or too expensive to study with experimental methods. A model is defined as a simplified/idealized description of a system or process, often in mathematical terms, making it possible to perform calculations and predictions on it.[42] Molecular modeling has three stages: (1) selection of model of interest, (2) calculations of intra- and intermolecular interactions, and (3) analysis of the calculations. Quantum mechanics and molecular mechanics are the two main approaches for the model description, and they are applied individually or in combination. Quantum mechanics takes the movement of electrons relative to the nucleus into consideration, and is a better description of a model in terms of accuracy of geometry and energy calculations compared to molecular mechanics. The disadvantage of quantum mechanics is the

computational costs, because the calculations are so time-consuming it is only applicable to small molecules.[43]

1.6.1. Molecular mechanics

When modeling large molecules, such as proteins, molecular mechanics is the most commonly used approach. Molecular mechanics apply the Born-Oppenheimer approximation to the representation of atoms, which ignores the movements of electrons and treats each atom as a particle. Each molecule is treated as a collection of particles interacting with each other via harmonic forces according to Hooke's law. Due to this simplification, molecular mechanics is a relatively fast computational method, making it a compromise between accuracy and computational efficiency.[42]

A molecular mechanics force field calculates the potential energy of a system of atoms. Molecular mechanics enables the calculation of the total steric energy of a molecule in terms of deviations from reference 'unstrained' bond lengths, angles and torsions plus nonbonded interactions (equation 1). The collection of unstrained values, together with empirically derived parameters for different types of atoms, makes up the force field.[43] The general form of the force field can be written as:

$$E_{tot} = E_{bonded} + E_{nonbonded}$$
$$E_{tot} = (E_{bond} + E_{angle} + E_{dihedral}) + (E_{vdw} + E_{elec}) \quad (1)$$

E_{tot} is the total potential energy, the sum of all bonded (E_{bonded}) and nonbonded ($E_{nonbonded}$) interactions. E_{bonded} is all the bond stretching (E_{bond}), angle bending (E_{angle}) and torsional ($E_{dihedral}$) energy terms, and $E_{nonbonded}$ is all the Van der Waals (E_{vdw}) and electrostatic (E_{elec}) energy terms.[43]

1.6.2. Homology modeling

Homology modeling, also known as comparative modeling, is the model construction of a target protein, based on its amino acid sequence and a three-dimensional (3D) structure of a related homologous protein (figure 6). In the absence of an experimentally determined structure of the target protein, homology modeling can be a useful tool in the study of structure and activity of a protein of interest. Homology modeling is possible because homologous proteins, related by divergence from a common ancestor, have a similar overall 3D folding. The 3D structure of a protein is more conserved through evolution than its amino

acid sequence. This makes it possible to use homologous proteins as templates in modeling, even if the amino acid sequences of target and template are dissimilar. The most conserved regions of a protein are in its hydrophobic core, where only small changes in sequence are tolerated to preserve the overall 3D fold, while loop regions and protein surface are more variable regions. The seven transmembrane helices of GPCRs comprise the hydrophobic core, and are the regions where sequence identity and structural similarity are highest. Loop regions are more flexible, thus are not as important for the overall 3D fold, and can tolerate more changes than the hydrophobic core without affecting structure and function.

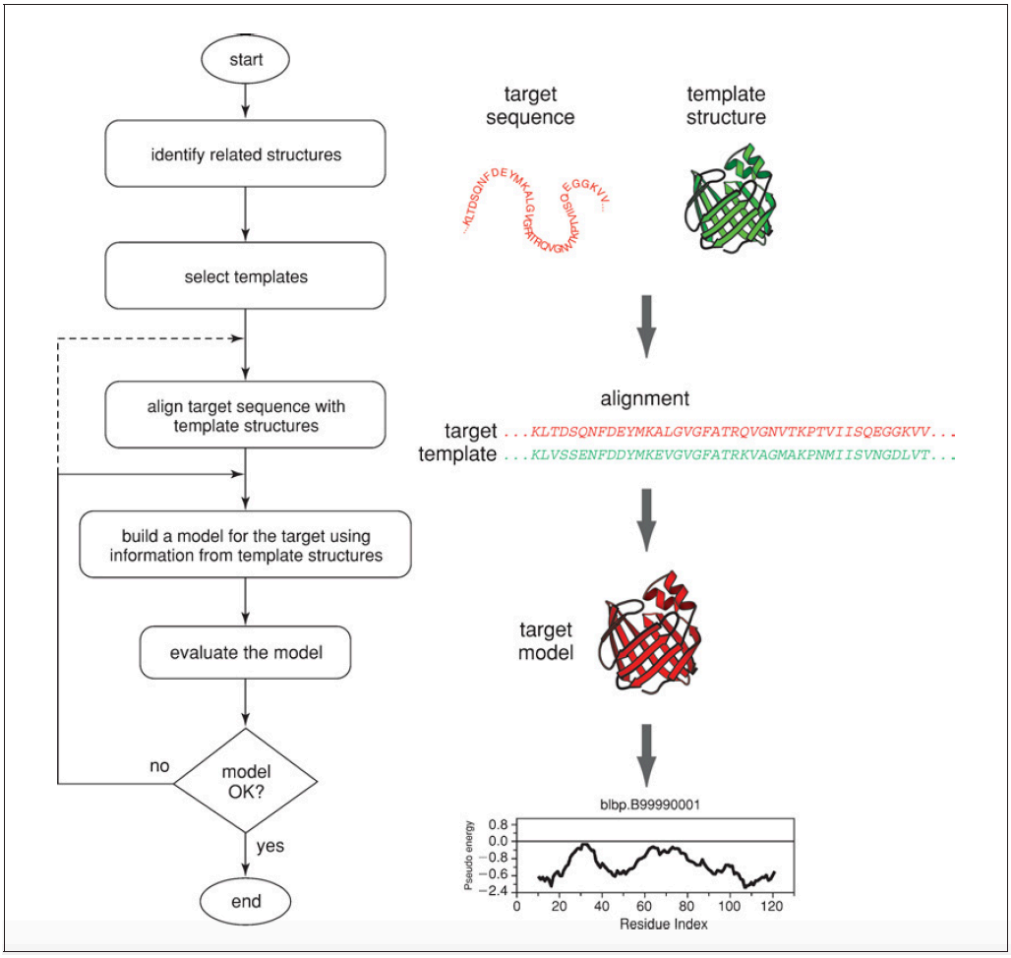


Figure 6. Steps in homology modeling.[44]

Homology modeling consists of four main steps: (1) identification of suitable template, (2) amino acid sequence alignment, (3) construction and refinement of model, and (4) evaluation and validation of the model.

Identifying suitable templates for family A GPCRs (e.g. 5-HTRs) is a pretty straightforward task. This is due to the large number of experimentally determined structures and the high degree of structure conservation in this family.

However, in some cases the target protein has no known related experimentally solved structures. Searching the PDB archive, comparing the target sequence to the sequence of all proteins that have an experimentally solved 3D structure, can identify a suitable template. This search can be done with tools such as FASTA[45] and Basic Local Alignment Search Tool (BLAST).[46]

An amino acid alignment between target sequence and template sequence is made after the identification of a template. The quality of the model depends on a correct alignment, and errors made in the sequence alignment can cause incorrect models. Having multiple homologous proteins in the alignment highlights evolutionary relationships, and increases the probability that corresponding sequence positions are aligned correctly.[47]

The construction of models is divided into three stages: (1) generation of amino acid backbone of structurally conserved regions, (2) generation of the non-conserved regions (loops), and (3) optimization of sidechains and backbone. Predicting loop regions is one of the difficulties of homology modeling, as templates often lack structural information on loops because solving a flexible structure by crystallization is problematic.

Refinement of the model can be performed by energy minimizations, Monte Carlo simulations and/or molecular dynamics calculation.[47] The refinement relaxes the energy of system, and corrects unfavorable positions of residues.

The last step of the homology modeling procedure is the evaluation of the model. Tools for assessing the 3D quality and stereochemistry of the model include the Structural Analysis and Verification server (SAVES, nihserver.mbi.ucla.edu/SAVES/) and the ModFOLD4 (http://www.reading.ac.uk/bioinf/ModFOLD/) server. Molecular docking studies to evaluate the ability of the models to discriminate between molecules known to interact with the target and molecules that most probably not bind are also useful as a test of the model quality. In addition, docking may confirm the binding mode and binding site conformation of the model. Available experimental information, such as mutagenesis studies, can be used to support the evaluation.

The accuracy of a homology model depends on: (1) the functional and sequential similarity between the template protein and the target, (2) the alignment of the template with the target, and (3) the resolution of the crystal structure of the template protein. A sequence identity between template and target of 50% is expected to give a model with a Cα-root mean square deviation (RMSD) of approximately 1Å.[48] RMSD is a measurement of structural similarity, with low RMSD representing structures with similar tertiary arrangement, based on the average distance between Cα-atoms of the template and corresponding Cα-atoms of the target

model. Acceptable models are generally achieved with a sequence identity of 30%, which is expected to give an RMSD of 2 Å, while lower sequence identity decreases the probability of correct homology models.[48]

1.6.3. Docking and scoring

Docking is the prediction of conformation and orientation of a ligand in the binding site of a target. While both ligand and target structures are flexible in nature, the complexity of large molecules makes this computationally exhausting in molecular modeling studies. Though the computational capacity has increased dramatically the last decade, the most common approach is the docking of a flexible ligand into a rigid macromolecule target (e.g. protein). Incorporating flexibility in docking studies can be achieved with a method like Induced Fit Docking (IFD)[49], developed to reflect the change in protein structure upon ligand binding. While some proteins do not change substantially upon ligand binding, others rearrange sidechains and/or backbone to get the correct binding conformation for a given ligand.

Another approach for incorporating protein flexibility is docking into multiple slightly different models of the target protein, where the variations among the models represent different conformations of the target. Generation of multiple structures can be achieved with modeling software like MODELLER[44], or by methods such as molecular dynamics and Monte Carlo sampling.

The description of interactions of ligand-protein complexes is often a goal in molecular docking. Traditionally these interactions are inspected visually, but can also be evaluated by schematic representations such as LIGPLOT[50], or automated approaches like Structural Interaction Fingerprint (SIFt).[51]

Scoring is the evaluation of the interaction between the protein and its docked ligand. Three important applications of scoring functions in molecular docking include: (1) determination of the binding mode and site of a ligand on a protein, by ranking ligand conformations based on binding tightness of the ligand-protein complexes, (2) prediction of binding affinity between ligand and protein, and (3) identification of potential drug hits for a target in a virtual screening, where the scoring function should be able to rank known binders highly.[52]

The binding affinity that scoring functions aim to predict is in free energy terms, and the free energy of the binding is given by the Gibbs-Helmholtz equation (equation 2):

$$\Delta G = \Delta H - T\Delta S \qquad (2)$$

ΔG is the free energy of binding, ΔH the enthalpy, T the temperature in Kelvin and ΔS the entropy. ΔG is related to the binding constant K_i by (equation 3):

$$\Delta G = -RT\ln K_i \quad (3)$$

R is the gas constant.[43] K_i is often used as a measurement of ligand binding affinity.

Multiple different scoring functions for protein-ligand interactions have been developed, with different accuracies and computational efficiencies. The most commonly used scoring functions can be divided into four categories: (1) force field scoring functions, (2) empirical scoring functions, (3) knowledge-based scoring functions, and (4) consensus scoring.[52] Force field scoring functions are based on the energy terms of the molecular mechanics force field (Equation 1). Empirical scoring functions estimate the binding affinity of the ligand-protein complex on the basis of a set of weighted energy terms such as Van der Waals energy, electrostatics, hydrogen bond, desolvation, entropy, hydrophobicity, etc. Compared to force field scoring functions, the empirical scoring functions have faster calculations because of simpler energy terms.[52] Knowledge-based scoring functions employ energy potentials derived from the structural information of experimentally determined structures, and offer a balance between accuracy and speed. Consensus scoring combines the three previously described scoring functions, to take the advantages and balance the deficiencies of the different scoring functions.[52]

2. Aim of study

Recently developed antidepressants combine the traditional SERT inhibition with activity at other 5-HTRs, and are known as multimodal antidepressants. Agonism of the 5-HT_{1A}R and antagonism of other 5-HTRs has been found to increase efficacy and accelerate onset of therapeutic action, and to produce fewer and less debilitating side effects.

A recent study identified 74 novel SERT compounds[53], and 18 of these were found to also have affinity for one or more 5-HTRs (supplementary information III and IV).

This study aims to examine and describe the interactions between the 18 multimodal compounds and the target 5-HTRs (human 5-HT_{1A}R, 5-HT_{2A}R, 5-HT_6R, and 5-HT_7R) by computational methods, and to possibly hypothesize if the 18 compounds elicit agonist or antagonist activity at the receptors. Because none of the 5-HTRs in this study have experimentally determined structures, the homology modeling technique was used to construct models of the receptors.

The sub-goals of the study were to:

a) Construct homology models with multiple conformations of the human 5-HT_{1A}R, 5-HT_{2A}R, 5-HT_6R, and 5-HT_7R targets, based on templates with active state and inactive state
b) Select models with preference to agonists or antagonists for each target
c) Study interactions of known agonists and antagonists with the targets, and analyze binding site interactions with SIFt
d) Perform docking of multimodal compounds in agonist/antagonist-selective models
e) Identify favorable binding modes correlating with high affinity for the multimodal compounds
f) Predict mode of action (agonist or antagonist) for the multimodal compounds, based on preference to agonist/antagonist-selective models and conformation in binding site

3. Methods

3.1. Software and databases

3.1.1. Software

Schrödinger Software Release 2015-4

The Schrödinger software package contains tools for molecular modeling and drug design. Schrödinger's small molecule discovery suite has a wide range of virtual screening options, advanced calculations for binding affinity estimation, analyses for target structure and binding modes, utilities for ligand structures, and visualization and automated workflow tools. The tools used in this study include docking module Glide[54], ligand preparation module LigPrep[55], the unified interface Maestro[56], protein structure prediction tool Prime[57], and the Protein Preparation Wizard.[58]

MolSoft Internal Coordinates Mechanics Software Version 3.8-4

Internal Coordinates Mechanics (ICM)[59] is a modeling software with features such as protein structure analysis, 3D interactive editor, crystallographic analysis tools, small molecule docking, protein-protein docking, protein structure prediction tools, electrostatic analysis, chemistry tools and molecular graphics. The chemistry tools of ICM were applied to the ligand sets in this study.

MODELLER Release 9.15

MODELLER[44] is a program used for homology, also known as comparative, modeling of protein 3D structures. It implements comparative protein structure modeling by satisfaction of spatial restraints, and can perform additional tasks such as *de novo* modeling of loops in protein structures, optimization of various models of protein structure with respect to a flexibly defined objective function, multiple alignments of sequences and structures, searching of databases, and comparison of protein structures.

3.1.2. Databases

The Universal Protein Resource (UniProt)

UniProt[60] (www.uniprot.org) is a comprehensive resource for protein sequence and annotation data, where many entries are derived from genome sequencing projects. The database contains information on structure and function of proteins acquired from research

literature. UniProt is a collaboration between the European Bioinformatics Institute, the Swiss Institute of Bioinformatics, and the Protein Information Resource.

The Protein Data Bank (PDB)

The PDB[61] (www.rscb.org) archive is a worldwide repository of information on the 3D structure of large biological molecules, including proteins and nucleic acids. Structures submitted to the PDB are obtained by X-ray crystallography, NMR spectroscopy, or cryo-electron microscopy. The PDB is a key resource in structural biology, and most scientific journals require submission of solved structures to the PDB.

The International Union of Basic and Clinical Pharmacology (IUPHAR)/British Pharmocological Society (BPS) Guide to PHARMACOLOGY database

The IUPHAR/BPS Guide to PHARMACOLOGY database[62] (www.guidetopharmacology.org) contains quantitative information on drug targets and the prescription medicines and experimental drugs that act on these targets. The database includes all the GPCRs, voltage-gated ion channels, nuclear receptors, and ligand-gated ion channels that are known in the human genome.

3.2. Homology modeling of 5-HTRs

The homology modeling of human 5-HT_{1A}R, 5-HT_{2A}R, 5-HT_6R, and 5-HT_7R was based on multiple templates, constructed using MODELLER. The models were evaluated with the ModFOLD4 server to assess the global 3D structure quality of the proteins.

3.2.1. Templates

The family A GPCR templates considered for the homology modeling of the targets were crystal structures from β_1-adrenergic receptor (β_1-AR), β_2-adrenergic receptor (β_2-AR), 5-HT_{1B}R, 5-HT_{2B}R, and dopamine D_3 receptor (D_3R) (table 2). All crystal structures were of the human proteins, except the turkey (*Meleagris gallopavo*) β_1 adrenergic receptor protein, and all were obtained by X-ray crystallography. The decision to use two structures of the β_1 adrenergic receptor as templates was made by the desire to have homology models based on structures of the same receptor bound to a ligand in two different conformational states (agonist/antagonist bound receptor conformation). In addition, the β_1-AR has successfully been used by others as template when modeling 5-HTRs.[63] In addition to using the β_1-AR as

template (PDB ID 2Y03 and PDB ID 2YCW) for all the targets, the template with the highest sequence identity for each target was chosen as basis for homology models; 5-HT_{1B}R (PDB ID 4IAR) for 5-HT_{1A}R, 5-HT_{2B}R (PDB ID 4IB4) for 5-HT_{2A}R, and dopamine D_3R for 5-HT_7R. As the β_1-AR had the highest sequence identity to 5-HT_6R, no additional models were made of this receptor.

Table 2. An overview of considered templates for homology modeling of human 5-HT_{1A}, 5-HT_{2A}, 5-HT_6, and 5-HT_7 receptors, and their sequence identity to the targets

PDB ID	Receptor	Ligand bound state	G-prot. coupling	5-HT_{1A}R (G_i/G_0)	5-HT_{2A}R (G_q/G_{11})	5-HT_6R (G_S)	5-HT_7R (G_S)
2Y00	β_1-AR	Agonist	G_S	31%	24%	29%	26%
2Y03	β_1-AR	Agonist	G_S				
2YCW	β_1-AR	Antagonist	G_S				
2RH1	β_2-AR	Agonist	G_S	28%	23%	24%	23%
3PDS	β_2-AR	Antagonist	G_S				
4IAR	5-HT_{1B}R	Agonist	G_i/G_0	38%	23%	21%	28%
4IAQ	5-HT_{1B}R	Agonist	G_i/G_0				
4IB4	5-HT_{2B}R	Agonist*	G_q/G_{11}	28%	40%	21%	23%
4NC3	5-HT_{2B}R	Agonist*	G_q/G_{11}				
3PBL	D_3R	Antagonist	G_i/G_0	32%	26%	27%	29%

*Biased agonist, activates arrestin pathway.

3.2.2. Amino acid alignment

An amino acid alignment of all 5-HTRs and considered templates were made with the Multiple Sequence Viewer tool of the Maestro interface.[56] α-helices and extra- and intracellular loops were aligned and minor errors were corrected manually, thus making sure that all conserved residues within family A GPCRs were aligned properly. All homology models were constructed on the basis of this multiple alignment.

3.2.3. Constructing models with MODELLER

Homology models were constructed using the program MODELLER, which implements an automated approach to protein structure modeling by satisfaction of spatial restraints.[64] The input to the program is an alignment of the sequence of the target protein with the template protein, and the output is a 3D model of the target (figure 7)

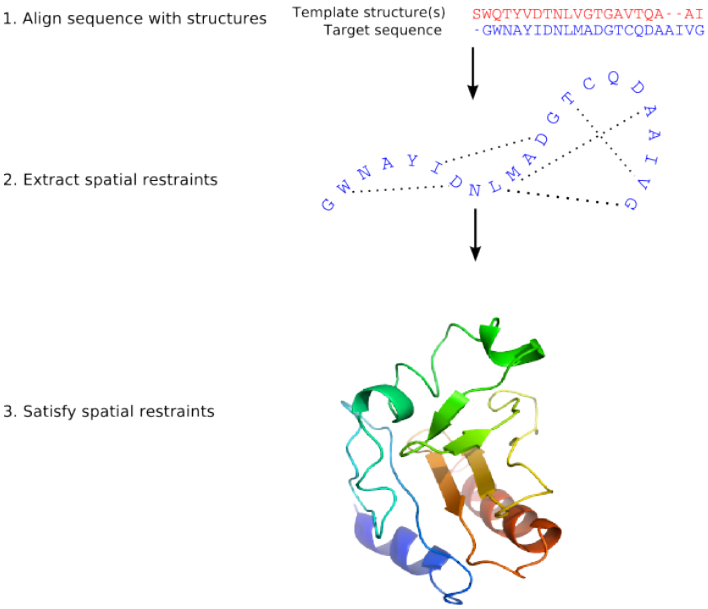


Figure 7. MODELLER comparative modeling by satisfaction of spatial restraints.[61] First, the template is aligned with the target sequence. Second, spatial features such as Cα - Cα atom distances, hydrogen bonds, and mainchain and sidechain dihedral angles, are transferred from the template to the target. This obtains spatial restraints on its structure. Third, the 3D model is obtained by satisfying the restraints as well as possible.[61]

The first step in constructing models by MODELLER is the calculation of distance and dihedral restraints from the target sequence from its alignment with the template 3D structure. The form of these restraints comes from a statistical analysis of the relationships between many pairs of homologous structures.[65] This analysis relies on a database containing 416 proteins, where correlations such as equivalent Cα - Cα atom distances and dihedral angles of mainchain residues from two related protein structures are quantified. An important feature of the MODELLER method is that the spatial restraints are obtained empirically, from a database of protein structure alignments.[65]

The next step combines spatial restraints and force field energy terms enforcing proper stereochemistry into an objective function.[66] Finally, the last step obtains the model by optimizing the objective function, employing methods of conjugate gradients and molecular dynamics with stimulated annealing.[67]

MODELLER can calculate several slightly different models by varying the initial structure, and the variability among these models can be used to estimate the errors in the corresponding regions of the fold.[44]

The inputs to the MODELLER software were alignments of the targets with their chosen templates (section 3.2.1.), extracted from the previously described multiple alignment (section 3.2.2.). N- and C-terminals, and ICL3, were not modeled because the crystal structures of the templates had limited structural information of these regions.

For every target-template alignment, 100 models were generated to increase the probability of achieving correct structures of the targets. Models were ranked by their discrete optimized protein energy (DOPE)[68] scores, and the 10 top-ranked models were kept for further evaluation and subsequent molecular docking. DOPE is a statistical potential included in the MODELLER software, used to assess the energy of the generated protein models.

3.2.4. Evaluation of 3D quality of models with ModFOLD4 server

For further assessment of the quality of the models generated with MODELLER, all models were evaluated with the ModFOLD4 server (www.reading.ac.uk/bioinf/ModFOLD/).[69] The server estimates both the global and local (per-residue) quality of 3D protein models, by predicting their similarity to the native structure.

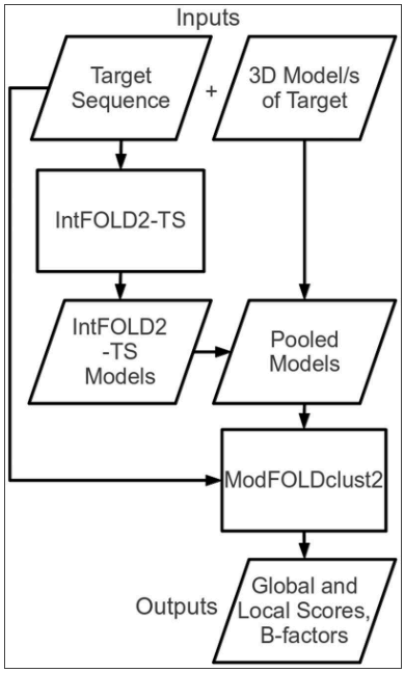


Figure 8. Flow chart outlining the principal stages of the ModFOLD4 server prediction pipeline.[66]

The input to the ModFOLD4 server is the full sequence of the target protein in the FASTA format and all the generated models for each target (figure 8). In short, the IntFOLD-TS protocol[70] generates multiple template models from the target sequence, which are pooled together with the input models. All models are then evaluated with the ModFOLDclust2 method[71], by pairwise structural comparisons of multiple models, often referred to as clustering. In ModFOLDclust2, global scores are calculated for each protein model. P-values, representing the probability that each model is incorrect, are calculated from the global scores.[69]

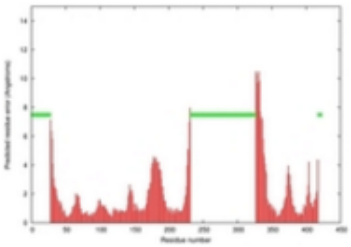
Model name	Confidence and P-value	Global model quality score	Residue error plot (click image for large version)	3D view of residue error (click image for large version)
prepd_5ht1a_2ycw_B99990078.pdb	HIGH: 1.038E-3	0.6510		

Figure 9. Example output from the ModFOLD4 server.

The output of the server is a table ranking models by global quality score (figure 9). The global model quality score ranges between 0 and 1, where scores less than 0.2 indicate there may be incorrectly modeled domains, and scores greater than 0.4 in general indicate more complete and confident models, which are highly similar to the native structure. P-values represent the probability that each model is incorrect, and P-values below 0.01 indicate high confidence of the model being similar to the native protein. Per-residue error plots are made from the predicted distance between Cα atoms of the model and the equivalent Cα atoms in the native structure, and are useful in identifying problematic regions in the modeled structure.

3.3. Ligand sets

Sets of ligands with known affinity and action for the target receptors were obtained from the Guide to PHARMACOLOGY database. The database contains ligands in the form of single molecule-input line-entry system (SMILES), and these were converted into two-dimensional form by the ICM software. The ligands were divided into sets of agonists and antagonists for each target receptor. An essential interaction between ligands and 5-HTRs is that between a protonated amine in the ligand and the receptor residue D3.32, and to ensure this interaction could be set as constraint in the docking, a protonation of an amine in each ligand were performed with ICM.

Novel SERT ligands with affinity for the target receptors were supplied from the author of the study identifying these compounds (supplementary information II-IV).[53] The ligands were imported to ICM and protonated like the aforementioned ligand sets of known agonists and antagonists.

All ligand sets were imported to the Maestro interface and prepared with the LigPrep wizard. LigPrep are designed to prepare high quality, all-atom 3D structures for large numbers of drug-like molecules, starting with 2D or 3D structures. The LigPrep process consists of steps that perform conversions, apply corrections to the structures, generate variations on the structures, eliminate unwanted structures, and optimize the structures. Many of the steps are optional, and are specified in the LigPrep panel.[55] The ligands were prepared with default settings, with the exception of no change in ionization, and no generation of tautomers. The force field applied was OPLS3[72], and specified chiralities were retained at one per ligand.

3.4. Docking

Docking studies were performed with the Virtual Screening Workflow (VSW) tool of the Schrödinger software. It is designed to run an entire sequence of jobs for screening large collections of compounds against one or more targets. As this experiment involved docking of ligands into multiple conformations of models, an automated approach was preferred. VSW includes multiple modules of the Schrödinger software including LigPrep, QikProp and Glide, but only Glide was employed in the docking studies. Glide is a method for docking ligands into rigid 3D structures of proteins; it does a systematic search of the conformational, orientational and positional space of the docked ligand.[54] The method evaluates hundreds of possible conformational poses of the ligand in multiple stages, each stage further refining the best candidates by energy optimization and Monte Carlo sampling. Selection of the best-docked pose uses a model energy function that combines empirical and force field-based terms.[54]

GlideScore is the empirically derived scoring function of Glide, used to score the best docking pose. It is based on the ChemScore function[73], adding new rewards and penalties, and modifying other terms.

The protein structures were prepared with the one-step protein preparation tool of Maestro, an automated approach to the Protein Preparation Wizard.[58] Protein preparation involves steps including assignment of hydrogens, removal of unwanted water molecules, and protein minimization.

A prerequisite of docking studies is the definition of a binding site, a region on the protein where ligands bind. Binding sites defined with Maestro are referred to as receptor grids. Receptor grids contain an inner and outer box, and the user specifies their placement and size. In this study, grids with an inner box of 10 Å and an outer box of 25 Å were defined around D3.32, the residue involved in the essential interaction of ligand binding to 5-HTRs.
Because there is no automated approach to generate receptor grids for multiple models in Maestro, a script was used to do this in the command line on Linux OS.

To ensure that all the docking poses had the interaction of a protonated amine in the ligands and D3.32, a constraint was applied to the docking procedure. The constraint anchors the protonated amine of the ligand to the carboxylate atom of D3.32, but otherwise move freely in the specified receptor grid. Because the constraint option is not available in the graphical interface of VSW, two scripts were made to include the constraint in the docking, and VSW was run from the command line. All docking runs were launched with standard precision (SP) and the OPLS3 force field.

Ligand sets of known antagonists and agonists were docked into each of the target receptor models to examine if differences in binding mode could be identified dependent on ligand action, and if there was a preference to receptor models based on specific templates. Novel SERT ligands with affinity towards the target receptors were subsequently docked into all models.

To determine if any of the models had a binding site conformation with agonist or antagonist preferences to, both agonists and antagonists were docked for each target and ranked by the GlideScore. The ranking and statistical analysis were performed by the Boltzmann-enhanced discrimination of receiver operating characteristics (BEDROC) metric.[74] BEDROC was originally developed to evaluate virtual screening, by applying statistical metrics to assess the ability of a model to distinguish between active compounds and decoys (compounds that are physiochemically similar to actives, but are presumed not to bind). The full statistical strength of BEDROC was not applicable in this study, as the number of ligands was too low, and its use was mainly to examine if any of the models ranked agonists higher than antagonists, or vice versa.

3.5. Structural Interaction Fingerprint

Structural Interaction Fingerprint (SIFt)[51] was used to analyze the ligand-protein interactions of the docked ligands in their target receptor. The method is based on a 9-digit binary interaction pattern that describes physical ligand-protein interactions in structures and models of ligand-protein complexes (figure 10).[75] SIFt is a rapid and computationally efficient method, suited to process large amounts of data. Traditional analysis of ligand-protein interactions, such as visual inspection, proves inadequate when dealing with the massive amount of structural information generated from docking studies. SIFt translates 3D structural binding information into a one-dimensional binary string.

The basic algorithm of SIFt is to find amino acids around the bound ligand, and to determine type of interaction from distance and residue types. Figure 10 displays the binary string of the SIFt, where each digit corresponds to an interaction between ligand and protein residue. The first digit represents any interaction between the ligand and a residue, the second and third digit specify if the interaction is to a sidechain or backbone atom, and the last six digits specify the type of interaction.

This study focused only on the presence of an interaction, not taking the type of interaction into consideration. All the residues found to have an interaction to each ligand were recorded and quantified, resulting in a table describing the occurrence of ligand interaction for each residue. No interaction was recorded as 0, and any type of interaction was recorded as 1, as SIFt is based on binary strings. A SIFt score of 1.0 for a residue indicates that 100% of the ligands interacted with that residue.

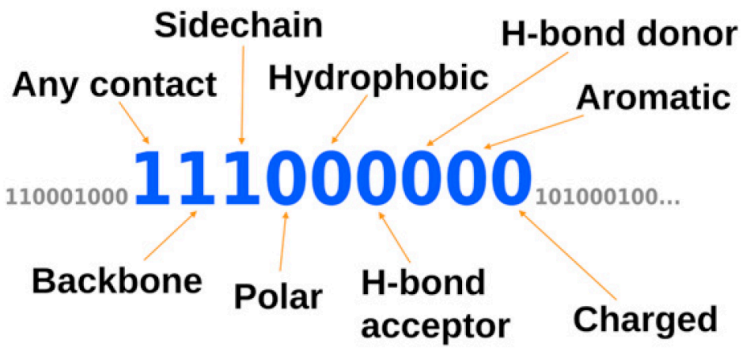


Figure 10. The 9-digit binary interaction pattern describing residue-ligand interactions.[72]

4. Results

4.1. Workflow of study

Computational methods were used to construct models of human 5-HT_{1A}R, 5-HT_{2A}R, 5-HT_6R and 5-HT_7R. Models with a preference for either agonists or antagonists were selected from docking studies of known agonists and antagonists for the targets, and docking of novel multimodal compounds into these models was the basis for the prediction of mode of action for the compounds. A workflow of the study is presented in figure 11.

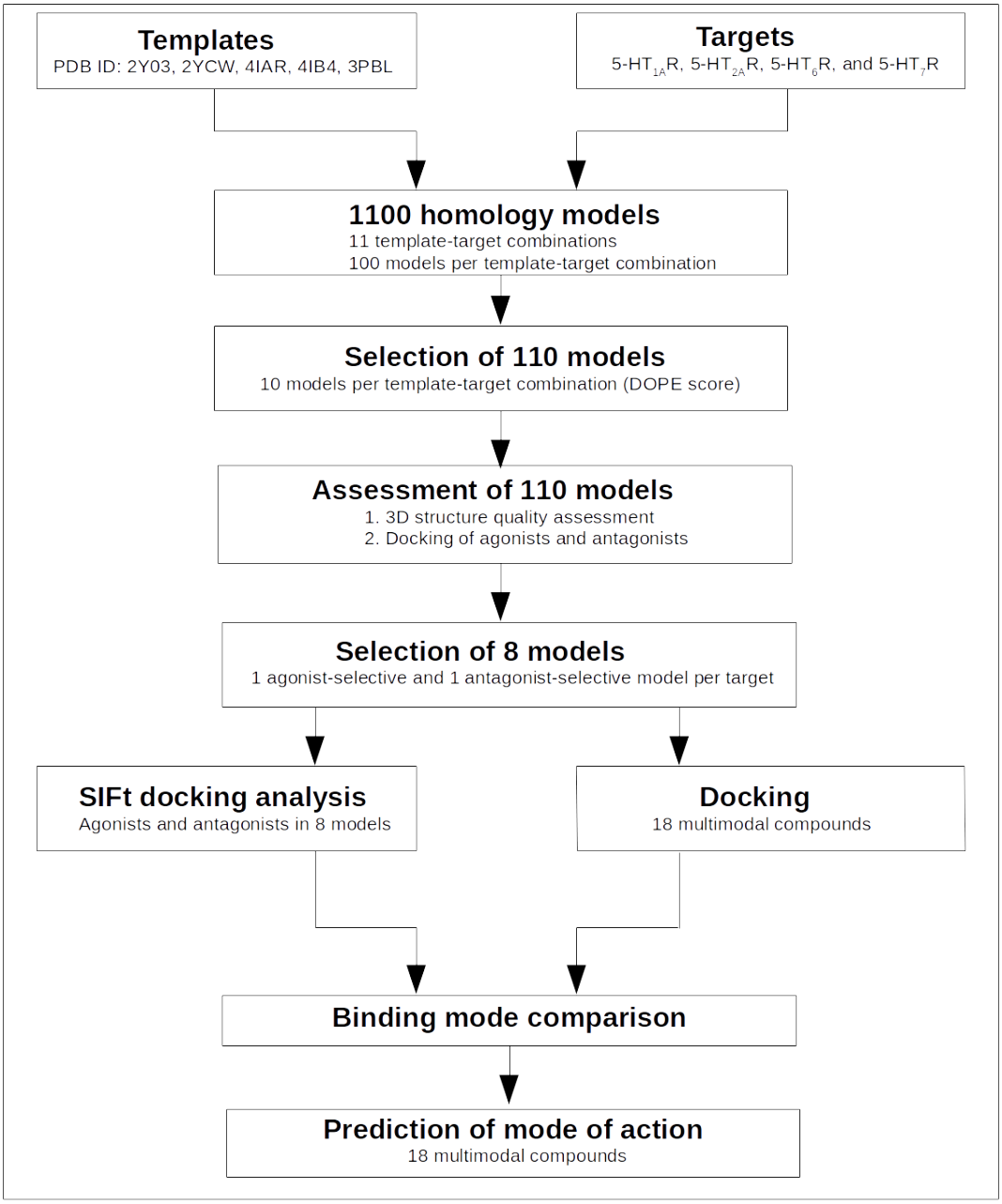


Figure 11. Workflow of study.

4.2. 3D structure quality assessment of 110 models

The 10 best models, based on DOPE score, for each template-target combination were assessed with the ModFOLD4 server. Global scores ranged from 0.54 to 0.65, which gave P-values from 3.2 x 10^{-3} to 1.1 x 10^{-3}. The per-residue plots showed high quality for core helical residues, while residues at the end of the helices and in the loop regions had lower quality.

4.3. Binding site residues of targets and templates

An extensive mapping of possible binding site residues in all template and target receptors was made (table 3), to determine important residues for ligand binding. The binding site residues were identified by displaying all residues within 5Å of the cocrystallized ligands in the crystal structures of the templates, and the corresponding residues in the targets were identified using the Ballesteros-Weinstein annotation. In addition, ECL2 residues in the targets were discovered by inspecting residues within 5Å of docked ligands.

Table 3. Binding site residues for templates and targets, extracted from residues within 5Å of cocrystallized ligands in templates and docked ligands in targets

	Templates				Targets			
Ballesteros-Weinstein	**β1-AR (ago)**	**β1-AR (ant)**	**5-HT_{1B}R**	**5-HT_{2B}R**	**5-HT_{1A}R**	**5-HT_{2A}R**	**5-HT_{6}R**	**5-HT_{7}R**
2.62	A99*	A99*	Y107	L112*	A94	M132	M84	S139
2.64	L101*	L101*	Y109	T114*	Y96	T134	N86	T141
3.28	W117	W117	W125	W131	F112	W151	W102	F158
3.29	T118	T118	L126	L132	I113	I152	T103	I159
3.32	D121	D121	D129	D135	D116	D155	D106	D162
3.33	V122	V122	I130	V136	V117	V156	V107	V163
3.36	V125	V125	C133	S139	C120	S159	C110	C166
3.37	T126*	T126	T134	T140	T121	T160	S111	T167
3.40	I129*	I129*	I137	I143*	I124	I163	I114	I170
4.56	V172*	V172*	I180	I186	I167	I206	A157	I213
ECL2	F201	F201	C199	C207		S226		
ECL2		T203	V200	V208	C187	C227		C231
ECL2			V201	L209	T188	L228	L182	L232
ECL2			V203	T210	I189	L229	L183	I233
ECL2				K211		A230	A184	S234
ECL2				F214				
5.36	R205*	R205*	I206*	G215	H193	D232	L186	F237
5.38	Y207	Y207	Y208*	F217	Y195	F234	F188	Y239
5.39	A208*	A208	T209*	M218	T196	V235	V189	T240
5.42	S211	S211	S212	G221	S199	G238	A192	S243
5.43	S212	S212	T213	S222	T200	S239	S193	T244

5.46	S215	S215	A216	A225	A203	S242	T196	A247
6.48	W303	W303	W327	W337	W358	W336	W281	W340
6.51	F306	F306	F330	F340	F361	F339	F284	F343
6.52	F307	F307	F331	F341	F362	F340	F285	F344
6.55	N310	N310	S334	N344	A365	N343	N288	S347
6.58	N313*	N313*	M337	L347	L368	A346	Q291	R350
6.59	V314*	V314*	ECL3*	V348	ECL3	V347	A292	ECL3
7.32	D332*	D332*	L348	Q359	T379	G359	P299	L363
7.35	F325*	F325*	F351	L362	G382	L362	F302	E366
7.36	V326*	V326*	D352	E363	A383	N363	D303	R367
7.39	N329	N329	T355	V366	N386	V366	T306	L370
7.43	Y333	Y333	Y359	Y370	Y390	Y370	Y310	Y374

*Residues in red were not within 5Å of the cocrystallized ligand in the specified template receptor.

4.4. Selected agonist/antagonist models

Two models of each target were selected, one model with a preference for agonists and one model with a preference for antagonists (table 4). The selection of these models was based on docking scores of known agonists and antagonists for the different receptors, and the BEDROC metric was used to identify models that scored agonists better than antagonists, and vice versa.

Table 4. Selected ligand-preference models for each target, based on BEDROC scores

Model	Template (receptor)	BEDROC	# Ligands*
5-HT_{1A}-agonist	2Y03 (β_1-agonist)	0.959	67 agonists
5-HT_{1A}-antagonist	4IAR (5-HT_{1B}-agonist)	0.983	44 antagonists
5-HT_{2A}-agonist	4IB4 (5-HT_{2B}-agonist)	0.770	47 agonists
5-HT_{2A}-antagonist	4IB4 (5-HT_{2B}-agonist)	0.957	66 antagonists
5-HT_6-agonist	2Y03 (β_1-agonist)	0.990	19 agonists
5-HT_6-antagonist	2Y03 (β_1-agonist)	0.679	33 antagonists
5-HT_7-agonist	2YCW (β_1-antagonist)	0.944	17 agonists
5-HT_7-antagonist	2YCW (β_1-antagonist)	0.957	36 antagonists

*# Ligands indicate the number of known agonists or antagonists docked into the model

4.5. SIFts of known agonists and antagonists in 8 selected models

SIFt was employed to the docking of known agonists and antagonists in the targets to determine important binding site residues, and to suggest specific residues involved in either agonist or antagonist binding in the receptors (table 5.1-5.4). The scores for the residues were color-coded by the occurrence of contact with the ligands, from green (1.0, 100% of ligands in contact with residue) to yellow (0.5, 50 % of ligands in contact with residue). Only scores ≥0.5 were kept.

Table 5.1. SIFt of known agonists and antagonists in selective models of 5-HT_{1A}R

Ballesteros-Weinstein	5-HT_{1A}R residue	Agonist-selective model	Antagonist-selective model
2.60	M92		0.62
2.62	A94		0.71
2.64	Y96	0.57	0.55
3.26	D110		0.76
3.27	L111		0.83
3.28	F112	0.72	
3.29	I113	0.98	1
3.30	A114		1
3.32	D116	1	0.83
3.33	V117	1	0.57
3.36	C120	0.92	
4.60 4.59	P171		0.55
ECL2	M172		0.86
ECL2	W175		0.74
ECL2	R176	0.68	
ECL2	T177		0.5
ECL2	P178		0.88
ECL2	R181		0.64
ECL2	A186		0.57
ECL2	C187	0.54	
ECL2	T188	0.94	0.71
ECL2	I189		1
ECL2	S190	0.88	0.83
5.38	Y195	0.92	
5.39	T196	0.68	0.5
5.42	S199	0.88	
5.43	T200	0.82	

5.46	A203	0.66	
6.48	W358	0.83	0.5
6.51	F361	1	
6.52	F362	0.95	1
6.55	A365	0.51	
6.56	L366		1
7.36	A383	0.51	
7.39	N386	1	
7.43	Y390	1	

Table 5.2. SIFt of known agonists and antagonists in selective models of 5-HT_{2A}R

Ballesteros-Weinstein	5-HT_{2A}R residues	Agonist-selective model	Antagonist-selective model
2.61	S131		0.5
2.64	T134		0.55
3.28	W151	0.63	0.86
3.29	I152	0.8	0.91
3.32	D155	1	1
3.33	V156	1	1
3.36	S159	0.93	0.91
3.37	T160	0.73	0.62
4.61 4.60	I210		0.55
ECL2	C227		0.83
ECL2	L228	0.54	1
ECL2	D231		0.52
5.38	F234	0.9	0.79
5.39	V235	0.95	0.86
5.42	G238	0.78	0.66
5.43	S239	0.83	0.69
5.46	S242	0.8	0.69
6.48	W336	0.93	0.86
6.51	F339	1	1
6.52	F340	0.88	0.81
6.55	N343	0.59	0.69
7.39	V366	0.98	1
7.43	Y370	1	1

Table 5.3. SIFt of known agonists and antagonists in selective models of 5-HT$_6$R

Ballesteros-Weinstein	5-HT$_6$R residue	Agonist-selective model	Antagonist-selective model
3.28	W102	0.77	0.96
3.29	T103	0.92	0.92
3.32	D106	1	1
3.33	V107	1	1
3.36	C110	0.85	0.81
4.61 4.60	L162	0.85	0.54
ECL2	H167	0.62	0.5
ECL2	R181		0.73
ECL2	L182	1	1
ECL2	L183		0.77
ECL2	A184	1	1
5.38	F188	1	1
5.39	V189	0.85	0.92
5.42	A192	1	0.81
5.43	S193	0.69	0.85
5.46	T196		0.77
6.48	W281	1	0.81
6.51	F284	1	1
6.52	F285	0.77	0.92
6.55	N288	0.85	0.92
7.35	F302		0.77
7.39	T306	0.92	1
7.43	Y310	1	1

Table 5.4. SIFt of known agonists and antagonists in selective models of 5-HT$_7$R

Ballesteros-Weinstein	5-HT$_7$R residue	Agonist-selective model	Antagonist-selective model
2.61	V138	0.57	0.6
2.64	T141		0.53
3.28	F158	0.86	0.73
3.29	I159	0.93	1
3.32	D162	1	1
3.33	V163	1	1
3.36	C166	0.93	1
ECL2	L232	0.64	0.6
ECL2	I233	1	1
ECL2	Q235	1	0.87
5.38	Y239	0.79	0.6
5.39	T240	0.79	0.73
5.42	S243	0.86	0.73
5.43	T244	0.93	0.8
5.46	A247	0.79	0.6
6.48	W340	0.93	1
6.51	F343	1	1
6.52	F344	1	1
6.55	S347		0.6
6.58	R350		0.53
7.39	L370	1	1
7.43	Y374	1	1

4.6. Favorable binding modes

Several ligand-receptor complexes from the docking of the multimodal compounds in targets with experimentally determined affinity were visually inspected. Comparisons between high ranked compounds and low ranked compound in the docking, correlated with experimental affinity (supplementary information IV), were made to highlight important residues and interactions, revealed by docking poses of the compounds in the targets (figures 12-15).

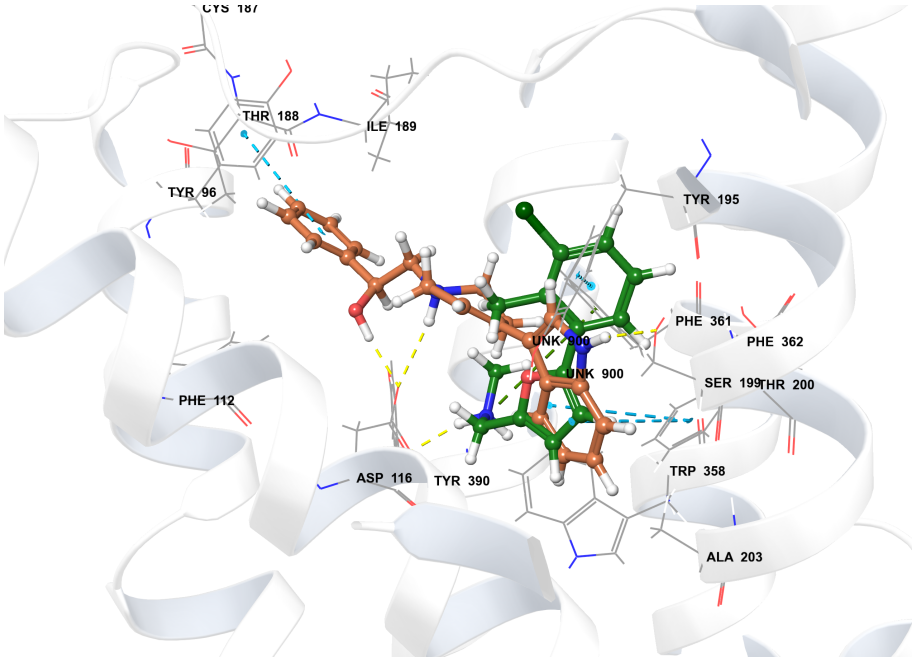


Figure 12. Docking pose of high ranked compound T6125232 (orange) and low ranked compound 9034414 (green) in 5-HT_{1A}R. Helices are represented in light gray, hydrogen bonds in yellow dashed lines, direct aromatic interactions in light blue dashed lines.

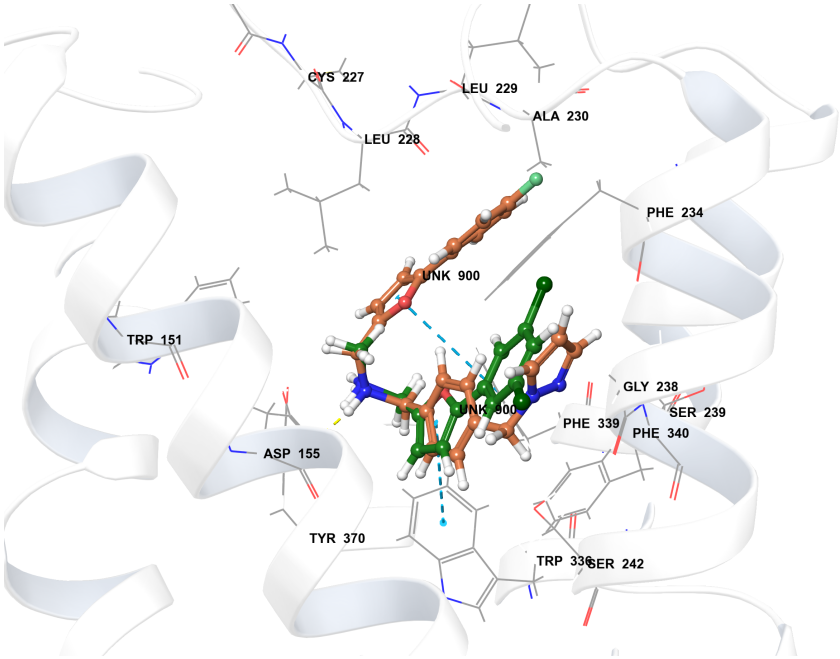


Figure 13. Docking pose of high ranked compound T6209417 (orange) and low ranked compound ASN13153175 (green) in 5-HT_{2A}R. Helices are represented in light gray, hydrogen bonds in yellow dashed lines, direct aromatic interactions in light blue dashed lines.

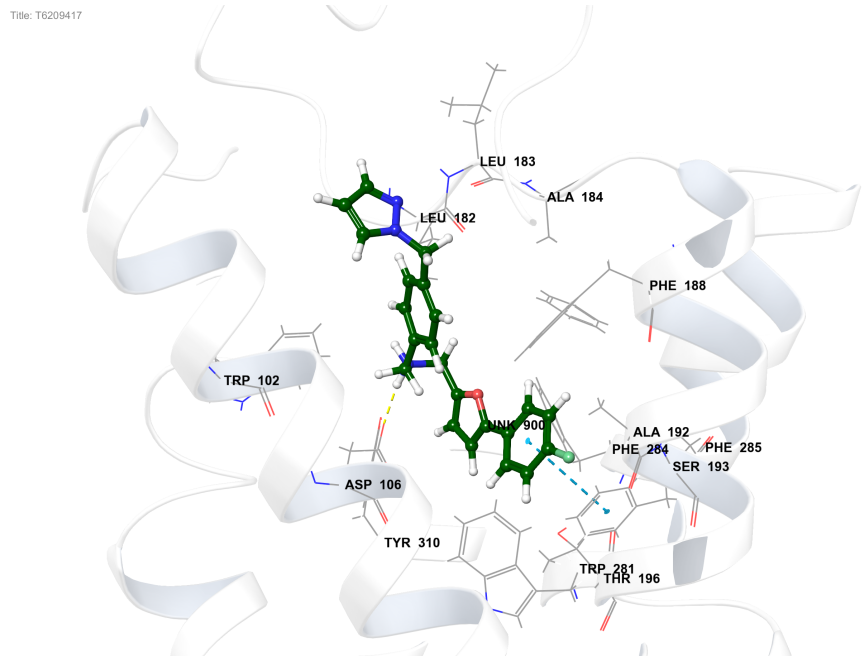


Figure 14. Docking pose of compound T6209417 in 5-HT_6R (low ranked). Helices are represented in light gray, hydrogen bonds in yellow dashed lines, direct aromatic interactions in light blue dashed lines.

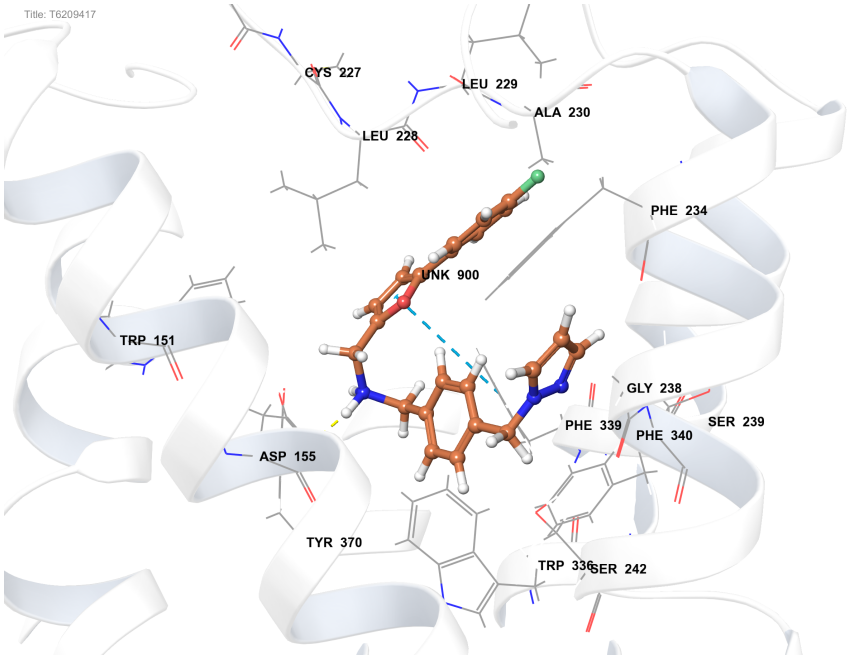


Figure 15. Docking pose of compound T6209417 in 5-HT_{2A}R (high ranked). Helices are represented in light gray, hydrogen bonds in yellow dashed lines, direct aromatic interactions in light blue dashed lines.

4.7. Prediction of mode of action for novel multimodal compounds

A mode of action was suggested for selected multimodal compounds for each target receptor, based on their preference to agonist- or antagonist-selective models, and their conformation in the binding site and interactions with important residues in the ligand-receptor complex.

4.7.1. Agonist/antagonist-model preference for 18 novel multimodal compounds

The 18 novel multimodal compounds were docked into the targets receptors. Compounds with an experimentally determined affinity for a target (supplementary information IV) were docked into agonist/antagonist-selective models, to see if there was a preference to either of them (table 6). The conformation of each compound in the binding site was examined to observe if a polar part of the compound was in contact with TM5.

Table 6. Agonist/antagonist-selectivity for 18 novel multimodal compounds

Cluster[a]	Ligand	Affinity (K_i)[b]	Model	Polar contact with TM5
C01	ASN13153175	292 nM	5-HT_6R-agonist	Yes
C01	9034414	846 nM	5-HT_{2A}R-agonist	Yes
		568 nM	5-HT_6R-agonist	Yes
C01	9013195	839 nM	5-HT_{2A}R-antagonist	Yes
		117 nM	5-HT_6R-agonist	Yes
		730 nM	5-HT_7R-agonist	Yes
C01	T6209417	655 nM	5-HT_{2A}R-agonist	Yes
C01	9134052	187 nM	5-HT_{1A}R-agonist	Yes
C03	5417988	860 nM	5-HT_{1A}R-antagonist	No
C04	T6125232	56 nM	5-HT_{1A}R-agonist	Yes
		217 nM	5-HT_{2A}R-antagonist	No
		569 nM	5-HT_6R: Not docked*	-
		314 nM	5-HT_7R-agonist	No
C04	T6275452	8 nM	5-HT_{1A}R-agonist	Yes
		101 nM	5-HT_6R-agonist	Yes
		126 nM	5-HT_7R-antagonist	No
C06	9066608	742 nM	5-HT_6R-agonist	No
		819 nM	5-HT_7R-antagonist	No

C06	7989485	818 nM	5-HT_{2A}R-antagonist	No
		420 nM	5-HT_6R-agonist	Yes
C07	5458751	596 nM	5-HT_6R-antagonist	No
C07	5456380	265 nM	5-HT_{2A}R-antagonist	No
		856 nM	5-HT_6R-antagonist	Yes
		830 nM	5-HT_7R-agonist	Yes
C07	SYN16295816	508 nM	5-HT_{2A}R-antagonist	Yes
		640 nM	5-HT_6R-agonist	Yes
		895 nM	5-HT_7R-agonist	Yes
C07	ASN16295801	543 nM	5-HT_{2A}R-antagonist	No
C07	SYN16295876	737 nM	5-HT_{2A}R-antagonist	Yes
C12	T0502-9459	309 nM	5-HT_{2A}R-antagonist	Yes
		69 nM	5-HT_6R-antagonist	No
C12	T0503-1300	315 nM	5-HT_{2A}R-antagonist	No
		46 nM	5-HT_6R-antagonist	No
		92 nM	5-HT_7R: Not docked*	-
C13	EN300-08612	581 nM	5-HT_7R: Not docked*	-

a. Compounds of the same cluster are structurally similar (supplementary information II). Compounds were clustered using Molprint2D fingerprints and Tanimoto metrics.[53]

b. The lower K_i values give greater binding affinity.

*Not docked. No acceptable conformation found for compound in receptor.

4.7.2. Predicted agonist/antagonist compounds for target receptors

Predicted agonists and antagonists were selected for each target receptor, based on their preference to agonist- or antagonist-selective models and docking pose in the target receptor. Docking poses of the compounds in their selective models show interactions important for the prediction of mode of action for the compounds (figures 16-23).

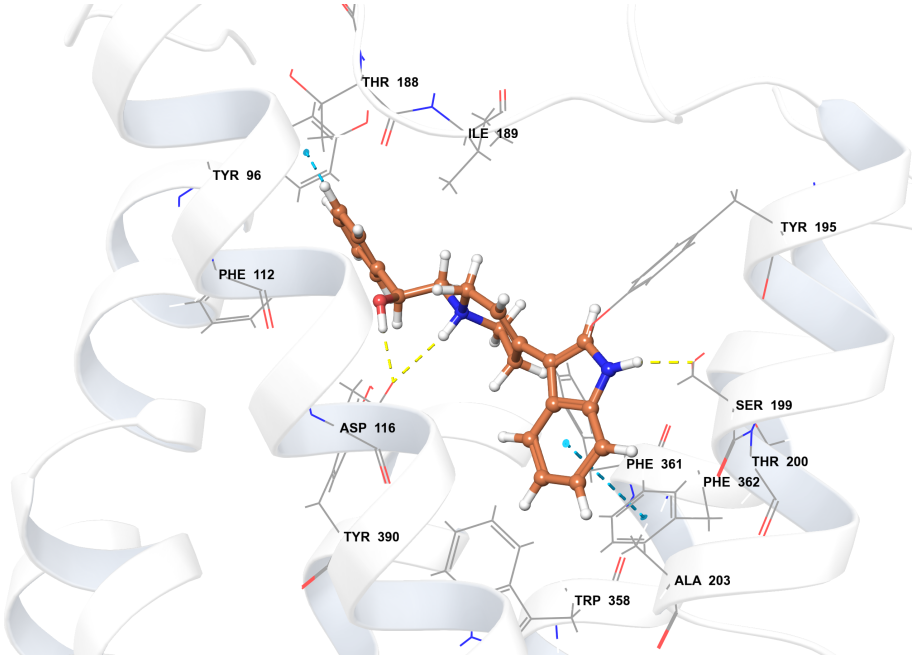


Figure 16. Docking pose of predicted agonist T6125232 in 5-HT_{1A}R. Helices are represented in light gray, hydrogen bonds in yellow dashed lines, direct aromatic interactions in light blue dashed lines.

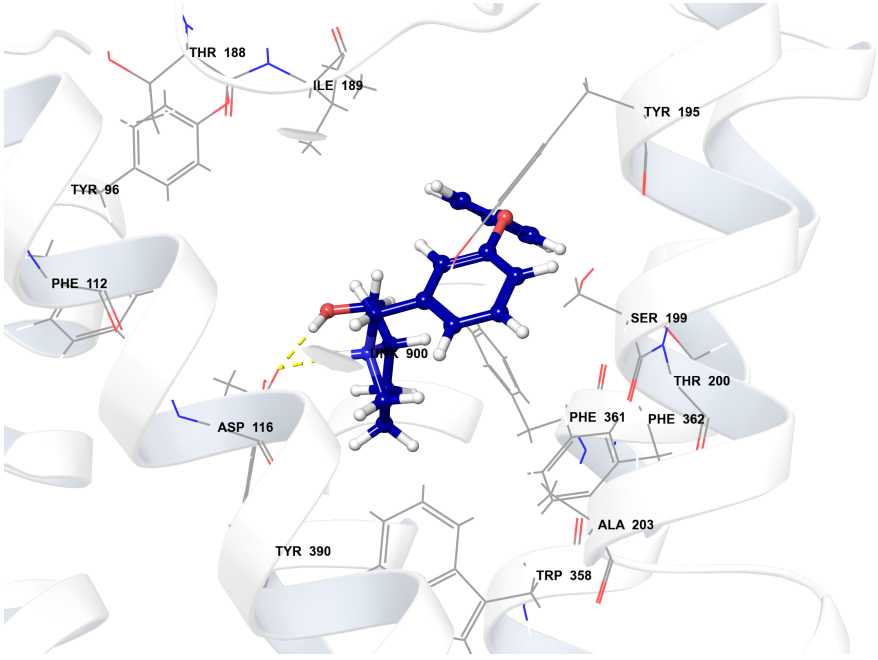


Figure 17. Docking pose of predicted antagonist 5417988 in 5-HT_{1A}R. Helices are represented in light gray, hydrogen bonds in yellow dashed lines, direct aromatic interactions in light blue dashed lines.

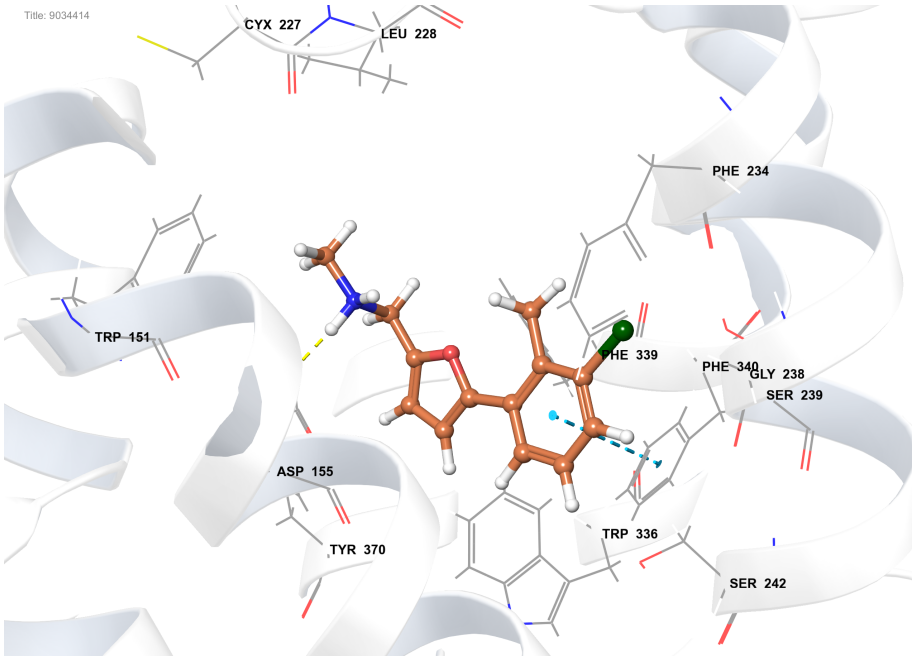


Figure 18. Docking pose of predicted agonist 9034414 in 5-HT_{2A}R. Helices are represented in light gray, hydrogen bonds in yellow dashed lines, direct aromatic interactions in light blue dashed lines.

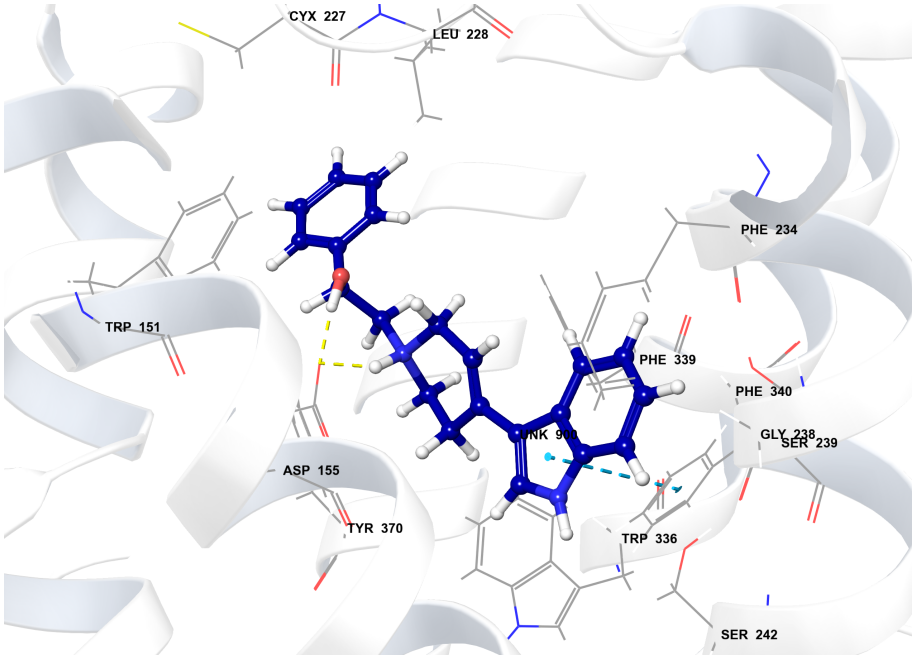


Figure 19. Docking pose of predicted antagonist T0503-1300 in 5-HT_{2A}R. Helices are represented in light gray, hydrogen bonds in yellow dashed lines, direct aromatic interactions in light blue dashed lines.

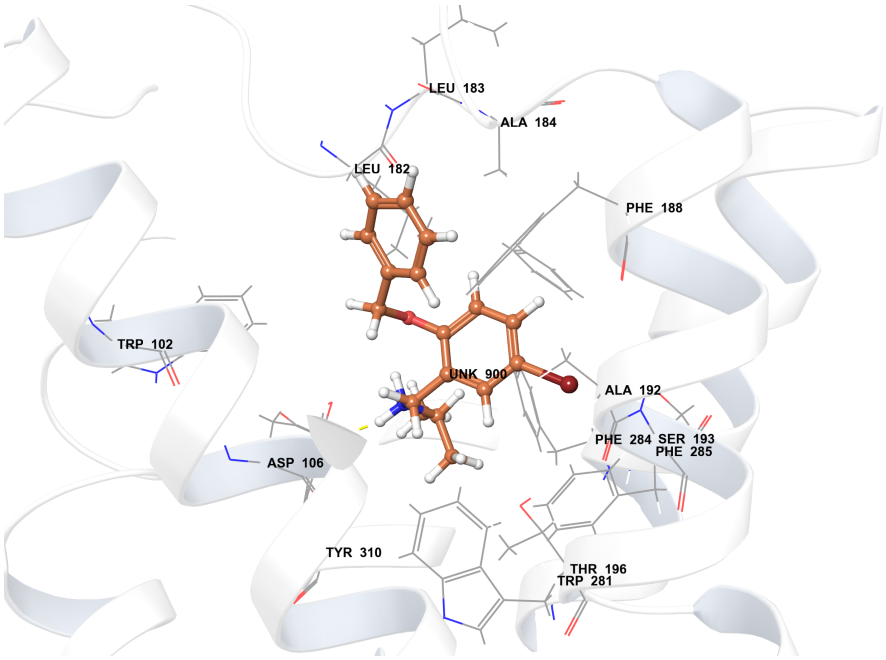


Figure 20. Docking pose of predicted agonist 7989485 in 5-HT_6R. Helices are represented in light gray, hydrogen bonds in yellow dashed lines, direct aromatic interactions in light blue dashed lines.

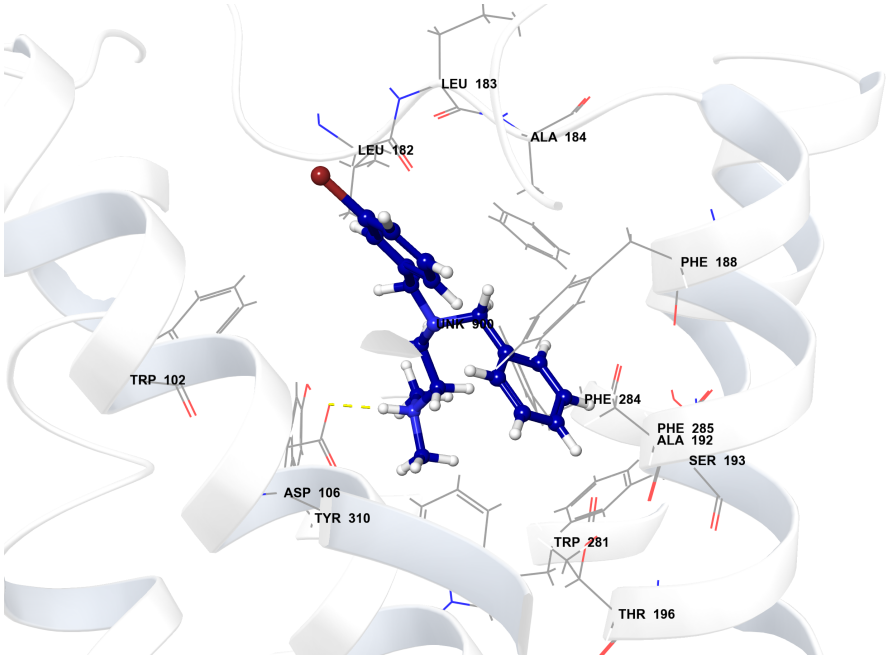


Figure 21. Docking pose of prediced antagonist 5458751 in 5-HT_6R. Helices are represented in light gray, hydrogen bonds in yellow dashed lines, direct aromatic interactions in light blue dashed lines.

,

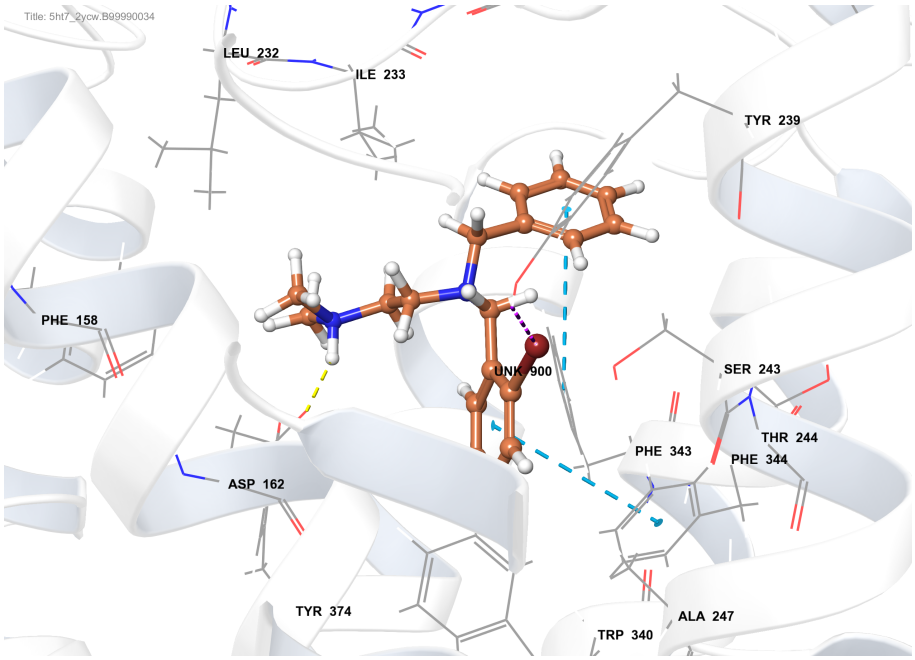


Figure 22. Docking pose of predicted agonist 5456380 in 5-HT$_7$R. Helices are represented in light gray, hydrogen bonds in yellow dashed lines, direct aromatic interactions in light blue dashed lines.

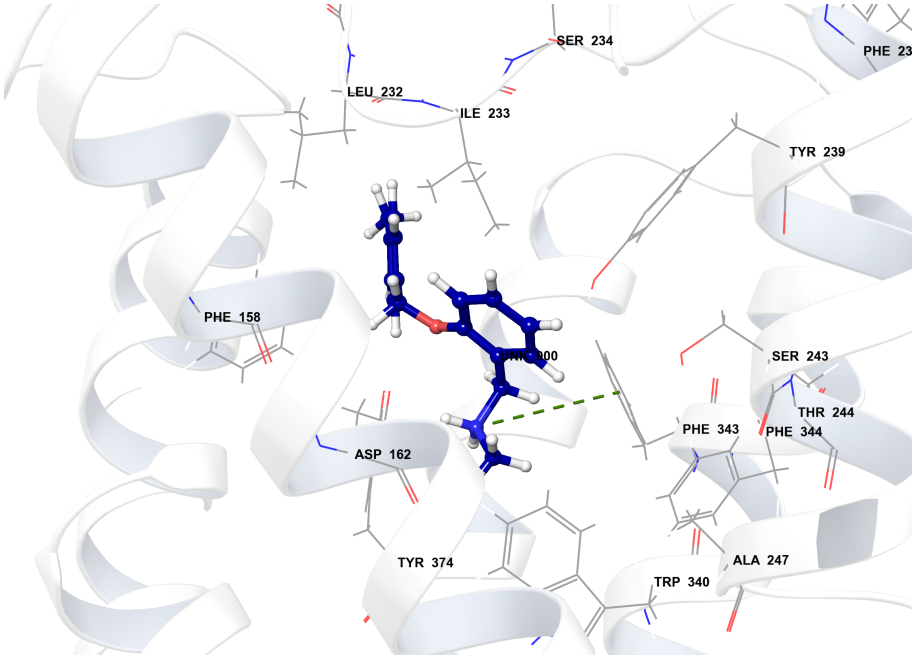


Figure 23. Docking pose of predicted antagonist 9066608 in 5-HT$_7$R. Helices are represented in light gray, hydrogen bonds in yellow dashed lines, direct aromatic interactions in light blue dashed lines.

5. Discussion

The construction of homology models of human 5-HT_{1A}R, 5-HT_{2A}R, 5-HT_6R and 5-HT_7R enabled structural analysis and interaction studies of novel multimodal compounds in the target receptors. Although the affinity to SERT and target 5-HTRs had been experimentally determined for the novel multimodal compounds, their mode of action was not established. This study aimed to give structural explanations for high and low affinity binding of the compounds for the target 5-HTRs, and to predict agonist or antagonist mode of action of the compounds.

The determination of mode of action for a compound by computational methods is an ambitious task. Compounds that are structurally similar can have completely different activity in a receptor, and a compound can have different activity in similar receptors of the same class. X-ray crystal structures of GPCRs with an agonist and an antagonist indicated slight structural differences between the agonist and antagonist states. The crystal structure of the β_1-AR in complex with the agonist isoprenaline showed approximately 1 Å contraction of the binding site (TM5 and TM7 closer to the ligand) compared with the β_1-AR structure in complex with the antagonist cyanopindolol.[25] Results from our study show that homology modeling can be a useful tool when studying proteins that do not have experimentally solved structures. Docking of the multimodal compounds in the constructed homology models was used to suggest if the compounds acted as agonists or antagonists for the different target receptors.

Evaluation of the models was performed with the ModFOLD4 server, and results showed that all the models scored in the “high confidence”-category. P-values indicated that there was less than 1% chance that the models were incorrect. Per-residue plots for the models indicated lower quality for the residues at the ends of the helices and the loop regions, than for the residues in the core of the helices. This is expected due to the lack of structural information of these regions in the template receptors. The validation of the models with the ModFOLD4 server, along with docking of known agonists and antagonists in the models, demonstrated that the models were predictive, and thus could be used for further analyses.

Agonist/antagonist-selective models were selected by their ability to rank agonists over antagonists (table 4), and vice versa. To determine this ability, the BEDROC metric was employed to analyze the ranking of agonists and antagonists docked in the target receptors.

Originally developed to evaluate virtual screening, BEDROC has its statistical strength in the first 8% of ranked compounds in virtual screening experiments with huge data sets. Considering the low number of known agonists and antagonists for the targets, the BEDROC scores were obtained on smaller data sets than the intended use of the metric.

Results from the selection of models showed that six of the eight models had BEDROC scores of >0.94, which gave these six models high probability to rank either agonists or antagonists high in docking studies. The 5-HT_{2A}R-agonist and 5-HT_6R-antagonist models scored slightly lower (0.77 and 0.68 respectively), making these two models moderately less probable to rank one type of ligand over another. All the agonist/antagonist-selective models scored sufficiently high to be used in studies to theorize mode of action for compounds not established to be agonists or antagonists.

There was no correlation between agonist/antagonist selectivity of the models and the ligand state of the templates; all targets had the same template state for each of their respective agonist/antagonist-selective models (e.g. agonist/antagonist-selective models for 5-HT_{2A}R were based on the same template, in agonist state). An examination of the crystal structures of the β_1-AR templates revealed only minor differences between agonist and antagonist states of the templates, and these differences were smaller than the structural variations achieved by generating multiple conformations with the MODELLER software. By using MODELLER to construct models, the template bias of the models was reduced significantly.

The identification of residues in the binding sites of target receptors is important in the determination of how a ligand binds to a receptor. Specific residues involved in ligand-receptor interactions are more favorable than others, and the ability of the ligand to interact with residues that promote favorable interactions is what gives the ligand its affinity for a receptor. The ability of an agonist to induce activity at a receptor might be owed to its interactions with specific residues in the receptor, whereas antagonists will inhibit activity by not interacting with these residues in the same manner, but will occupy the binding site such that the competing endogenous agonist is not able to bind. A complete overview of binding site residues in both template and target receptors (table 3) shows that some residues are conserved in all 5-HTRs, and are thus thought to be important in ligand binding. Conserved residues include D3.32, W6.48, F6.51, F6.52 and Y7.43, while residues (W/F)3.28, (V/I)3.33, (T/S)3.37, (Y/F)5.38, and (S/T)5.43 are conserved by residues of the same properties. Other residues vary among the different receptors, and these variations can be part of the ligand specificity for each receptor. The information gathered from the identification of binding site

residues, of both conserved and variable residues, was used to support the study of the binding mode of multimodal compounds in the target receptors.

SIFts were generated for the docking of known agonists and antagonists in the selective models (table 5.1-5.4), to show how many of the ligands were in contact with each residue in the binding site of each model. Residues found to have contact with ligands generally independent of mode of action include D3.32 and surrounding residues in TM3, TM5 residues (Y/F)5.38 and (T/V)5.39, aromatic residues of TM6 W6.48, F6.51, and F6.52, and TM7 residues (N/V/T/Y)7.39 and Y7.43. The only outlier of the models in terms of SIFT scores was 5-HT_{1A}R-antagonist, where the ligands had few contacts with TM5, TM6 and TM7, indicating that results from this model need to be viewed critically.

Several ECL2 residues were found to have differences in contact with agonists and antagonists, which suggests that this loop plays a role in the differentiation between agonists and antagonists in the binding site. Minor differences were found for TM5 residues (S/A/G)5.42, (S/T)5.43, and (S/T/A)5.46, where more agonists than antagonists were in contact with these residues.

Only the first of the nine digits of the SIFt was used in this study, recording only contact or no contact of the ligands with the residues. To examine the interactions more closely, especially those with residues of TM5, all the digits can be used to define the type of interaction between ligand and residue.[75]

To analyze what determined high affinity and good docking score (supplementary information IV) for a ligand to a receptor, several ligand-receptor complexes were examined in light of interactions and ligand conformation. The identification of receptor residues involved in optimal ligand binding is valuable in structure-based drug discovery.

Docking pose comparison of high ranked compound T6125232 with low ranked compound 9034414 in 5-HT_{1A}R (figure 12) shows that T6125232 had favorable interactions with more binding site residues than 9034414. Interactions for T6125232 include hydrogen bond to D3.32, aromatic interactions with residues of TM2 (Y2.64), TM3 (F3.28), TM5 (Y5.38), TM6 (W6.48, F6.51, and F6.52) and TM7 (Y7.43), and contact with residue T188 in ECL2. 9034414 is a smaller compound, and its conformation in the binding site showed that it lacked interaction with residues in ECL2, which may contribute to its low rank in 5-HT_{1A}R.

A comparison made for compounds in 5-HT_{2A}R (figure 13) supports findings for 5-HT_{1A}R, where the smaller low-ranked compound ASN13153175 failed to make aromatic interactions

with TM5, and had no contact with residues in ECL2. The binding mode of the high ranked compound T6209417 shows a better fit in the binding pocket, interacting with more residues (F5.38 and residues ECL2 residues L228, L229, and A230) than 9034414.

An examination of binding mode for compound T6209417 in 5-HT_{2A}R and 5-HT_6R (figure 14 and 15), where it is high ranked and low ranked respectively, revealed that the best binding mode in the receptors showed a different conformation for the compound in the two receptors. The compound had a more advantageous conformation in the binding site of 5-HT_{2A}R, where the best conformation had favorable interactions with several important binding site residues (F5.38 and ECL2 residues L228, L229, and A230). In 5-HT_6R, T6209417 had a conformation that extended the compound out of the binding site, thus limiting its interactions with some of the important residues.

These comparisons indicate that contact with as many residues as possible is important for the binding affinity of a ligand to 5-HTRs, and that favorable interactions include aromatic contacts with residues in multiple helices (Y/F 5.38, W6.48, F6.51, F6.52, and Y7.43), and contacts with ECL2. The findings also suggest that the binding sites of the models may not be optimal for smaller ligands, as they seem to be able to make contact with only part of the binding site. An explanation for this can come from the selection of models, where the BEDROC metric generally favors models able to dock as many ligands as possible. This implies that the binding sites of the selected models by BEDROC scores have space for the larger compounds, thus making them sub-optimal for smaller compounds unable to make contact with all important binding site residues.

The prediction of mode of action for the novel multimodal compounds was based on two observations; their preference to either agonist- or antagonist-selective models, and their ability to form polar interactions to TM5 residues. The hypothesis of activation of GPCRs by agonists includes strong polar interactions between agonists and TM5.[25–27] Thus, to predict agonist activity for a compound it is essential that it has a conformation in which polar parts of the compound make contacts with polar side chains of residues of TM5 (S/T 5.42, 5.43 and 5.46). Conversely, relying on the hypothesis suggests antagonist activity for compounds unable to form polar interactions with residues of TM5.

Results from the prediction of mode of action for the 18 novel multimodal compounds show a correlation between preference to agonist/antagonist-selective models and polar contacts to residues of TM5 (table 6). There were 16 predicted agonist activities for compounds in target receptors, and 14 of these showed a conformation in which they could form polar interactions

to polar side chains of residues at positions 5.42, 5.43, and 5.46. The correlation for predicted antagonists was more moderate; 11 of the 16 predicted antagonist activities for the compounds showed conformations where they were unable to have polar interactions to TM5 residues. The total numbers show that 25 of the 32 predicted modes of action are supported by the conformation of the compounds in the binding site.

Three of the compounds were not docked in their target receptor, meaning no acceptable conformation could be found for the compounds. As the compounds were experimentally determined to have affinity for their target receptors, their lack of acceptable conformations exposes one of the limitations of molecular modeling; docking into rigid receptor structures will not capture a binding site that fits ligands of all sizes. Receptors are flexible in nature, and will alter their binding site to accommodate ligands of different size. Although generating multiple conformations of the receptor with MODELLER allowed selection of models suited for most ligands, docking was still performed into rigid structures. Protein flexibility is one of the challenges of molecular modeling, the complexity of macromolecules is difficult to handle computationally, and methods to account for this include induced fit docking and molecular dynamics.

Another problem of molecular modeling is the task of correlating docking scores with experimental affinity. Docking modules of modeling software are often better at determining correct ligand conformation in the binding site, and separating active compounds from decoys, than to calculate affinity. The determination of which selective model (either agonist or antagonist) each compound had a preference to was made on the basis of their GlideScore in each model. In some of the cases the difference in GlideScore was minimal, thus making visually inspecting the docking pose essential in the binding mode analysis.

Docking poses of predicted agonists for each target receptor (figure 16, 18, 20, 22) show that all have polar moieties in contact with polar residues of TM5, and strong anchoring to D3.32. T6125232 in 5-HT_{1A}R has a hydrogen bond from an amine to S5.42, and favorable interactions with aromatic residues (Y2.64, F.328, Y5.38, W6.48, F6.51, and F6.52), and contact to ECL2 residues (T188 and I189). The predicted agonist 9034414 for 5-HT_{2A}R has chlorine in contact with S5.43, but lacks interaction with residues of ECL2, probably due to the large size of the binding pocket. For 5-HT_6R, the predicted agonist 7989485 has a hydroxyl moiety in contact with S5.43, and has aromatic interactions to important residues (F5.38, W6.48, F6.51, and F6.52). Predicted agonist 5456380 for 5-HT_7R has a hydroxyl moiety in contact with S5.42, and multiple favorable interactions with aromatic residues (F5.38, W6.48, F6.51, and F6.52).

For the predicted antagonists in their target receptors, selected docking poses show that none of them have polar moieties to the polar residues of TM5 (figure 17, 19, 21, 23). Instead, they have a hydrophobic phenyl group facing towards the polar side chains of residues at positions 5.42, 5.43, and 5.46 at TM5.

The implication of predicted agonist activity for compounds 9134052, T6125232 and T6275452 in 5-HT_{1A}R is that they may be developed further as SPARIs, combining SERT inhibition with partial agonism at 5-HT_{1A}R. Partial agonism of 5-HT_{1A}R is associated with an accelerated onset of therapeutic action.[33] This study only predicts agonist or antagonist activity for the compounds; partial agonism needs to be examined further experimentally. T6125232 was predicted to be an antagonist for 5-HT_{2A}R additionally, and antagonism at this receptor has been found to reduce adverse effects often experienced with present antidepressants.[31]

Combining SERT inhibition with predicted agonism of 5-HT_{1A}R and antagonism at the novel target 5-HT_7R, compound T6275452 is interesting for further studies. Antagonism at the 5-HT_7R has been theorized to increase efficacy of antidepressants.[40]

Compound T0503-1300 had predicted antagonist activity for 5-HT_{2A}R and 5-HT_6R. Antidepressants combining SERT inhibition with antagonist activity at 5-HTRs are labeled as SARIs, and are associated with fewer and less debilitating side effects.[31] 5-HT_6R antagonism is theorized to both accelerate onset of therapeutic action and produce fewer side effects[41], compounds with this activity are thus of interest for future studies.

Future directions for this study include experimental testing of the compounds, in order to determine the actual mode of action for the compounds in the target receptors. Agonist or antagonist activity is only predicted on the basis of computational methods in this study, and the verification of results from molecular modeling is always made by experimental methods.

Further modeling procedures can include induced fit docking of each compound into the target receptors. Induced fit docking optimizes the binding site residues of the receptors to fit the ligand, by allowing the residues to alter conformation in regard to the ligand. This can provide further information of binding mode for the compounds in the target receptors, and increase the likelihood of getting correct ligand conformation for every compound.

Molecular dynamics is a computer simulation method, developed to study the movements of atoms and molecules over time, which may provide insights into the binding mode of the compounds. The application of molecular dynamics to this study can be to analyze

conformational changes of ligand and receptor over time, when the receptor is bound to either a predicted agonist or antagonist. Observing the movement of TM5 of the receptors during the simulation can support the results of this study.

6. Conclusion

The constructed homology models of 5-HT_{1A}R, 5-HT_{2A}R, 5-HT_6R and 5-HT_7R performed well in both the structural quality assessment and the docking of known agonists and antagonists, indicating that the models were predictive. Selective models able to differentiate between agonists and antagonists were determined from the docking of known binders, while calculation of SIFts for known agonists and antagonists in the 5-HTRs identified important binding site residues. The docking indicated that agonists formed polar interactions with amino acid residues in TM5, which is in agreement with previous observations from X-ray crystal structure complexes of GPCRs.

The docking of multimodal compounds with affinity for SERT and different 5-HTRs determined favorable binding modes, and highlighted interactions with residues important for affinity. The mode of action of the multimodal compounds for the different 5-HTR was predicted based on their preference to agonist- or antagonist-selective models, and by the binding mode in the targets.

7. References

1. Belmaker, R. H. & Agam, G. Major Depressive Disorder. *N. Engl. J. Med.* **358,** 55–68 (2008).
2. Kringlen, E., Torgersen, S. & Cramer, V. A Norwegian Psychiatric Epidemiological Study. *Am. J. Psychiatry* **158,** 1091–1098 (2001).
3. Leonard, B. E. Neuropsychopharmacology—The fifth generation of progress. Edited by K. L. Davis, D. Charney, J. T. Coyle, C. Nemeroff. Lippincott, Williams and Wilkins: Philadelphia, 2002. ISBN: 0-7817-2837-1. Price: $189. Pages: 2080. *Hum. Psychopharmacol. Clin. Exp.* **17,** 433–433 (2002).
4. Kessler, R. C. & Bromet, E. J. The epidemiology of depression across cultures. *Annu. Rev. Public Health* **34,** 119–138 (2013).
5. Rihmer, Z. Can better recognition and treatment of depression reduce suicide rates? A brief review. *Eur. Psychiatry* **16,** 406–409 (2001).
6. Kalia, M. Neurobiological basis of depression: an update. *Metabolism.* **54,** 24–27 (2005).
7. Hamet, P. & Tremblay, J. Genetics and genomics of depression. *Metabolism.* **54,** 10–15 (2005).
8. Varghese, F. P. & Brown, E. S. The Hypothalamic-Pituitary-Adrenal Axis in Major Depressive Disorder: A Brief Primer for Primary Care Physicians. *Prim. Care Companion J. Clin. Psychiatry* **3,** 151–155 (2001).
9. Lee, B.-H. & Kim, Y.-K. The Roles of BDNF in the Pathophysiology of Major Depression and in Antidepressant Treatment. *Psychiatry Investig.* **7,** 231–235 (2010).
10. Hirschfeld, R. M. A. History and Evolution of the Monoamine Hypothesis of Depression. *J. Clin. Psychiatry* **61,** 4–6 (2000).
11. Fiedorowicz, J. G. & Swartz, K. L. The Role of Monoamine Oxidase Inhibitors in Current Psychiatric Practice. *J. Psychiatr. Pract.* **10,** 239–248 (2004).

12. Delgado, P. L. Depression: the case for a monoamine deficiency. *J. Clin. Psychiatry* **61 Suppl 6,** 7–11 (2000).

13. Duman, R. S. & Li, N. A neurotrophic hypothesis of depression: role of synaptogenesis in the actions of NMDA receptor antagonists. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **367,** 2475–2484 (2012).

14. Warner-Schmidt, J. L. & Duman, R. S. Hippocampal neurogenesis: opposing effects of stress and antidepressant treatment. *Hippocampus* **16,** 239–249 (2006).

15. Torres, G. E., Gainetdinov, R. R. & Caron, M. G. Plasma membrane monoamine transporters: structure, regulation and function. *Nat. Rev. Neurosci.* **4,** 13–25 (2003).

16. Cherezov, V. *et al.* High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318,** 1258–1265 (2007).

17. Drews, J. Drug discovery: a historical perspective. *Science* **287,** 1960–1964 (2000).

18. Kristiansen, K. 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 (2004).

19. Landry, Y. & Gies, J. P. Heterotrimeric G proteins control diverse pathways of transmembrane signaling, a base for drug discovery. *Mini Rev. Med. Chem.* **2,** 361–372 (2002).

20. Juan A. Ballesteros, H. W. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* **25,** 366–428 (1995).

21. Barnes, N. M. & Sharp, T. A review of central 5-HT receptors and their function. *Neuropharmacology* **38,** 1083–1152 (1999).

22. Nichols, D. E. & Nichols, C. D. Serotonin receptors. *Chem. Rev.* **108,** 1614–1641 (2008).

23. Neves, S. R., Ram, P. T. & Iyengar, R. G Protein Pathways. *Science* **296,** 1636–1639 (2002).

24. Wettschureck, N. & Offermanns, S. Mammalian G Proteins and Their Cell Type Specific Functions. *Physiol. Rev.* **85,** 1159–1204 (2005).

25. Warne, T. *et al.* The structural basis for agonist and partial agonist action on a β1-adrenergic receptor. *Nature* **469,** 241–244 (2011).

26. Venkatakrishnan, A. J. *et al.* Molecular signatures of G-protein-coupled receptors. *Nature* **494,** 185–194 (2013).

27. Gandhimathi, A. & Sowdhamini, R. Molecular modelling of human 5-hydroxytryptamine receptor (5-HT2A) and virtual screening studies towards the identification of agonist and antagonist molecules. *J. Biomol. Struct. Dyn.* **34,** 952–970 (2016).

28. Squire, Larry & et. al. *Fundamental Neuroscience (3rd ed.)*. (Elsevier / Academic Press).

29. Coleman, J. A., Green, E. M. & Gouaux, E. X-ray structures and mechanism of the human serotonin transporter. *Nature* **advance online publication,** (2016).

30. Gabrielsen, Mari. Structure, function and inhibition of the serotonin transporter studied by molecular docking, -dynamics and virtual screening. (2011).

31. Stahl, S. M. *Stahl's Psychopharmacology*. (Cambridge University Press, 2013).

32. Nutt, D. J. Beyond psychoanaleptics - can we improve antidepressant drug nomenclature? *J. Psychopharmacol. Oxf. Engl.* **23,** 343–345 (2009).

33. Stahl, S. M. Mechanism of action of the SPARI vilazodone: serotonin 1A partial agonist and reuptake inhibitor. *CNS Spectr.* **null,** 105–109 (2014).

34. Harmer, C. J., Goodwin, G. M. & Cowen, P. J. Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action. *Br. J. Psychiatry* **195,** 102–108 (2009).

35. Bradley N Gaynes, A. J. R. The STAR*D study: treating depression in the real world. Cleve Clin J Med. *Cleve. Clin. J. Med.* **75,** 57–66 (2008).

36. Pigott, H. E., Leventhal, A. M., Alter, G. S. & Boren, J. J. Efficacy and effectiveness of antidepressants: current status of research. *Psychother. Psychosom.* **79,** 267–279 (2010).

37. Bang-Andersen, B. *et al.* Discovery of 1-[2-(2,4-dimethylphenylsulfanyl)phenyl]piperazine (Lu AA21004): a novel multimodal compound for the treatment of major depressive disorder. *J. Med. Chem.* **54,** 3206–3221 (2011).

38. Schreiber, R. & De Vry, J. 5-HT1A receptor ligands in animal models of anxiety, impulsivity and depression: multiple mechanisms of action? *Prog. Neuropsychopharmacol. Biol. Psychiatry* **17,** 87–104 (1993).

39. De Boer, S. F. & Koolhaas, J. M. 5-HT1A and 5-HT1B receptor agonists and aggression: a pharmacological challenge of the serotonin deficiency hypothesis. *Eur. J. Pharmacol.* **526,** 125–139 (2005).

40. Katona, C. L. & Katona, C. P. New generation multi-modal antidepressants: focus on vortioxetine for major depressive disorder. *Neuropsychiatr. Dis. Treat.* **10,** 349–354 (2014).

41. Artigas, F. Serotonin receptors involved in antidepressant effects. *Pharmacol. Ther.* **137,** 119–131 (2013).

42. Leach, A. R. *Molecular Modelling, Principles and Applications. 2nd Edition*. (Pearson Education Limited).

43. Höltje, H.-D. & Folkers, G. *Molecular Modeling: Basic Principles and Applications*. (John Wiley & Sons, 2008).

44. Eswar, N. *et al.* in *Current Protocols in Protein Science* (John Wiley & Sons, Inc., 2001).

45. Pearson, W. R. Finding Protein and Nucleotide Similarities with FASTA. *Curr. Protoc. Bioinforma. Ed. Board Andreas Baxevanis Al* **53,** 3.9.1–3.9.25 (2016).

46. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402 (1997).

47. Ravna, A. W. & Sylte, I. Homology modeling of transporter proteins (carriers and ion channels). *Methods Mol. Biol. Clifton NJ* **857,** 281–299 (2012).

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

49. Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A. & Farid, R. Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. *J. Med. Chem.* **49,** 534–553 (2006).

50. Laskowski, R. A. & Swindells, M. B. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51,** 2778–2786 (2011).

51. Deng, Z., Chuaqui, C. & Singh, J. Structural interaction fingerprint (SIFt): a novel method for analyzing three-dimensional protein-ligand binding interactions. *J. Med. Chem.* **47,** 337–344 (2004).

52. Huang, S.-Y., Grinter, S. Z. & Zou, X. Scoring functions and their evaluation methods for protein–ligand docking: recent advances and future directions. *Phys. Chem. Chem. Phys.* **12,** 12899–12908 (2010).

53. Gabrielsen, M. *et al.* Identification of novel serotonin transporter compounds by virtual screening. *J. Chem. Inf. Model.* **54,** 933–943 (2014).

54. Friesner, R. A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **47,** 1739–1749 (2004).

55. Schrödinger Release 2015-4: LigPrep, version 3.6, Schrödinger, LLC, New York, NY, 2015.

56. Schrödinger Release 2015-4: Maestro, version 10.4, Schrödinger, LLC, New York, NY, 2015.

57. Jacobson, M. P., Friesner, R. A., Xiang, Z. & Honig, B. On the Role of the Crystal Environment in Determining Protein Side-chain Conformations. *J. Mol. Biol.* **320,** 597–608 (2002).

58. Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* **27,** 221–234 (2013).

59. Abagyan, R., Totrov, M. & Kuznetsov, D. ICM—A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* **15,** 488–506 (1994).

60. Consortium, T. U. UniProt: a hub for protein information. *Nucleic Acids Res.* **43,** D204–D212 (2015).

61. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28,** 235–242 (2000).

62. Southan, C. *et al.* The IUPHAR/BPS Guide to PHARMACOLOGY in 2016: towards curated quantitative interactions between 1300 protein targets and 6000 ligands. *Nucleic Acids Res.* **44,** D1054–1068 (2016).

63. Kufareva, I., Katritch, V., Participants of GPCR Dock 2013, Stevens, R. C. & Abagyan, R. Advances in GPCR modeling evaluated by the GPCR Dock 2013 assessment: meeting new challenges. *Struct. Lond. Engl. 1993* **22,** 1120–1139 (2014).

64. Šali, A. & Blundell, T. L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **234,** 779–815 (1993).

65. Sali, A. & Overington, J. P. Derivation of rules for comparative protein modeling from a database of protein structure alignments. *Protein Sci. Publ. Protein Soc.* **3,** 1582–1596 (1994).

66. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L. & Evanseck, J. D. All-atom empirical potential for molecular modeling and dynamics studies of proteins. (1998).

67. Braun, W. & Go, N. Calculation of protein conformations by proton-proton distance constraints. A new efficient algorithm. *J. Mol. Biol.* **186,** 611–626 (1985).

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

69. McGuffin, L. J., Buenavista, M. T. & Roche, D. B. The ModFOLD4 server for the quality assessment of 3D protein models. *Nucleic Acids Res.* **41,** W368–372 (2013).

70. Buenavista, M. T., Roche, D. B. & McGuffin, L. J. Improvement of 3D protein models using multiple templates guided by single-template model quality assessment. *Bioinforma. Oxf. Engl.* **28,** 1851–1857 (2012).

71. McGuffin, L. J. & Roche, D. B. Rapid model quality assessment for protein structure predictions using the comparison of multiple models without structural alignments. *Bioinforma. Oxf. Engl.* **26,** 182–188 (2010).

72. Harder, E. *et al.* OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* **12,** 281–296 (2016).

73. Eldridge, M. D., Murray, C. W., Auton, T. R., Paolini, G. V. & Mee, R. P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.* **11,** 425–445 (1997).

74. Truchon, J.-F. & Bayly, C. I. Evaluating Virtual Screening Methods: Good and Bad Metrics for the 'Early Recognition' Problem. *J. Chem. Inf. Model.* **47,** 488–508 (2007).

75. Mordalski, S., Kosciolek, T., Kristiansen, K., Sylte, I. & Bojarski, A. J. Protein binding site analysis by means of structural interaction fingerprint patterns. *Bioorg. Med. Chem. Lett.* **21,** 6816–6819 (2011).

Supplementary information

I. Multiple sequence alignment of all 5-HTRs and considered templates

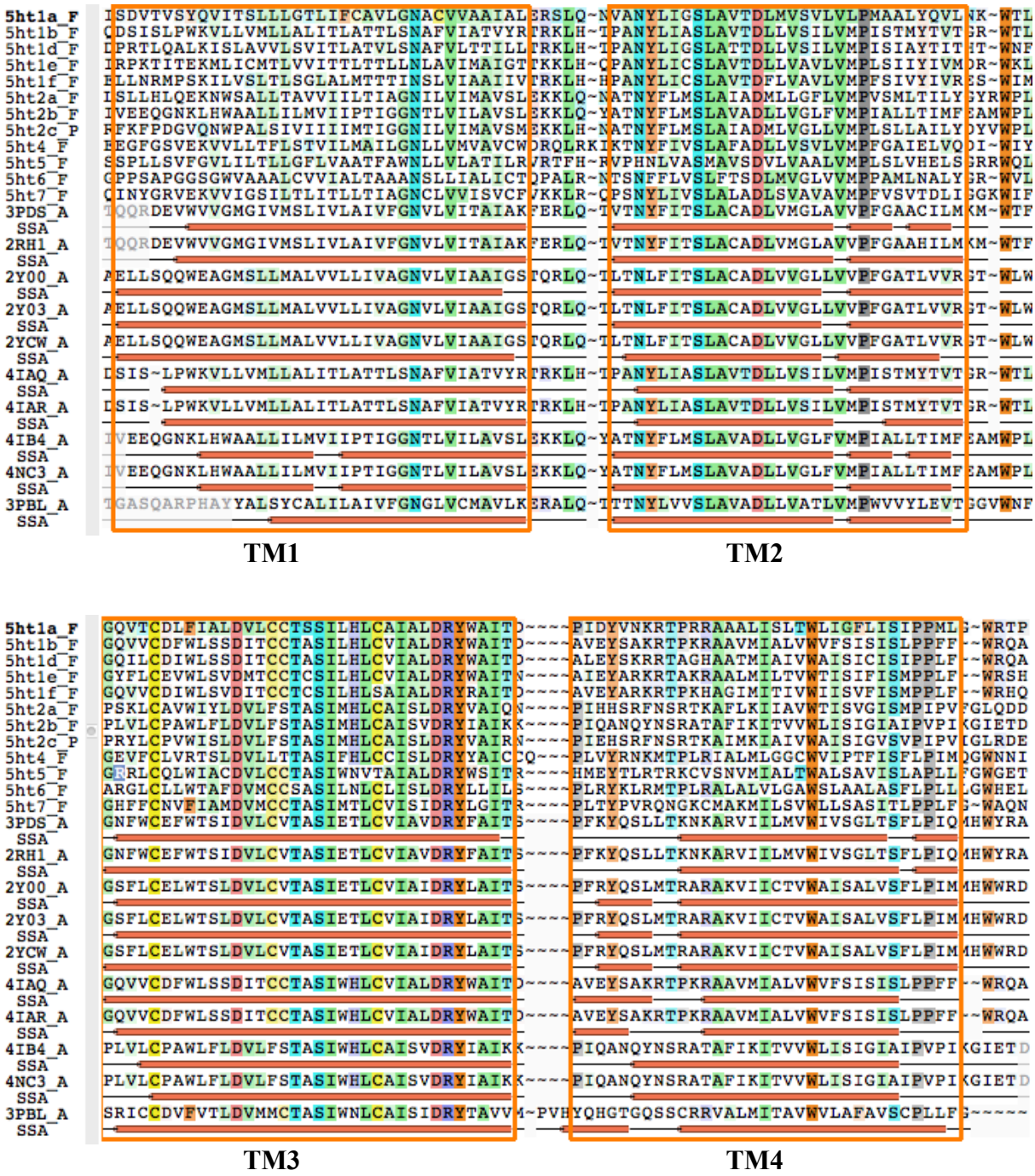
II. 2D structures of the 18 multimodal compounds

III. Experimental testing of compounds

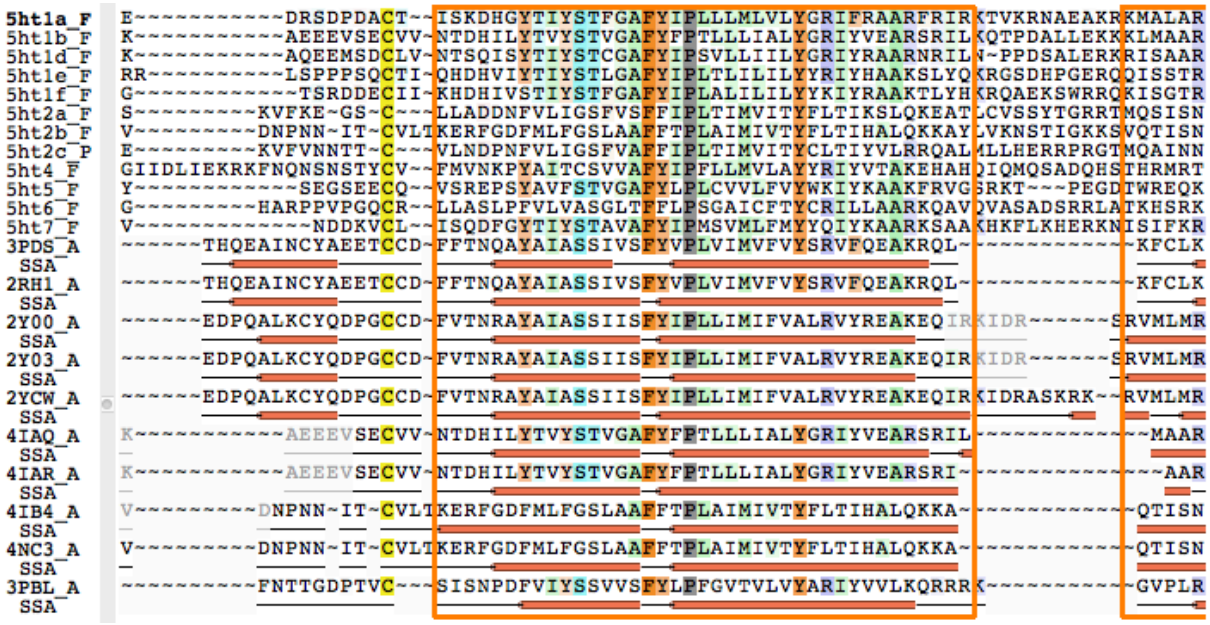
IV. Experimental affinities and GlideScores of 18 multimodal compounds

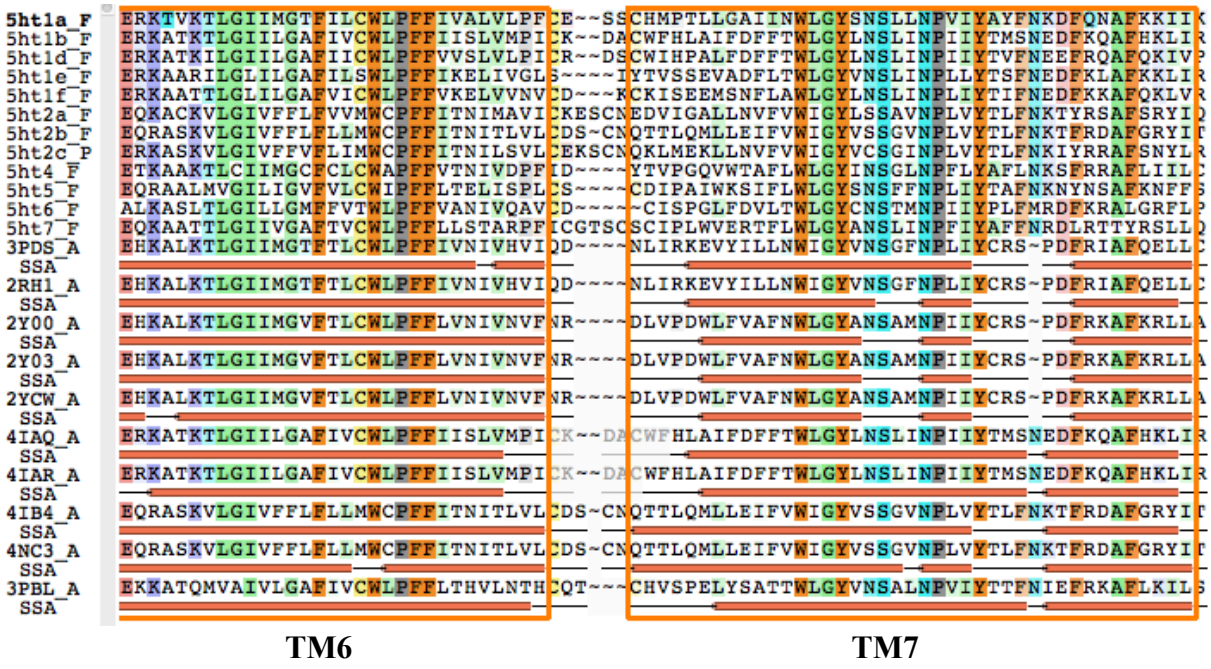
I. Multiple sequence alignment

Amino acid alignment of all 5-HTRs (except 5-HT$_3$) and all considered templates.



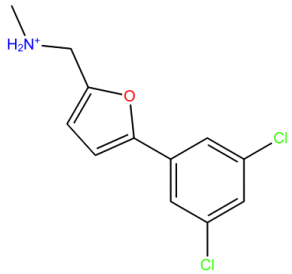
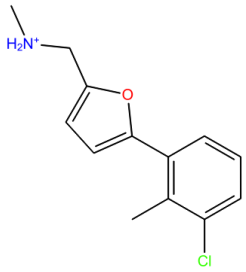
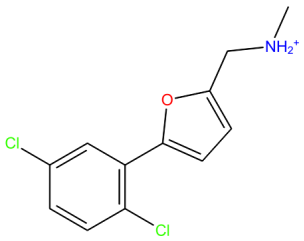
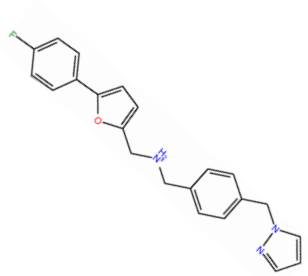
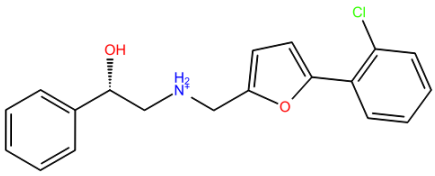
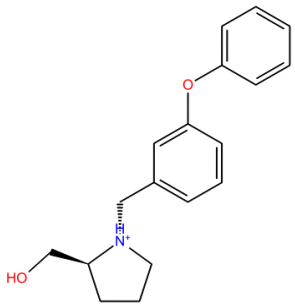
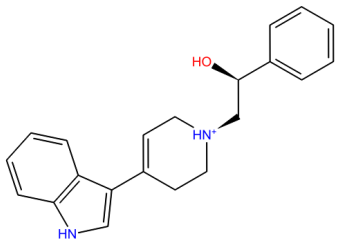
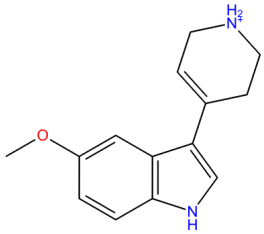
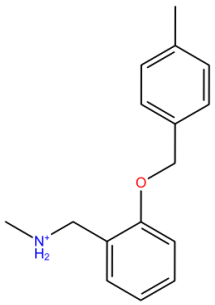
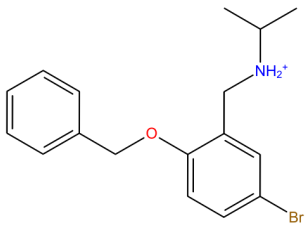
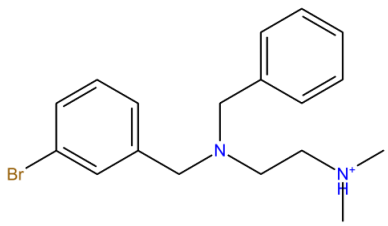
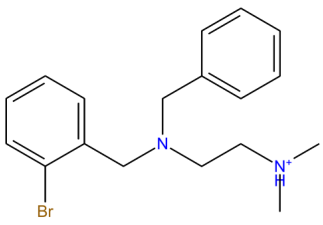
A multiple sequence alignment of all 5-HTRs (excluding the ion channel 5-HT$_3$R) and considered family GPCR templates. The seven transmembrane α-helices of the GPCRs are annotated as TM1 – TM7, represented by orange rectangles. Non-modeled regions C-terminal, N-terminal and ICL3 are excluded from the alignment. SSA, secondary structure; 3PDS/2RH1, β_2 adrenergic receptor crystal structures; 2Y00/2Y03/2YCW, turkey β_1 adrenergic receptor crystal structures; 4IAQ/4IAR, 5-HT$_{1B}$ receptor crystal structures; 4IB4/4NC3, 5-HT$_{2B}$ receptor crystal structures; 3PBL, dopamine D$_3$ receptor crystal structure.

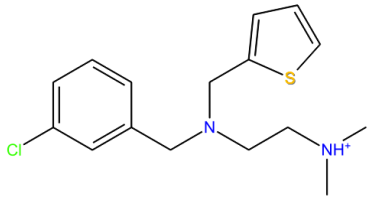
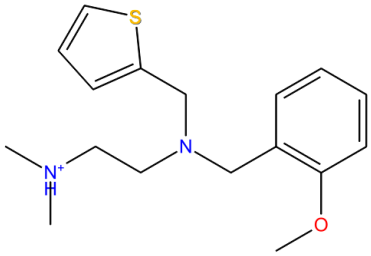
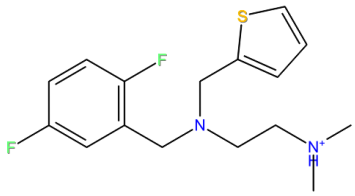
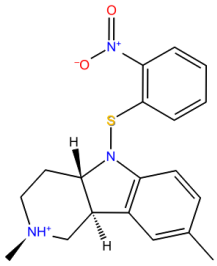
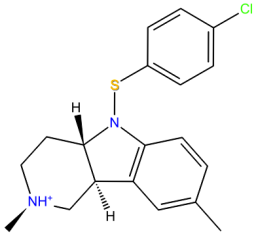
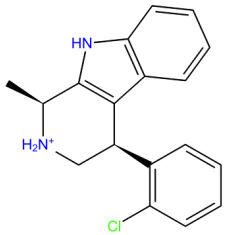




(continued)

II. 2D structures of 18 multimodal compounds

ASN 13153175  **C01**	**9034414**  **C01**	**9013195**  **C01**
T6209417  **C01**	**9134052**  **C01**	**5417988**  **C03**
T6125232  **C04**	**T6275452**  **C04**	**9066608**  **C06**
7989485  **C06**	**5458751**  **C07**	**5456380**  **C07**

SYN 16295816  **C07**	**ASN 16295801**  **C07**	**SYN 16295876**  **C07**
T0502-9459  **C12**	**T0503-1300**  **C12**	**EN300-08612**  **C13**

III. Experimental testing of compounds

Our collaborating partners in Krakow performed testing of the compounds experimentally. Prof. Andrzej Bojarski supplied the description of the methods used.

In vitro pharmacology

Cell culture

HEK293 cells with stable expression of human serotonin 5-HT_{1A}R, 5-HT_6 and 5-HT_{7b}R (obtained with the use of Lipofectamine 2000, Invitrogen) or CHO-K1 cells with plasmid containing the sequence coding for the human serotonin 5-HT_{2A} receptor (Perkin Elmer) were maintained at 37 °C in a humidified atmosphere with 5% CO_2 and were grown in Dulbecco's Modified Eagle Medium containing 10% dialyzed fetal bovine serum and 500 μg/ml G418 sulfate. For membranes preparations, cells were subcultured into 150 cm^2 cell culture flasks, grown to 90% confluence, washed twice with prewarmed to 37 °C phosphate buffered saline (PBS) and were pelleted by centrifugation (200 g) in PBS containing 0.1 mM EDTA and 1 mM dithiothreitol, and stored at –80 °C.

5-HT_{1A}/5-HT_{2A}/5-HT_6/5-HT_7 radioligand binding assays

Cell pellets were thawed and homogenized in 10 volumes of assay buffer using an Ultra Turrax tissue homogenizer and centrifuged twice at 35,000 g for 15 min at 4 °C, with incubation for 15 min at 37 °C in between the rounds of centrifugation. The composition of the assay buffers was as follows: for 5-HT_{1A}R: 50 mM Tris–HCl, 0.1 mM EDTA, 4 mM $MgCl_2$, 10 μM pargyline and 0.1% ascorbate; for 5-HT_{2A}R: 50 mM Tris–HCl, 0.1 mM EDTA, 4 mM $MgCl_2$ and 0.1% ascorbate; for 5-HT_6R: 50 mM Tris–HCl, 0.5 mM EDTA and 4 mM $MgCl_2$, for 5-HT_{7b}R: 50 mM Tris–HCl, 4 mM $MgCl_2$, 10 μM pargyline and 0.1% ascorbate.

All assays were incubated in a total volume of 200 μL in 96-well microtiter plates for 1 h at 37 °C, except for 5-HT_{1A}R and 5-HT_{2A}R, which were incubated at room temperature and 27 °C, respectively. The equilibration process was terminated by rapid filtration through Unifilter plates with a 96-well cell harvester and the radioactivity retained on the filters was quantified using a Microbeta plate reader (PerkinElmer, USA).

For the displacement studies, the assay samples contained the following as radioligands (PerkinElmer, USA): 1.5 nM [^{3}H]-8-OH-DPAT (135.2 Ci/mmol) for 5-HT_{1A}R; 2 nM [^{3}H]-ketanserin (53.4 Ci/mmol) for 5-HT_{2A}R; 2 nM [^{3}H]-LSD (83.6 Ci/mmol) for 5-HT_6R or 0.6

nM [^{3}H]-5-CT (39.2 Ci/mmol) for 5-HT_7R. Non-specific binding was defined using 10 μM of 5-HT in 5-HT_{1A}R and 5-HT_7R binding experiments, whereas 20 μM of mianserin or 10 μM of methiothepine was used in the 5-HT_{2A}R, and 5-HT_6R assays, respectively. Each compound was tested in triplicate at 7–8 concentrations (10^{-11}–10^{-4} M). The inhibition constants (K_i) were calculated from the Cheng-Prusoff equation [1]. Results were expressed as the means of at least two independent experiments.

[1] Y. Cheng, W. Prusoff, Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction, Biochem. Pharmacol. 22 (1973) 3099–3108.

IV. Experimental affinities and GlideScores for 18 multimodal compounds

Experimental affinities (K_i) were obtained from the lab testing of compounds, and GlideScores were from the docking of multimodal compounds in target receptors.

Cluster	Ligand	SERT	5-HT1A		5-HT2A		5-HT6		5-HT7	
		Ki	Ki	Score	Ki	Score	Ki	Score	Ki	Score
C01	ASN13153175	91	2368	-6.397	17750	-7.110	292	-7.173	17670	-6.948
C01	9034414	134	13820	-6.907	846	-7.298	568	-7.816	1049	-7.189
C01	9013195	164	1013	-6.501	839	-7.495	117	-7.387	730	-6.902
C01	T6209417	427	2891	-8.681	655	-9.762	2650	-7.973	1181	-8.116
C01	9134052	825	187	-8.239	1929	-9.186	1861	-7.760	11610	-8.405
C03	5417988	926	860	-7.667	5108	-7.491	6408	-6.896	2726	-6.046
C04	T6125232	2	56	-8.130	217	-9.145	569	-7.377	314	-7.773
C04	T6275452	288	8	-7.114	1216	-8.432	101	-8.075	126	-8.102
C06	9066608	28	2660	-8.169	4033	-7.719	742	-7.389	819	-8.077
C06	7989485	657	1203	-8.220	818	-7.844	420	-6.578	3527	-6.120
C07	5458751	50	5946	-7.714	1113	-8.477	596	-7.378	1082	-6.904
C07	5456380	84	2470	-7.281	265	-8.670	856	-7.376	830	-7.339
C07	SYN16295816	107	20590	-7.619	508	-8.906	640	-7.288	895	-8.174
C07	ASN16295801	364	12340	-7.298	543	-8.333	11680	-6.961	1556	-7.749
C07	SYN16295876	441	19040	-7.816	737	-9.377	1046	-8.227	2213	-7.694
C12	T0502-9459	268	6292	-7.527	309	-8.190	69	-8.106	1460	-7.260
C12	T0503-1300	790	1153	-7.431	315	-8.352	46	-7.980	92	-5.906
C13	EN300-08612	322	6320	-7.303	1460	-8.246	1433	-7.030	581	-6.944