UiT
THE ARCTIC UNIVERSITY OF NORWAY

Faculty of Health Sciences
Department of Medical Biology
Medical Pharmacology and Toxicology Research Group

Modelling of serotonergic receptors and molecular optimization of X-ray crystal structures of serotonin transporter and their interactions with exogenous compounds

—

Dušan Milićević

Master thesis in Pharmacy [FAR-3911] May 2016

UNIVERSITY OF TROMSØ

Acknowledgements

This master thesis was written at the Department of Medical Biology at the University of Tromsø in collaboration with the Department of Pharmacy (IFA) from August 2015 to May 2016.

I would like to express my deepest gratitude to the main supervisor Associate Professor Kurt Kristiansen and co-supervisor Professor Ingebrigt Sylte for their dedicated assistance, guidance and support during the course of this research. With the help of these two fantastic supervisors, I have deepened my knowledge and skills in the molecular modelling field.

I am also immensely thankful to Isak Andreas Bøgwald for teaching me how to use different software programs and handle the computational commands and scripts.

I would also like to extend my deepest gratitude to my parents Goran and Vesna, and my brothers Matija and Danilo for encouraging me to study pharmacy and for supporting me through all these years.

Tromsø, May 2016

Dušan Milićević

Abstract

The serotonin (5-hydroxytryptamine, 5-HT) receptors and transporter are in the serotonergic neurotransmission system, and believed to have a major role in pathology of depression. They are of pharmacological importance, being targeted by many nowadays antidepressants. It is therefore of great interest to understand their structural and functional properties for development of future drugs.

There is generally little knowledge today about the effects of environmental toxicants on the human brain. If the exogenous compounds interact with the serotonin receptors and transporter, they may interfere with the serotonergic neurotransmission in the brain and interfere with the effects of the CNS drugs.

Homology modelling is an *in silico* method used for prediction of the 3D structure of structurally unknown proteins. Models of serotonergic receptors (5-HT_{1A}, 5-HT_{2A}, 5-HT_{2C}) were constructed by the homology approach with known structures in the PDB. The newly released X-ray crystal structures of the human serotonin transporter (SERT) were also imported from the PDB and optimized with molecular modelling techniques. Molecular docking was utilized to predict putative harmful effects and drug interactions of the toxicants in the Tox21 database with these protein targets. Many toxic compounds were predicted to interact with serotonergic receptors and the SERT and many of these had physiochemical properties that suggest that they may act in the CNS. Detailed interaction analysis of the selected compounds of serotonergic receptors and the SERT indicated that besides the crucial interaction with an aspartic acid, aromatic interactions with phenylalanine residues are also very important. The obtained high CNS MPO scores and similar Glide scores between the known high affinity binders and toxicants could suggest harmful effects and drug interactions in serotonergic system of the CNS.

Table of Contents

Acknowledgements **3**

Abstract **5**

List of abbreviations **8**

1. Introduction **11**

1.1 Environmental toxicants and their effects on the central nervous system (CNS) **11**

1.2 Blood-brain barrier **12**

1.3 G-protein coupled receptors **12**

1.3.1 Family A of G-protein coupled receptors 14

1.4 Serotonin as a neurotransmitter and its receptors **15**

1.4.1 Serotonin receptors 5-HT_{1A} and 5-$HT_{2A/2C}$ 16

1.4.2 Serotonin transporter (SERT) 17

1.5 Future aspects in pharmacological treatment of depression and anxiety **18**

1.5.1 Psychotherapy 20

1.6 Computational science and drug discovery **21**

1.6.1 Molecular modelling 21

1.6.2 Homology modelling 22

1.6.3 Docking and scoring 24

2. Aim of the study **26**

3. Methods **27**

3.1 Software and databases **27**

3.1.1 Molsoft Internal Coordinates Mechanics (Version 3.8.4) 27

3.1.2 Schrödinger (Version 2015.3) 27

3.1.3 The Protein Data Bank (PDB) 27

3.1.4 UniProt Knowledge-Base 28

3.1.5 ChEMBL 28

3.1.6 IUPHAR/BPS Guide to PHARMACOLOGY 28

3.1.7 Database of Useful Decoys-Enhanced (DUD-E) 29

3.1.8 Tox21 database (Version: Spring 2012) 29

3.2 Homology modelling **30**

3.2.1 Template identification 30

3.2.2 Sequence alignment 31

3.3 Molecular docking **33**

3.3.1 Generation of binders and decoys 33
3.3.2 Preparation of ligands, receptors and transporter for docking 33
3.3.3 Induced fit docking (IFD) 34
3.3.4 Virtual screening workflow (VSW) 35
3.3.4 Evaluation of models with BEDROC 36
3.3.5 Glide docking calculations with exogenous toxicants 37
3.4 Prediction of the BBB passage by CNS MPO tool 38
4. Results 39
4.1 Homology models 39
4.2 Molecular docking 41
4.2.1 Results from BEDROC calculations 41
4.2.2 Induced fit docking 43
4.2.3 Docking of known binders into receptor models obtained by IFD 47
4.2.4 Docking of known binders into crystal structures of SERT 49
4.2.5 Screening scores for exogenous toxicants 52
5. Discussion 59
6. Conclusion 64
References 65
Appendix 69

List of abbreviations

2D	two-dimensional
3D	three-dimensional
5-HIAA	5-hydroxyindolacetic acid
5-HT	5-hydroxytryptamine
5-HTP	5-hydroxytryptophan
7TM	seven-transmembrane
AADC	aromatic amino acid decarboxylase
ADMET	absorption, distribution, metabolism, excretion and toxicity
β_2	beta-2 adrenergic receptor
BBB	blood-brain barrier
BEDROC	Boltzmann-Enhanced Discrimination of Receiver-operating characteristics
BPS	British Pharmacological Society
CNS	central nervous system
CNS MPO	central nervous system multiparameter optimization
D_3	dopamine 3 receptor
DHE	dihydroergotamine
DUD	directory of useful decoys
DUD-E	database of useful decoys-enhanced
E_{angle}	angle energy
E_{bond}	bonding energy
ECL	extracellular loop
E_{tors}	torsional energy
E_{elec}	electrostatic energy
$E_{non\text{-}bond}$	non-bonding energy
E_{tot}	total potential energy
E_{vdw}	van der Waals energy
EPA	U.S. Environmental Protection Agency
ERG	ergotamine
GDP	guanosine diphosphate

GPCR	G-protein coupled receptor
GTP	guanosine-5'-triphosphate
GUI	graphical user interface
IFD	induced fit docking
ICL	intracellular loop
ICM	internal coordinate mechanics
IUPHAR	International Union of Basic and Clinical Pharmacology
LeuT	leucine transporter
LigPrep	ligand preparation
LSD	lysergic acid diethylamide
MAO	monoamine oxidase
MDD	major depressive disorder
MM	molecular mechanics
mRNA	messenger ribonucleic acid
NIH	National Institutes of Health
NMR	nuclear magnetic resonance
NSS	sodium symporter
PDB	protein data bank
PDB ID	protein data bank identification code
pIC_{50}	logarithmic half maximal inhibitory concentration
p*K*a	logarithmic acid dissociation constant
p*K*i	logarithmic binding affinity constant
PPCR	protein preparation constrained refinement
ProteinPrep	protein preparation
RIE	robust initial enhancement
ROC	receiver operating characteristic curve
SAVES	structural analysis and verification server
SB-CADD	structure-based computer-aided drug design
SERT	serotonin transporter
SMILES	simplified molecular-input line-entry system
SNRI	serotonin and norepinephrine reuptake inhibitors
SSRI	selective serotonin reuptake inhibitors
TCA	tricyclic antidepressants
TM	transmembrane

TMH	transmembrane helix
Tox21	Toxicology in the 21st Century
Trp	tryptophan
QM	quantum mechanics
UniProtKB	UniProt knowledgebase
VS	virtual screening
VSW	virtual screening workflow
Å	angstrom

1. Introduction

1.1 Environmental toxicants and their effects on the central nervous system (CNS)

Environmental toxicants include organic and inorganic substances that are harmful to human health and development. To which degree individuals are exposed to these chemicals is depended on factors as lifestyle, living and working place, foods and drinks, pharmaceutical consumption and radiation. The most commonly studied environmental toxicants are heavy metals, air pollutants and pesticides (1).

Environmental toxicants are seen as a major public health issue, though little is known about their effects on the human brain. Still, there is convincing evidence that chemicals in the environment can alter functions of the nervous system (2).

Dementia and Parkinson's syndrome have been associated with aluminium toxicity, while cerebellar ataxia with dementia have been associated with lithium overdose, and severe psychotic disorders with lysergic acid diethylamide (LSD). However, putative psychiatric and psychological problems associated with exposure of environmental chemicals are not much studied and even trivialised and more research is needed into both the acute and chronic effects of neurotoxic exposure on mental functions (3).

To get insight into if and how environmental toxicants may affect mechanisms in the human body, we need to understand their molecular interactions with human proteins including receptors involved in cellular signalling. Environmental toxicants may interact with human receptors and transporters and thereby affect the actions of neurotransmitters, hormones and inflammatorial mediators. They may also resemble the interactions of drugs with their targets and thereby interfere with pharmacological effects of drugs.

Sedation may be a form of neurotoxicity and it was suggested that one of the reasons for sedation is caused by interaction of toxicants with cell membranes in the CNS (either directly with membrane lipids or with membrane proteins), which impairs their electrical and chemical cell signalling (4). All toxicants that are able to cross the blood brain barrier (BBB) could possibly interact with receptors and transporters in the CNS and thereby give physiological effects.

In 2008, the National Institutes of Health (NIH) and the U.S. Environmental Protection Agency (EPA) announced a new toxicity testing initiative called Toxicology in the 21st Century (Tox21). This collaboration has contributed to the establishment of a database named

Tox21 library, which is now a collection of approximately 10 000 unique environmental chemicals and approved drugs. Tox21 researchers aim to develop better toxicity assessment methods to quickly and efficiently test whether certain chemical compounds have the potential to disrupt processes in the human body that may lead to negative health effects (5).

1.2 Blood-brain barrier

Not all the chemicals can reach and affect the CNS. They are prevented by the blood-brain barrier (BBB) and endothelial cell-astrocyte interactions.

The BBB consists of unusual tight junctions formed by endothelial cells surrounding capillaries that supply the brain tissue. This barrier prevents the free passage of the most blood-borne substances and thus maintains the control of what is entering the brain.

Some chemicals can, however, cross this barrier and enter the CNS. Compounds may either use special transport mechanisms or they could cross the membrane by passive diffusion if they are lipid soluble. The required nutrients, amino acids, fatty acids, hormones etc., are reaching the brain by these mechanisms. So, the environmental toxicants that structurally resemble these substances, or are lipid-soluble, may enter the CNS (2).

It is important to mention, that the BBB's permeability for chemicals is age-related and the BBB is not complete for until about 6 months after birth. The human brain continues to develop in the postnatal period, and is highly vulnerable over many months, through infancy and into early childhood (6,7). This means that some chemicals that are not harmful to adults, may be susceptible to injury children's brain, and other substances that are harmful in the mature brain at certain doses, may need even smaller doses to give neurotoxicity in children.

1.3 G-protein coupled receptors

At least 30% of all drugs on the market are targeting G-protein coupled receptors (GPCRs). GPCRs is the largest superfamily of membrane-bound receptors numbering around 800 members in the human genome (8). They are divided into several families (A-F), where A, B and C are the three main families.

All GPCRs share a common architecture of seven-transmembrane (7TM) α-helices. They are cell surface receptors (located in the lipid bilayer) that mediate biological responses to external stimuli by transducing signals across the plasma membrane to heterotrimeric G-proteins and arrestins, which in turn activate cellular signalling cascades. Their signal transduction is fundamental for most physiological processes and these receptors mediate the actions of neurotransmitters, hormones and paracrines. This makes them important therapeutic targets for drugs acting at receptors as agonists or antagonists (8,9).

Different experiments have shown that GPCRs can undergo conformational changes between active and inactive conformations (10). The binding of antagonists or inverse agonists stabilizes an inactive conformational state of the receptor, while the binding of agonists stabilizes an active conformational state of the receptor.

GPCRs interact with heterotrimeric G-proteins located in the intracellular part of a cell, made up of a Gα-subunit and a Gβγ heterodimer. When an agonist binds to a GPCR, the GPCR changes its conformation from an inactive to an active conformational state. Following this, the GPCR interacts with the appropriate G-proteins. This causes changes in the conformation of the Gα-subunit and the release of guanosine diphosphate (GDP).

Guanosine-5'-triphosphate (GTP) binds to the ternary complex consisting of the agonist, the GPCR and the G-protein. Subsequently, conformation of Gα-subunit changes and the heterotrimeric G-protein-complex dissociates into a Gα-GTP and a Gβγ-complex. The Gα-GTP and a Gβγ-complex stimulate and inhibit specific effector proteins (enzymes and ion channels) leading to cellular effects (Fig. 1) (11).

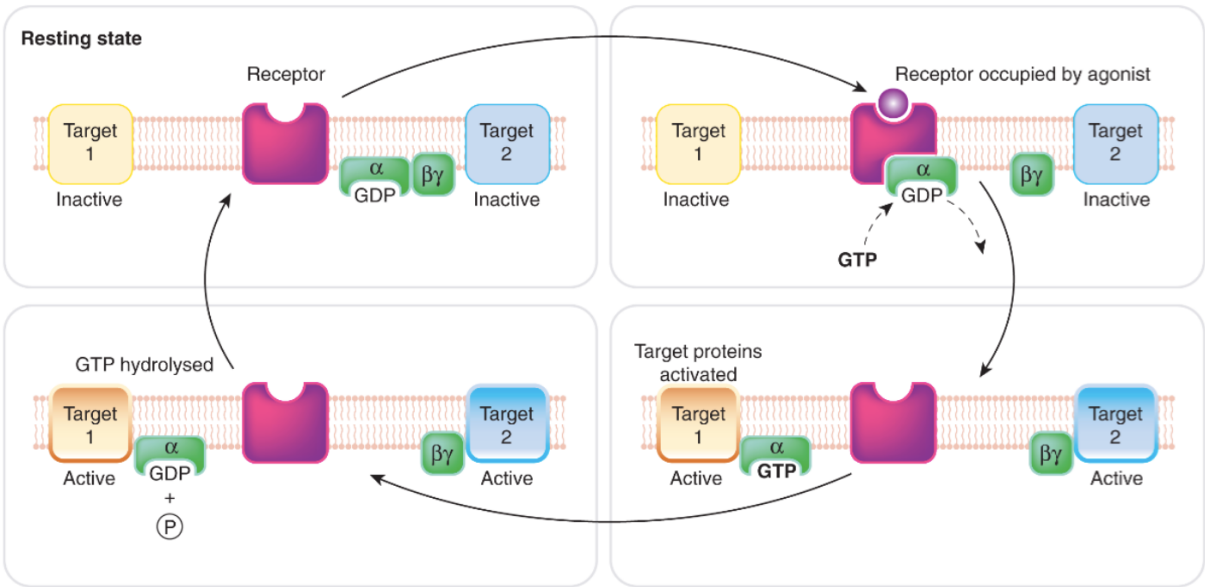


Fig. 1. The figure illustrates the function of the GPCRs (12).

1.3.1 Family A of G-protein coupled receptors

Family A of GPCRs makes up the vast majority of the GPCRs. They consist of a 7TM helical bundle in the membrane, three extracellular and three intracellular loops (ECLs and ICLs) connecting the individual helices and an extracellular N-terminus and an intracellular C-terminus. Typical for most of the family A GPCRs is a disulfide bridge between the ECL2 and the upper part of TM helix III. Most family A receptors also have a palmitoylated cysteine in the C-terminus (Fig. 2) (13).

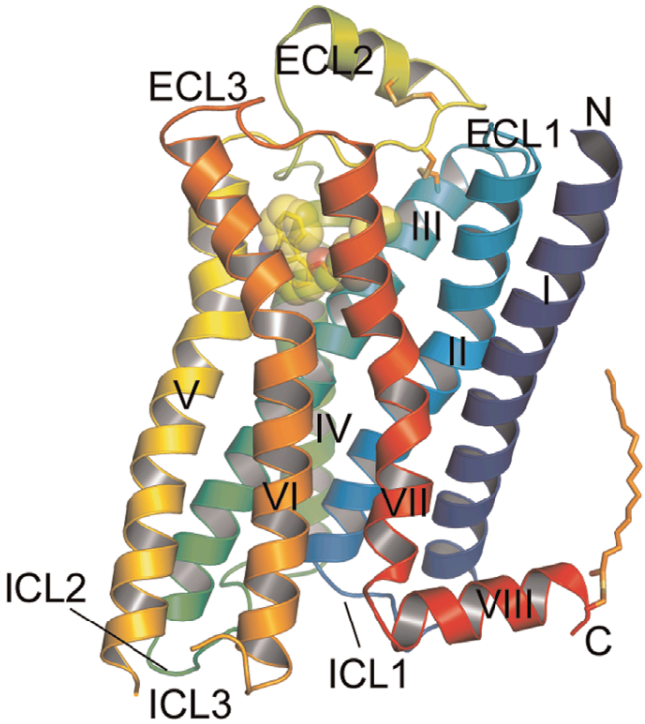


Fig. 2. Structure of the β_2-adrenergic receptor (family A of GPCRs) with a bound inverse agonist carazolol (PDB ID: 2RH1), Cherezov et al. (8); I-VII TM helices; VIII helix; 1-3 ICL – intracellular loops; 1-3 ECL – extracellular loops. The TM helices are marked with Roman numerals I-VII. The backbone atoms of the receptor are shown.

Different experimental approaches have shown that TM helix III plays a decisive role in ligand binding and receptor activation (8). Additionally, X-ray structures complexes have shown that ligands are making contact with residues in helices III, VI and VII, which define the orthosteric binding pocket, and in some cases with helix V, which is an important determinant for activation and ligand efficacy (8).

1.4 Serotonin as a neurotransmitter and its receptors

Serotonin, also known as 5-hydroxytryptamine (5-HT), is a major neurotransmitter found in the CNS and many peripheral tissues (14). It regulates several functions such as dopamine release, cognitive function, learning, memory, appetite, immune function, arousal, sexual desire, vascular tone, and coagulation (15). Serotonin plays an important role in controlling behaviour and mental status and has been implicated in the pathogenesis of many psychiatric disorders, including depression and anxiety (16).

The biosynthesis of serotonin takes place in the presynaptic serotonergic neurons of the CNS, where it is synthesized from the essential amino acid tryptophan (Trp). Tryptophan is converted to 5-hydroxytryptophan (5-HTP) by tryptophan hydroxylase and then to serotonin by aromatic amino acid decarboxylase (Fig. 3) (17).

The effects of serotonin are mediated by serotonin receptors. During the past two decades, multiple 5-HT receptors subtypes have been characterized ($5\text{-HT}_{1\text{-}7}$), based on their amino acid sequence homology, ligand affinity and signalling pathways. All the serotonin receptors belong to family A of GPCRs, with one exception the 5-HT_3 receptor, which is a ligand-gated ion channel receptor (8,14).

In 2013, the crystal structures of the 5-HT_{1B} and the 5-HT_{2B} receptor in complex with two full agonists ergotamine (ERG) and dihydroergotamine (DHE), respectively, were reported by Wang et al (18). These receptor structures represent an agonist bound state of the 5-HT_{1B} receptor and an arrestin biased state (the receptor conformation for interactions with arrestin) of the 5-HT_{2B} receptor. This has enabled construction of homology models of other serotoninergic GPCRs. Molecular docking of 5-HT into the orthosteric binding pocket of the 5-HT_{1B} and 5-HT_{2B} receptors revealed important residues involved in the recognition of 5-HT (18).

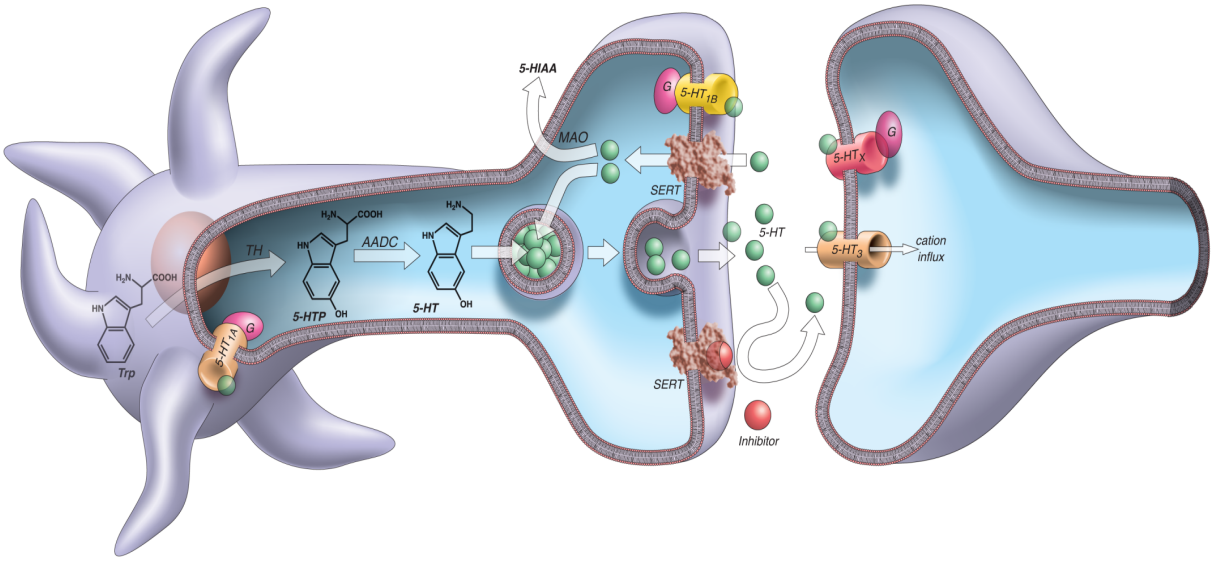


Fig. 3. A simplified illustration of the serotonergic system. Trp, tryptophan; TH, tryptophan hydroxylase; 5-HTP, 5-hydroxytryptophan; AADC, aromatic amino acid decarboxylase; 5-HT, 5-hydroxytryptamine; SERT, serotonin transporter; MAO, monamine oxidase; 5-HIAA, 5-hydroxyindolacetic acid; 5-HT_X, 5-HT receptors (17).

1.4.1 Serotonin receptors 5-HT_{1A} and 5-$HT_{2A/2C}$

The 5-HT_{1A} receptor inhibits adenylyl cyclase, increases the potassium conductance by regulating inward rectifying potassium channels, and decreases the opening of voltage gated calcium channels, and mainly function as an inhibitory presynaptic and postsynaptic receptor. The receptor has an important role in depression (19). When activated by serotonin, the 5-HT_{1A} receptor triggers the opening of potassium channels in the cell membrane and hyperpolarisation of the cell, which further results in a reduction in the discharge rate.

The 5-HT_2 receptor subfamily preferentially couples to $G_{q/11}$ to increase inositol phosphates and cystolic [Ca^{2+}].

The 5-HT_{2A} receptor is an important member of this family and this subtype of 5HT_2 receptors is widely distributed at varying densities throughout the brain. As for the 5-HT_{1A}, there are cumulative evidences indicating a role in depression (19). Many antidepressants and antipsychotic agents bind with high affinity to this receptor. Blockade of 5-HT_{2A} receptor might enhance the 5-HT_{1A} receptor-mediated neurotransmission in cortical and limbic areas. This effect is likely to be linked to the efficacy of antidepressants.

The 5-HT_{2C} receptor is predominantly located in the choroids plexus, cerebral cortex, hippocampus, substantia nigra and cerebellum. The 5-HT_{2C} receptor also plays a role in

depression and is involved in the actions of several antidepressant drugs. An altered editing of the mRNA encoding this receptor has been reported in the prefrontal cortex of depressed suicide victims, suggesting an abnormal function of the receptor protein. Preclinical studies have shown that selective and nonselective 5-HT_{2C} antagonists potentiate the neurochemical effects of SSRIs on hippocampal and cortical extracellular serotonin levels (19).

1.4.2 Serotonin transporter (SERT)

The serotonin transporter (SERT) belongs to the neurotransmitter sodium symporter (NSS) family and is one of the most widely studied NSS transporters. The SERT plays a key role in serotonergic signalling by removing 5-HT from the synaptic cleft into the presynaptic neuron. In presynaptic neurons, serotonin is either recycled into storage vesicles or converted to the inactive metabolite 5-hydroxyindolacetic acid (5-HIAA) by the monoamine oxidase (MAO) (Fig. 3) (17,20). The SERT can be inhibited by drugs, such as antidepressants, hence enhancing the serotonergic neurotransmission (21,22). This makes SERT an important pharmacological target.

X-ray crystal structures of the human SERT in complex with two antidepressants:
(S)-citalopram and paroxetine, were recently reported by Coleman et al (23).

Both an allosteric and an orthosteric binding site were identified in the X-ray crystal structures. These SERT structures define the mechanism of antidepressant action and give hopes for future drug design, and must be regarded as an important breakthrough in the search for new antidepressant drugs.

Until 2013, only one member of NSS transporters was solved by X-ray crystallography, the prokaryotic Aquifex aeolicus leucine transporter (LeuT). This is why homology modelling has been an important technique to study NSS transporters. Gabrielsen et al used the LeuT structure as template for constructing SERT homology models. The project's purpose was the identification of novel SERT compounds and several high affinity binders were identified by using the homology models for structure based Virtual Screening (VS) and experimental verification (20).

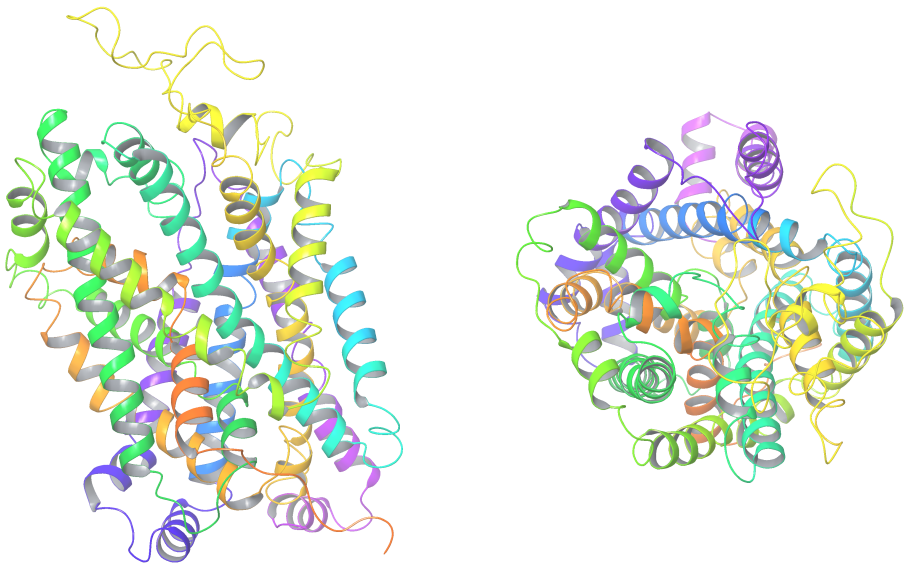


Fig. 4. Structure of human SERT imported from the PDB (PDB ID: 5I71) and protein prepared in Maestro. The backbone atoms of the transporter are shown. Left – viewed in the membrane plane. Right – viewed from the extracellular side.

1.5 Future aspects in pharmacological treatment of depression and anxiety

Clinical depression or major depressive disorder (MDD) is the most common psychiatric disorder worldwide. It is a medical condition that may cause serious, long-lasting symptoms and often disrupts a person's ability to perform routine tasks. It can take different forms and has varying levels of severity. People with untreated depression have lower quality of life and an increased risk for suicide (24).

It is reported 5-600 suicides every year in Norway, because of major psychiatric disorders. The society pays up to $8.5 billion (70 billion Norwegian Krone) annually, mostly on treatment of depression and anxiety (25).

The main theory of depression is the monoamine hypothesis proposed by Schildkraut in 1965 (26). It states that depression is caused by a functional deficit of the monoamine transmitters, noradrenaline and serotonin at certain sites in the brain, while mania results from the functional excess. The pharmacological evidence for that is summarized in Table 1.

Table 1. Pharmacological evidence supporting the monoamine hypothesis of depression (12).

Drugs	Principal action	Effect in depressed patients
Tricyclic antidepressants (TCA)	Block noradrenaline and 5-HT reuptake	Mood ↑
Monoamine oxidase (MAO) inhibitors	Increase stores of noradrenaline and 5-HT	Mood ↑
Selective serotonin reuptake inhibitors (SSRI)	Block 5-HT reuptake	Mood ↑
Reserpine	Inhibits noradrenaline and 5-HT storage	Mood ↓
Methyldopa	Inhibits noradrenaline synthesis	Mood ↓
α-Methyltyrosine	Inhibits noradrenaline synthesis	Mood ↓
Tryptophan depletion	Decreases brain 5-HT synthesis	Induces relapse in SSRI-treated patients
Tryptophan	Increases 5-HT synthesis	Mood ↑ (in some studies)

An important question is if environmental toxicants can contribute to the impairments of neurotransmission by interacting with serotonergic receptors and transporters, and thereby possibly could modulate the actions of antidepressant drugs.

Perhaps the strongest evidence for the role of the serotonergic system in MDD is the efficacy of antidepressants that target the SERT. These are the selective serotonin reuptake inhibitors (SSRIs) and the dual serotonin and norepinephrine reuptake inhibitors (SNRIs). The two mentioned groups of antidepressants account for more than 90% of the global antidepressant market (19).

SSRIs show selectivity with respect to serotonin over noradrenaline reuptake. They are less likely to cause anticholinergic side effects and are less dangerous than tricyclic antidepressants (TCAs) (12).

However, the antidepressant effect is not seen immediately. Beneficial effects from the treatment appear within one or two weeks, while the full effect may be seen after 6 to 12 weeks. It is important to mention that the direct neurochemical effects of antidepressant drugs appear very rapidly (within minutes to hours), while the adaptive changes in the brain responsible for the clinical improvements, take much longer time to obtain (12,24).

The antidepressant market is believed to be almost saturated with antidepressants, and new drugs can only achieve success if they prove to be faster acting and/or more efficacious than

SSRIs and SNRIs. Future development can rely on knowledge of the role played by the different 5-HT receptors in depression.

The presynaptic 5-HT_{1A} receptor is suspected to play an important detrimental role in the slow onset of action due to activation of negative feedback mechanisms taking place in 5-HT neurons. On contrary, activation of postsynaptic 5-HT_{1A} receptor in corticolimbic networks has positive antidepressant action. It is also shown that blockade of 5-HT_{2A} and 5-HT_{2C} improves the action of SSRIs.

One of the strategies in the development of new antidepressants drugs might include combinations of SERT blockade and agonising/antagonizing activities at most relevant serotonergic receptors. Vilazodone, a drug approved in 2011 in the United States, is demonstrating this ability by combining SERT inhibition with partial agonism at 5-HT_{1A} receptors (19).

1.5.1 Psychotherapy

Antidepressants are not the only method in treatment of the MDD. The first line of treatment of depression in Norway is non-medicament psychotherapy, assisted programmes for physical activity and advices on how to solve the problems on daily basis. Antidepressants are considered as an option only if psychotherapy attempt is not sufficient (27).

There are many different forms of psychotherapy, and a variety of techniques are used to treat their patients, depending on their life situation and degree of depression.

Options for treatments include: cognitive-behavioural therapy, problem solving therapy, supportive psychotherapy, psychodynamic psychotherapy, family and couples therapy.

Psychotherapy is generally not used alone for patients with severe depression. Major depression can be treated with antidepressants or psychotherapy, or a combination of both. Studies have shown that combination treatment is more effective than either treatment on its own (24).

1.6 Computational science and drug discovery

In drug discovery, computational methods have been playing a major role in the development of therapeutically important molecules for several years. Computer based methods allow rapid screening of large compound libraries and determination of potential binders through modelling, simulation and visualization techniques. These computational methods are classified either as structure-based or ligand-based. Ligand based methods take into account only ligand information, and predict activity based on its similarity or dissimilarity to previous known active ligands. On contrary, in structure-based methods both the structure of the target and of ligands binding to the target are known. Structure-based computer-aided drug design (SB-CADD) relies on the ability to determine and analyse 3D structures of biologic molecules. This approach predicts the ability of a molecule to favourably interact with a particular binding site on a specific protein.

In this project SB-CADD method will be applied through molecular modelling, including homology modelling and docking and scoring to examine if putative toxic compounds interact with serotonergic receptors and the SERT.

Software and databases are also a part of computational science. Structural information about the target protein is necessary for SBB-CADD approach. The Protein Data Bank (PDB) contains target structures experimentally determined through X-ray crystallography or NMR techniques (28).

1.6.1 Molecular modelling

Molecular modelling is a collection of computer based techniques for deriving, representing and manipulating the structures and reactions of molecules, and properties that are dependent on 3D structures. It generally accounts two computational approaches, molecular mechanics (MM) and quantum mechanics (QM) (29).

The molecular mechanics consider the atomic composition of a molecule to be a group of masses interacting with each other via harmonic forces. The method enables the calculation of the total steric energy (E_{tot}) of a molecule in terms of deviations from reference unstrained bond lengths (E_{bond}), angles and torsions (E_{angle}, E_{tors}), plus non-bonded interactions (E_{vdw}, E_{elec}), as shown in the following equation:

$E_{tot} = E_{bonded} + E_{non\text{-}bonded}$; or

$E_{tot} = (E_{bond} + E_{angle} + E_{tors}) + (E_{vdw} + E_{elec})$

The QM approaches are very valuable tools in computational chemistry. Properties like molecular geometry and relative conformational energies can be calculated with high accuracy for a broad variety of structures. However, QM is disadvantageous relative to other methods because of the computational costs and the limitation to rather small molecules.
Unlike the MM, QM is not applicable to large biological molecules such as proteins, DNA and lipid membranes (30).

1.6.2 Homology modelling

Homology modelling, also called comparative modelling is an *in silico* method used to predict the 3D structure of a query amino acid sequence based on a homologous experimentally determined structure (the template) (31).
This approach aims to establish the relationship between residues of the target and those of the template. For protein receptor models, the correspondence of binding pocket residues is of special concern.
Proteins that have evolved from a common ancestor are said to be homologous. Two homologous sequences can be nearly identical, similar to varying degrees or dissimilar because of extensive mutations. When a sequence with unknown protein 3D structure is found homologous to another protein sequence, and there is a 3D structure available for that sequence, the homology modelling approach is the method of choice for predicting the 3D structure of the protein sequence of unknown 3D structure. The prediction is based upon amino acid sequence alignments between the target and the template (30).
The first step in homology modelling is to determine which protein structures that can be used as templates. Templates for homology modelling can be found in the PDB. These are structures solved by X-ray crystallography or NMR. When choosing a template of known 3D structure, it is important to consider its crystallographic resolution if it is an X-ray structure (the R-factor). A lower R-factor indicates more detailed structural information which will give a more accurate 3D model. Structures with R-factor resolution of 2.0 Å or lower are considered to be reliable. Structure comparison of putative templates is also necessary in the process of template selection and alignment correction. The template has to be closely related

to the modelling target, having significant similar structural and functional properties. The sequence identity of the template and the target should be high, especially in the binding pocket, as it will provide a more reliable model conformation. If the template chosen is crystallized with a ligand, the target model will also be in a more appropriate state for ligand docking (30).

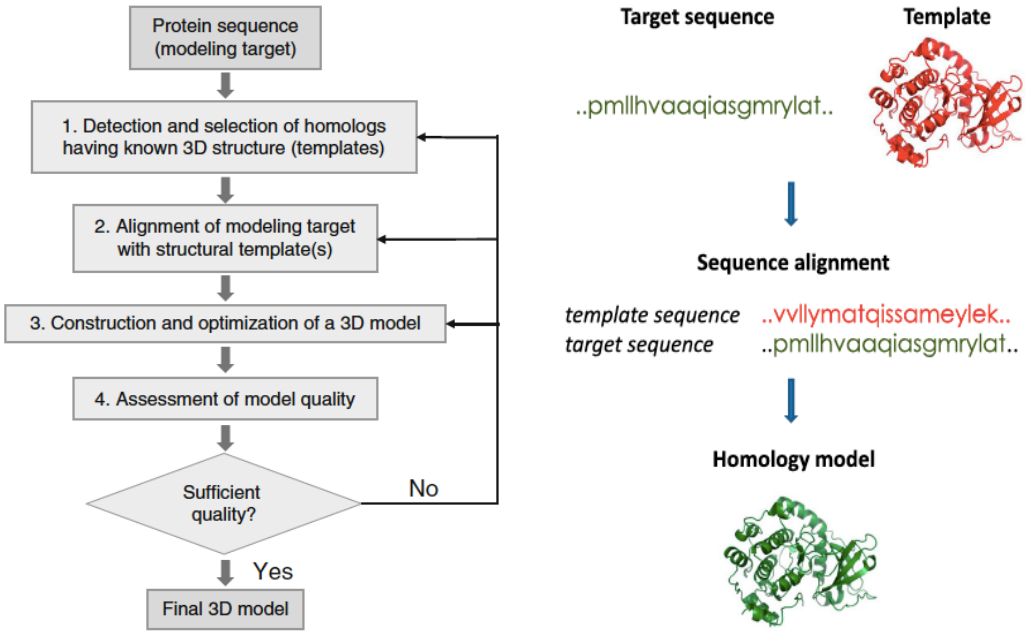


Fig. 5. Steps in homology modelling (31).

The second step in homology modelling is mapping of corresponding residues between the target sequence and template structure, the process often referred to as sequence–structure alignment. The sequence–structure alignment aims to reproduce this correspondence as accurately as possible, but without the benefit of knowing the "real" (experimental) structure of the modelling target. An accurate sequence–structure alignment should include all the structurally and evolutionary equivalent residue pairs, at the same time leaving out structurally different regions (31).

The third step in homology modelling is generation of a 3D model of a target protein on the basis of the sequence–structure alignment. At this stage, 3D properties of the modelling target can be adjusted, to resemble the real structure as much as possible. For instance, the family A of GPCRs is characterized by the presence of cysteine residues involved in the formation of extracellular disulfide bridges. It is of importance that the presence of cysteine residues and their putative involvement in the formation of disulfide bridges is identified prior to the

construction of the model (31). Following the construction, the receptor model can be refined using energy minimizations, Monte Carlo simulations, or molecular dynamics simulations. These simulations are designed to explore as much of the relevant regions of conformation spaces as possible (29). Model refinement is performed to relax high-energy structures and remove close contacts between atoms.

The fourth and final step in homology modelling is estimating the correctness of the resulting model. Once a protein model has been built, it is necessary to assess its quality and reliability. This can be verified by examining its stereochemical accuracy, packing quality and folding reliability. The accuracy of bond lengths, bond angles, torsion angles and the correctness of amino acid chirality has to be proved. The interior packing quality of a protein model is a major contributor to the stability of the overall conformation and can thus be used to estimate its reliability. The folding correctness can be measured by the 3D-Profiles method, an approach based on the general principle that the 3D structure of a protein must be compatible with its own amino acid sequences (30).

Evaluation of the model can be performed by online analyzing tools such as SAVES (Structural Analysis and Verification Server) at http://services.mbi.ucla.edu/SAVES/, by site directed mutagenesis studies based on the model and docking of known binders and decoys to the target protein.

1.6.3 Docking and scoring

The docking process involves the prediction of ligand conformation and orientation within a binding site in a target protein structure or homology model. There are two aims of docking studies: accurate structural modelling and correct prediction of affinity/activity. The process of docking begins with the application of docking algorithms to place small molecule ligands into the active site of a transporter, enzyme or the binding site of a receptor (termed ligand poses). These algorithms are additionally complemented by scoring functions that are designed to predict the affinity through the evaluation of interactions between compounds and potential targets (32,33). The scoring functions can be grouped into three basic types according to how they are derived: force field-based, empirical, and knowledge-based. Force field-based scoring derives from physical atomic interactions, including van der Waals interactions, electrostatic interactions and bond-stretching, bending and torsional forces.

Empirical scoring functions estimate the binding affinity of a complex on the basis of a set of weighted energy terms

$$\Delta G = \sum_i W_i \cdot \Delta G_i$$

where ΔG_i represents different energy terms such as van der Waals, electrostatics, hydrogen bonding, entropy and hydrophobicity energy terms. Knowledge-based scoring functions employ energy potentials derived from structural information embedded in experimentally determined atomic structures. Pairwise potentials are directly obtained from the occurrence frequency of atom pairs in a database using the inverse Boltzmann relation. The potentials are calculated by

$$w(r) = -k_B T \ln[g(r)], \quad g(r) = \rho(r)/\rho^*(r)$$

where k_B is the Boltzmann constant, T is the absolute temperature of the system, $\rho(r)$ is the number density of the protein–ligand atom pair at distance r, and $\rho^*(r)$ is the pair density in a reference state where the interatomic interactions are zero (34).

2. Aim of the study

The main aim of the project is to use a structure based molecular modelling approach to predict the putative interactions of 8 194 chemical substances in the Tox21 (Spring 2012) database with serotonergic GPCRs and with the SERT. If the chemical substances interact with these proteins, they may interfere with the serotonergic neurotransmission in the brain. These proteins are also important CNS drug targets and the chemicals may interfere with the effects of CNS drugs.

Models of serotonergic receptors (5-HT_{1A}, 5-HT_{2A}, 5-HT_{2C}) will be constructed by homology modelling with known structures in the PDB database. Newly released X-ray crystal structures of the human SERT will also be imported from the PDB and optimized with molecular modelling techniques. Molecular docking will be utilized to predict putative harmful effects and drug interactions of the toxicants in the Tox21 database with serotonergic receptors and the SERT. Known agonists and antagonists will be docked into the receptor models as well as inactive compounds with similar physiochemical properties to predict how well the models discriminate known binders from decoys. The CNS MPO tool will give ranking scores for toxicants, to predict their physicochemical abilities to reach the CNS. This study will provide insight into detailed prediction of ligand interactions of serotonergic receptors and the SERT including the molecular interactions of antidepressant drugs.

3. Methods

3.1 Software and databases

3.1.1 Molsoft Internal Coordinates Mechanics (Version 3.8.4)

The Molsoft molecular modelling technology is based on the Internal Coordinate Mechanics (ICM) approach which gives a general modelling and structure prediction framework for many tasks of structural biology and rational drug design (35). The ICM program was used to build homology models of the receptors, and to convert receptor binders and decoys from their SMILES code to 2D structures.

3.1.2 Schrödinger (Version 2015.3)

Schrödinger is a software company that has developed products for practicing computational chemistry. The company provides several software products for molecular modelling and molecular design for the use in drug development. Maestro is the graphical user interface (GUI) for nearly all of Schrödinger's computational programs (36). This program was used to refine the previously constructed homology models from ICM, and build new receptor models based on induced fit docking (IFD). The virtual screening workflow (VSW) option was used in the docking of sets of ligands and decoys into the homology models. Python-scripts developed by collaborators in Krakow, were used through Maestro to handle the ligand-protein interactions.

3.1.3 The Protein Data Bank (PDB)

The crystal structures of serotonin receptors, SERT and other homologous proteins of known structure (adrenergic and dopaminergic receptors) were downloaded from the PDB (http://www.rcsb.org/pdb/home/home.do). The PDB is an archive of information about the experimental determined 3D structures of large biological molecules, such as proteins and nucleic acids, mostly solved by X-ray crystallography techniques (37). At present, there is about 117 000 3D structure complexes in the PDB, of which only around 1 000 are membrane

proteins. This makes homology modelling a very important alternative for studying the structure of membrane proteins. The PDB archive contains receptor structures that provided templates for the receptors studied.

3.1.4 UniProt Knowledge-Base

UniProtKB is a database that provides resources of protein sequence and functional information (http://www.uniprot.org) (38). Unlike the PDB, which provides the 3D structure templates, UniProtKB gives access to the amino acid sequences of the target receptors. All sequences are given a unique accession number for each receptor.

3.1.5 ChEMBL

ChEMBL is a database containing binding, functional and ADMET (absorption, distribution, metabolism, excretion and toxicity) information for a large number of drug-like bioactive compounds (39). The ChEMBL database is accessible via a simple interface at: https://www.ebi.ac.uk/chembldb. This interface allows users to search for compounds, targets or assays of interest in a variety of ways.

3.1.6 IUPHAR/BPS Guide to PHARMACOLOGY

The IUPHAR/BPS Guide to PHARMACOLOGY is a database created by the British Pharmacological Society (BPS) and the International Union of Basic and Clinical Pharmacology (IUPHAR) (http://www.guidetopharmacology.org). This database provides pharmacological information of receptors, prescription medicines and experimental drugs that act on them (40).

This database was used to generate ligands known to be active with high affinity for serotonin receptors or the SERT.

3.1.7 Database of Useful Decoys-Enhanced (DUD-E)

The Database of Useful Decoys-Enhanced (DUD-E) (http://dude.docking.org) was developed by Mysinger et al. in 2012 and is an improved version of The Directory of Useful Decoys (DUD). The DUD-E is designed to help benchmark molecular docking programs by providing challenging decoys (41). It is largely based on the intersection of ChEMBL, for ligand annotations and affinities, and the PDB, for structures. This database was used to generate the decoys with similar physicochemical properties, assumed to be inactive for the serotonergic receptors.

3.1.8 Tox21 database (Version: Spring 2012)

Tox21 is a database library that contains 8 194 environmental chemicals (release Spring 2012) and approved drugs with a potential to disrupt biological pathways that may result in toxicity. Tox21 researchers aim to develop better toxicity assessment methods to quickly and efficiently test whether certain chemical compounds have the potential to disrupt processes in the human body that may lead to negative health effects (5). The potential binding of the chemicals in the Tox21 database for the modelled receptors and SERT was predicted by docking.

3.2 Homology modelling

3.2.1 Template identification

The X-ray structures chosen as templates were structures with high resolution and a bound ligand (agonist or antagonist) included at the binding site.

In order to construct high quality homology models of the 5-HT_{1A}, 5HT_{2A} and 5-HT_{2C} receptors, X-ray structures with high sequence similarities with target receptors were chosen. The sequence similarity between the target receptors and the templates are shown in Table 3.

Table 3. Sequence similarity between the targets and their putative templates obtained in Maestro. Templates with the highest sequence similarity were used (highlighted in red).

Targets	Agonist bound templates		Antagonist bound templates	
	5-HT_{1B}	**5-HT_{2B}**	**D_3**	**β_2**
5-HT_{1A}	17%	15%	47%	23%
5-HT_{2A}	17%	27%	20%	35%
5-HT_{2C}	17%	21%	38%	13%

Based on the percent sequence similarity, the following templates were downloaded from the PDB:

- Biased agonist bond state: Crystal structure of the chimeric protein of 5-HT_{2B} in complex with ergotamine, R-factor of 2.7 Å; PDB ID: 4IB4
- Agonist bound state: Crystal structure of the chimeric protein of 5-HT_{1B} in complex with ergotamine (PSI community target), R-factor of 2.7 Å; PDB ID: 4IAR
- Antagonist bound state: Crystal structure of human β_2-adrenergic G-protein coupled receptor, R-factor of 2.4 Å; PDB ID: 2RH1
- Antagonist bound state: Crystal structure of the human dopamine D_3 receptor in complex with eticlopride, R-factor of 2.89 Å; PDB ID: 3PBL

The X-ray crystal structures of human SERT were also downloaded from the PDB:

- X-ray structure of the ts3 (human SERT with three thermostabilizing mutations Tyr110Ala, Ile291Ala and Thr439Ser) complexed with (S)-citalopram at the substrate binding (central) site, R-factor of 3.15 Å; PDB ID: 5I71
- X-ray structure of the ts3 human SERT complexed with (S)-citalopram at the central and allosteric sites, R-factor of 3.24; PDB ID: 5I73
- X-ray structure of the ts3 human SERT complexed with paroxetine at the central site, R-factor of 3.14 Å; PDB ID: 5I6X

3.2.2 Sequence alignment

Sequence alignments of serotonergic and melatonin receptors were prepared by using the Molsoft ICM program. The templates were aligned to show the conserved regions within the serotonergic and melatonin receptors. Melatonin receptors are believed to be structurally quite similar to the serotonergic receptors, and it is of interest to study the sequential conservation between melatonin receptor and the serotonergic receptors and they were therefore included in the alignment.

For some templates, the T4-lysozyme part of the structure had to be removed by the text editor and the gaps in TM helices were adjusted manually in the ICM, all this to make sure that the conserved residues in helices and cysteine bridges between the helices were in correct position (Fig. 6).

Fig. 6. Sequence alignment obtained by the ICM software.

3.3 Molecular docking

3.3.1 Generation of binders and decoys

The known binders for the target receptors were downloaded from IUPHAR/BPS database into the ICM software. The chemistry module in ICM was used to convert SMILE codes to ligand 2D structures. The same tool was used to protonate the aliphatic amino groups of the ligands, important for ionic ligand-receptor interactions. The charges were set at a pH of 6.5.
To generate decoys, the SMILE codes had to be uploaded to the website in the panel: Generate → SMILES input (41). According to the website, there are 50 decoys obtained for each ligand, and for this to be statistically significant, decoys were generated for 4 times. The collected decoy compounds were then imported to the ICM software, where the same procedure was performed as for the known binders. Additionally, the duplicate decoys had to be selected and removed from the file via the option: Chemistry → Select duplicates → Delete.

3.3.2 Preparation of ligands, receptors and transporter for docking

The known binders (previously generated in ICM) for serotonin receptors and the SERT were imported to Maestro as an sdf file. These known binders were firstly protonated in ICM (pH set to 6.5, as described), such that they could interact with an aspartic acid residue. A salt bridge interaction between an Asp present in serotonin receptors and transporter, and the protonated amine moiety in ligands is important for binding. The task Ligand Preparation (LigPrep) was then used to convert 2D ligand structures to 3D conformations.
The homology models of serotonin receptors previously constructed in ICM software, and the X-ray structures of the SERT (obtained directly from the PDB) were also imported to Maestro. These were then prepared by the task Protein Preparation (ProteinPrep). ProteinPrep adjusts missing hydrogen atoms, builds in missing residues and loops, identifies overlapping atoms, assigns missing bond orders, and optimizes the hydrogen-bonding network.

3.3.3 Induced fit docking (IFD)

Induced fit docking (IFD) is a molecular docking method that can predict ligand binding modes and concomitant structural changes in the receptor. Unlike other docking methods that assume a rigid receptor, in IFD receptors adopt the binding site to conform to the shape and binding mode of the ligand. Schrödinger's Maestro possesses Prime and Glide modules that account for possible binding modes and associated conformational changes within receptor binding sites. These two modules were applied in the IFD protocol.

When ligand and protein preparation were conducted, we were ready to run an IFD.

To perform IFD, known binders with high affinity for each of the 5-HT_{1A}, 5-HT_{2A} and 5-HT_{2C} receptors were selected. The purpose was to optimize the binding pocket in each of the receptor models in complex with a high affinity binder. The important interaction between the carboxylate oxygen atom(s) in Asp3.32 (Ballesteros-Weinstein numbering scheme) in TM helix III (the amino acid conserved in all amine GPCRs) and the protonated amine of ligands was included. Inserting a constraint (Asp), implies requirement for a ligand to make a hydrogen bond with the carboxylate oxygen atom(s) of Asp3.32 in the receptor. The choice of ligands was made upon their mechanism of action (agonist/antagonist), affinity and sometimes selectivity for the target receptor. An overview of the target receptors and the ligands acting on these receptors used for IFD are presented in Table 4.

Table 4. Showing high affinity binders chosen for IFD and constraints of target receptors. Affinity values imported from www.guidetopharmacology.org (40).

Receptor	Constraint*	Ligand	Pharmacological activity	Affinity (p*K*i)
5-HT_{1A}	Asp 116	LY293284	Full agonist	10.1
		rec 15/3079	Antagonist	9.7
5-HT_{2A}	Asp 155	methylergonovine	Full agonist	9.4
		asenapine	Antagonist	10.2
5-HT_{2C}	Asp 134	YM348	Full agonist	9.0
		serindole	Inverse agonist	9.0 – 9.2

* carboxylate oxygen – protonated amine hydrogen in all aspartic acid residues.

IFD was run by the commands in Maestro as follow: Task → Docking → IFD → Browse high affinity ligand → Receptor centroid of selected residues: Constraint Asp3.32 → Protein preparation constrained refinement (PPCR) → Glide redocking: 20 → Jobs: Number of Glide 50; Number of Prime 50 → Run.

3.3.4 Virtual screening workflow (VSW)

The IFD generated conformations of the models were then used for docking the entire library of binders and decoys using the virtual screening workflow (VSW). Like regular docking and IFD, the VSW is also a molecular docking method for predicting ligand binding to a target receptor. The VSW assumes a rigid receptor and can assay a very large number of compounds.

Docking of decoys and active ligands was performed in Maestro by the following commands: Docking → VSW → Input of prepared actives/decoys → Remove LigPrep option → Receptor: Add in grids → Docking (SP Glide): Option 100% → Job write.

The job was then written in the directory as an input file. The input file was run on the command line in Linux (computer operating system) terminal, where the add_constraint.sh and features.txt scripts were added. The add_constraint.sh file was added to the input file to ensure the crucial interactions between the conserved Asp3.32 carboxylate group and the protonated amine in the actives/decoys, when docked. On the other hand, feature.txt is the constraint for both receptor and actives/decoys, where CONSTRAINT_GROUP is the constraint from the receptor side and FEATURE for actives/decoys. The feature.txt script had to be edited in the following way:

```
[[CONSTRAINT_GROUP:1]]
    USE_CONS   hbond1:1,
    NREQUIRED_CONS   ALL
  [[FEATURE:1]]
    PATTERN1   "[#1][NH+] 1 include"
      PATTERN2   "[#1][NH2+] 1 include"
      PATTERN3   "[#1][NH3+] 1 include"
```

3.3.4 Evaluation of models with BEDROC

BEDROC is a statistical approach used to evaluate the performance of ranking methods in virtual screening (VS) (42). To create BEDROC, the robust initial enhancement (RIE) and receiver operating characteristic curve (ROC) metrics were combined. This implied the BEDROC metric to receive the discrimination power of the RIE metric and the statistical significance from ROC. The problem of the ROC metric alone was the "early recognition problem" and this disadvantage is compensated by the RIE. A successful VS must rank actives early in the set of compounds since a very small proportion of the compounds will be tested experimentally (42).

BEDROC was used to statistically evaluate how the different homology models differentiate between actives and decoys. This task had to be performed in the Linux main terminal. The following command is an example of how the task was performed in the main terminal:

```
$SHRODINGER/run ~/script_used/enrichment_runner.py –n 20 –l ../../ligands/*.mae –d ../../decoys/*mae ../../docking/vsw_agonist/*SP*maegz ../../docking/vsw_decoys/*SP*maegz
```

3.3.5 Glide docking calculations with exogenous toxicants

The homology models with the highest BEDROC scores (after the VSW of binders and decoys) were used for docking of the compounds in the Tox21 database (9 757 after LigPrep) by using Glide. The protein prepared structures of SERT were directly Glide docked with both the known binders (inhibitors) and toxicants, but without use of the constraint option in Maestro. Glide approximates a complete systematic search of the conformational, orientational, and positional space for the docked ligands (43).

The toxicants were imported to Maestro as an sdf file and the compounds were prepared following the same procedure as for the binders and decoys (described in section 3.3.2).

This was achieved by choosing the following options in Maestro: Tasks → LigPrep → Browse sdf file of the toxicants → Generate possible states at target pH – 6.5 → Desalt and generate tautomers options – Removed → Generate specified chiralities at most 2 per ligand → Run.

The toxicants that have obtained the best Glide score were then visually inspected to study their interaction mode with the target models and the SERTs, and if these interactions resemble those of the known binders.

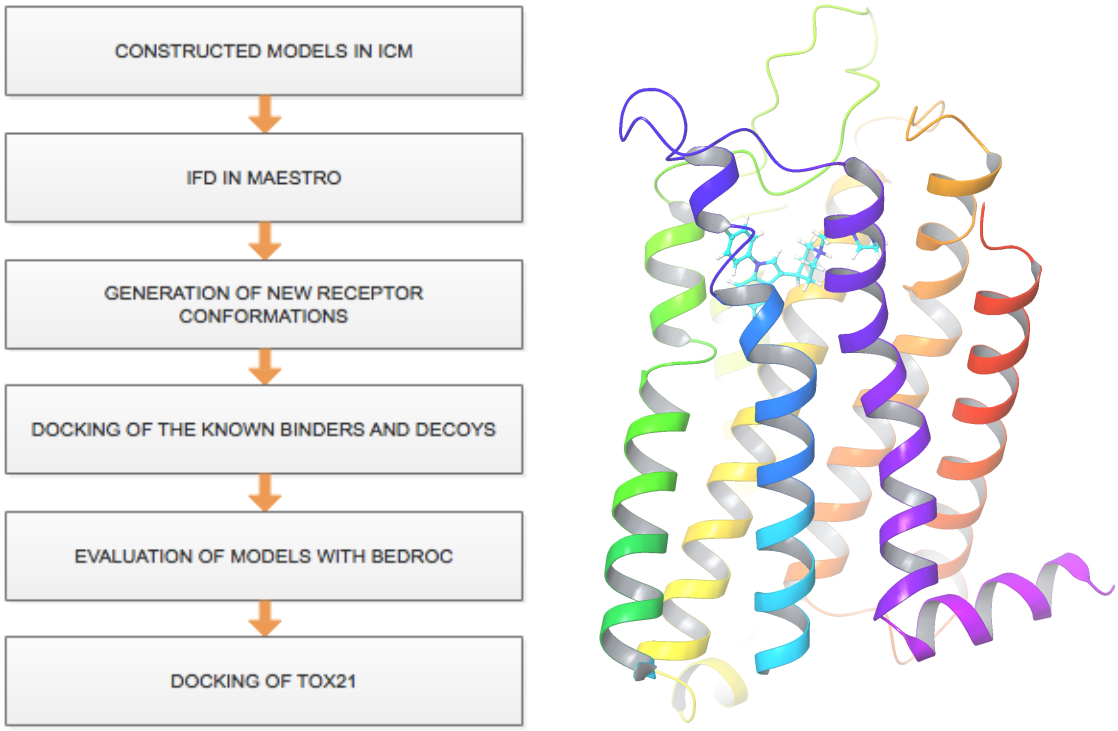


Fig. 7. Right – An overview of the steps utilized in the project described in sections 3.2 and 3.3. Left – An example of a receptor-ligand complex in 5-HT_{2C} receptor modelled (final stage). The backbone atoms of the receptor are shown.

3.4 Prediction of the BBB passage by CNS MPO tool

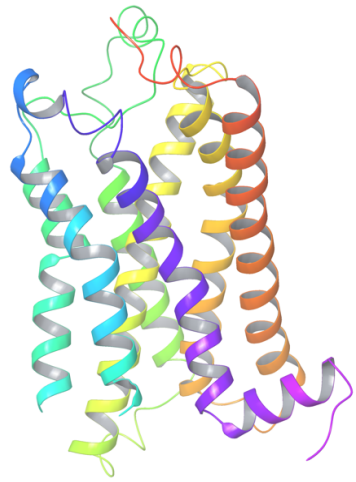
In 2010 Wager et al developed a novel medicinal chemistry tool named Central Nervous System Multiparameter Optimization (CNS MPO) (44). This tool is based on a set of physicochemical properties with purpose of enabling greater flexibility in CNS compound design beyond the use of a single parameters, expanding design space, and enhancing the odds of identifying compounds. The CNS MPO algorithm accounts a set of 6 fundamental physicochemical parameters (ClogP, ClogD, MW, TPSA, HBD, and p*K*a) and a variation of Harriongton's optimization method, which is a summation of the individual components to yield a composite desirability score.

The CNS MPO tool was applied in this project to predict the physicochemical properties of the environmental toxicants and drugs of the Tox21 database in order to predict if they have physiochemical properties that enable them to reach the CNS.

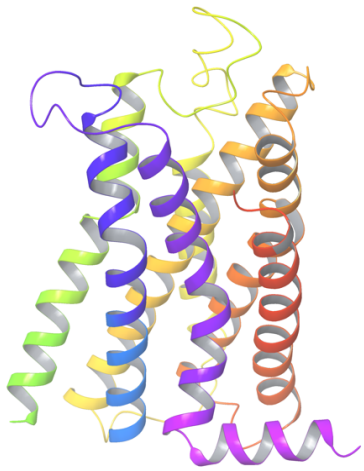
4. Results

4.1 Homology models

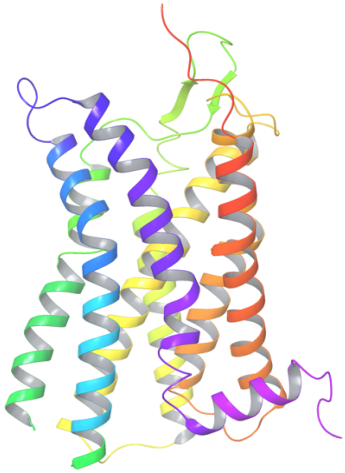
A total of 6 receptor models were built by using the homology modelling in ICM (Fig. 8). The models were agonist and antagonist bound states of 5-HT_{1A}, 5-HT_{2A} and 5-HT_{2C} receptors.
The agonist bound 5-HT_{1A} receptor model was constructed with the 5-HT_{1B} receptor structure (PDB ID: 4IAR) as the template, while the other two agonist bound models were constructed with the 5-HT_{2B} receptor structure (PDB ID: 4IB4). The antagonist bound 5-HT_{2A} receptor model was constructed from a β_2 receptor structure (PDB ID: 2RH1), while the antagonist bound 5-HT_{1A} and 5-HT_{2C} receptor models were constructed from the dopamine D_3 receptor structure (PDB ID: 3PBL). A model refinement was performed for all these models, and protein health score was predicted. The protein health calculates the energy strain of a structure. The protein health score was < 6 for all residues in binding pockets, which is considered as good, while some residues in loops showed protein health score > 6. However, the binding pockets were in focus, and loop modelling to improve loop quality was not prioritized. The scores indicate high quality models that can be used for docking.
For further optimizations of the models and binding pockets, the models were imported to Maestro to perform IFD with high affinity binders as given in Table 4.



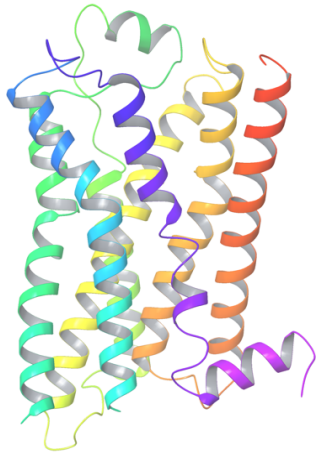
5-HT_{1A} agonist bound state



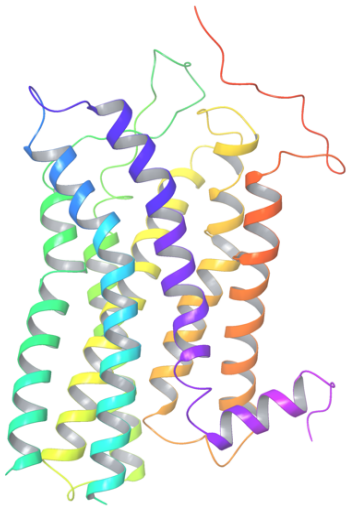
5-HT_{1A} antagonist bound state



5-HT_{2A} agonist bound state



5-HT_{2A} antagonist bound state



5-HT_{2C} agonist bound state

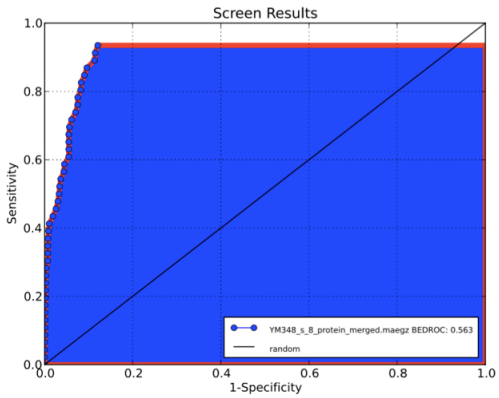
5-HT_{2C} antagonist bound state

Fig. 8. Models of agonist and antagonist bound states of serotonin receptors obtained in ICM and protein prepared in Maestro. The backbone of the receptor models is shown. The extracellular side is up in the illustrations.

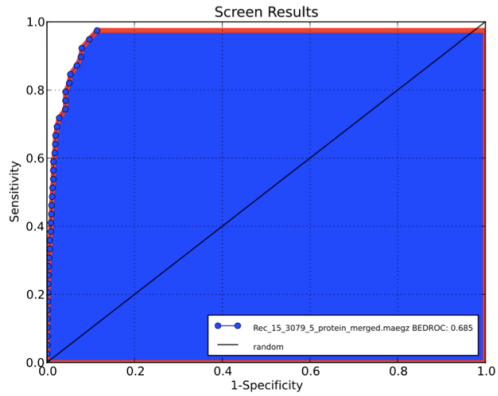
4.2 Molecular docking

4.2.1 Results from BEDROC calculations

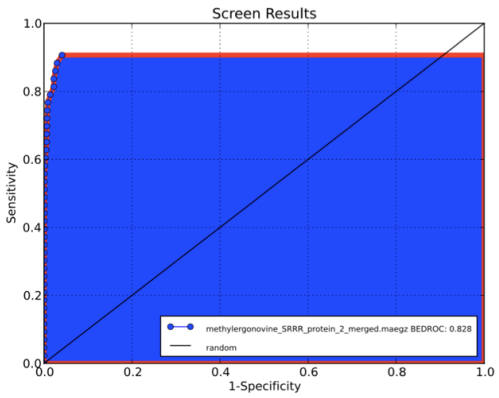
BEDROC calculations were used to evaluate how the homology models differentiate between actives (known binders) and decoys. The BEDROC values were obtained for each of the receptor output model from IFD, and the 6 best BEDROC-curves (one for each receptor model) are presented in Fig. 9. The best model of 5-HT_{1A} agonist bound state had a BEDROC value of 0.563, while the best model for antagonist bound state of the same receptor had a BEDROC value of 0.685. The best model of 5-HT_{2A} agonist bound state had a BEDROC value of 0.828, while the best model for antagonist bound state had a BEDROC value of 0.498. The best model of 5-HT_{2C} agonist bound state had a BEDROC value of 0.563, while the best model for antagonist bound state had a BEDROC value of 0.664.



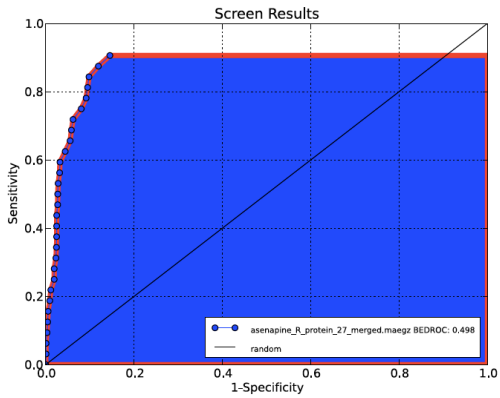
5-HT$_{1A}$ agonist bound state: Model 7



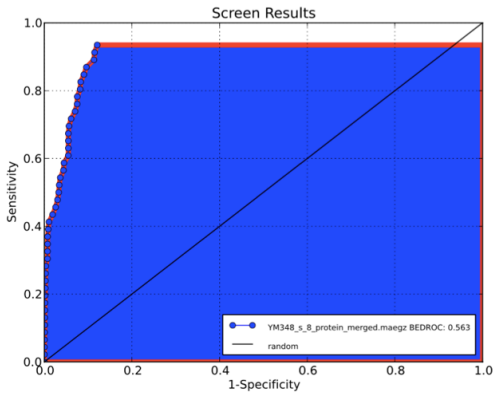
5-HT$_{1A}$ antagonist bound state: Model 5



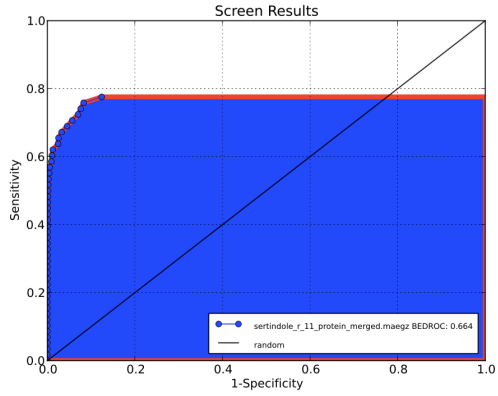
5-HT$_{2A}$ agonist bound state: Model 2



5-HT$_{2A}$ antagonist bound state: Model 27



5-HT$_{2C}$ agonist bound state: Model 8



5-HT$_{2C}$ antagonist bound state: Model 11

Fig. 9. BEDROC-curve for each of the three serotonin receptors (agonist/antagonist bound), showing the scoring values of ligands and decoys. X-axis represents the ranking of the ligand set (specificity). Y-axis represents the number of actives scored (sensitivity).

4.2.2 Induced fit docking

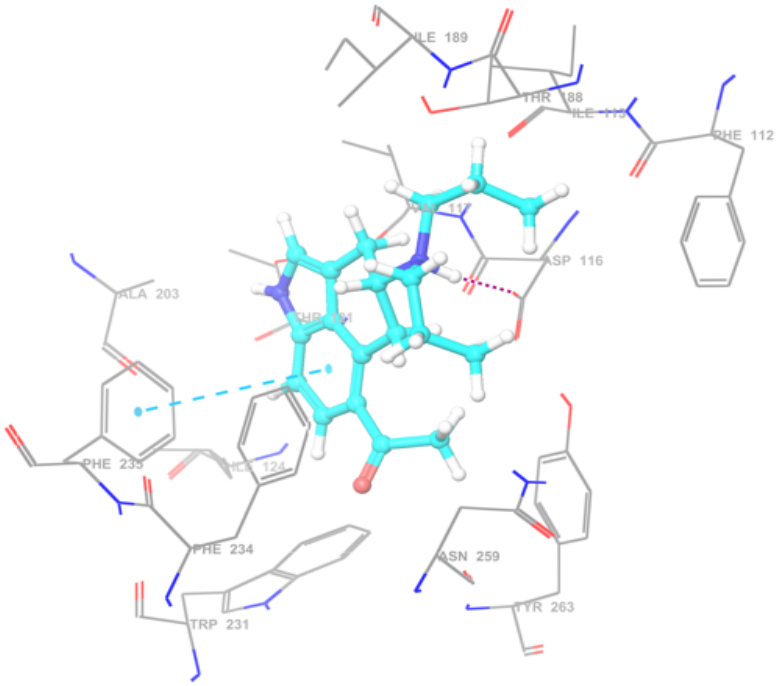
The IFD was performed by choosing known high affinity binders for receptor models. The scoring results of the best models (selected based on the highest BEDROC score) are given in Table 5.

Table 5. Scoring values of known binders in homology models. Negative scores indicate stronger interactions with the receptor. Affinity values imported from www.guidetopharmacology.org (40).

Receptor	Ligand	IFD score	Glide score	Affinity (p*K*i)
5-HT$_{1A}$	LY293284	-478.63	-6.35	10.1
	rec 15/3079	-467.76	-9.00	9.7
5-HT$_{2A}$	methylergonovine	-521.70	-9.21	9.4
	asenapine	-504.67	-7.09	10.2
5-HT$_{2C}$	YM348	-541.02	-8.23	9.0
	serindole	-485.75	-9.83	9.0 – 9.2

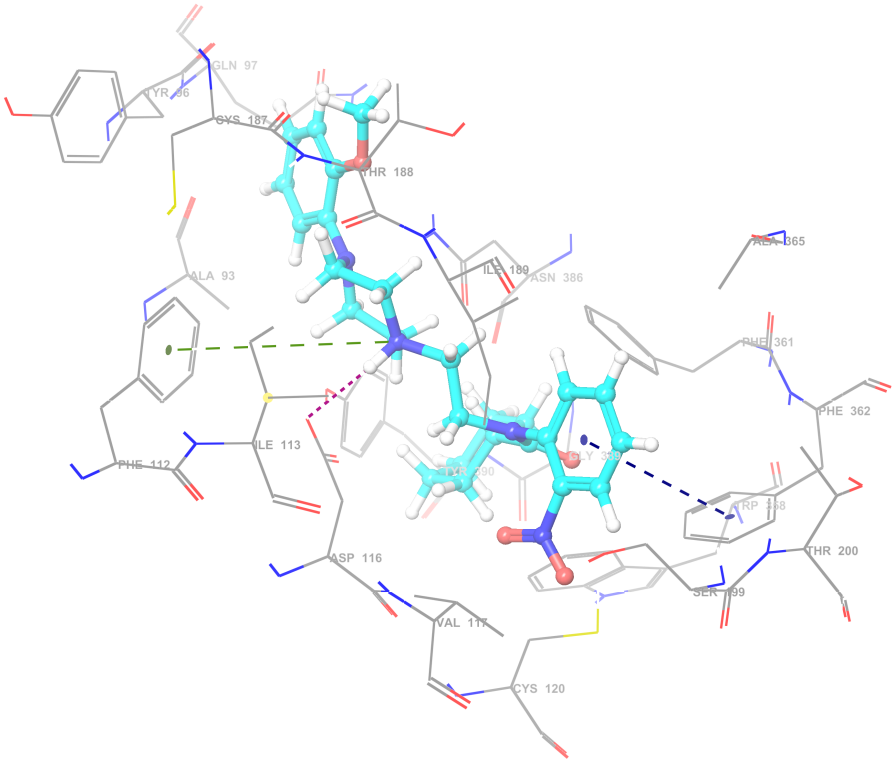
The IFD step was important for optimizing the binding pockets of the models and for clarifying the models for docking of known binders and toxicants into the best outputs. The IFD resulted in the generation of new receptor complexes (outputs), which were used for further docking. The IFD gave a varying number of receptor-ligand complexes as output between the different receptors. The 5-HT$_{1A}$ antagonist bound receptor yielded 16 new model complexes, the 5-HT$_{1A}$ antagonist bound receptor yielded 17 new model complexes, the 5-HT$_{2A}$ agonist bound receptor yielded 32 new model complexes, the 5-HT$_{2A}$ antagonist bound receptor yielded 36 new model complexes, the 5-HT$_{2C}$ agonist bound receptor yielded 32 new model complexes and the 5-HT$_{2C}$ antagonist bound receptor yielded 41 new model complexes.

One agonist and one antagonist conformational state for each receptor model (selected by the best BEDROC score) was later utilized for docking of the toxic compounds in the Tox21 database. These models in complex with their known binders are shown in Figs. 10, 11 and 12.



Known binder: LY293284

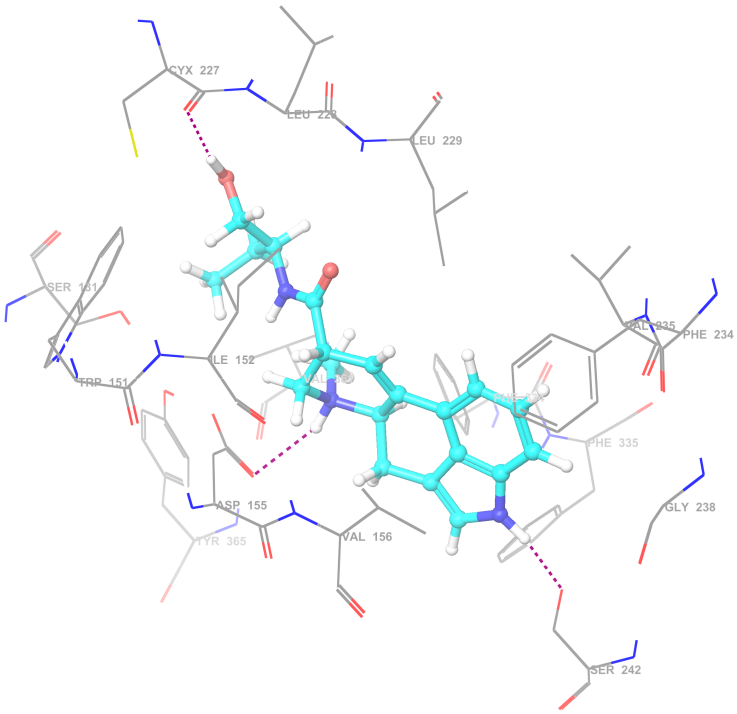
5-HT$_{1A}$ agonist bound receptor: Model 7 from IFD



Known binder: rec 15/3079

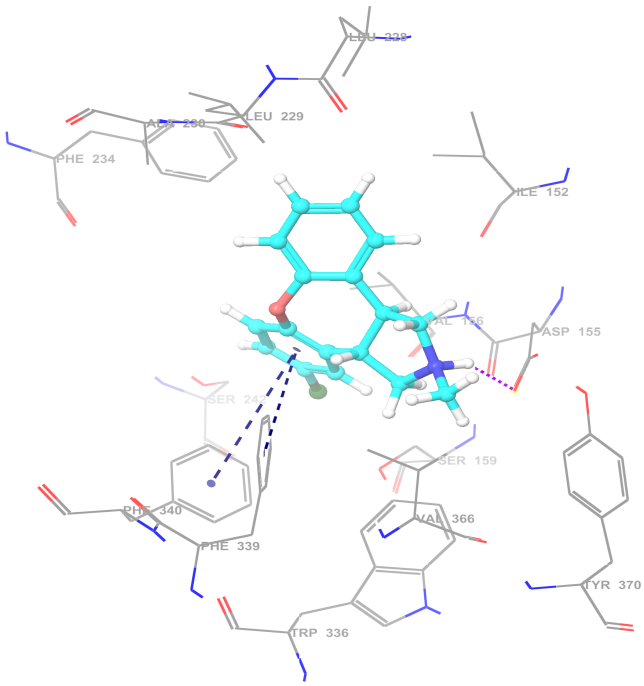
5-HT$_{1A}$ antagonist bound receptor: Model 5 from IFD

Fig. 10. 3D representation of ligand interactions of the best homology models obtained by IFD with known binders. Residues within 3Å sphere radius of the docked ligand are shown. Interactions are marked off with dashed line: blue/turquoise = aromatic, green = π-cation, purple = H-bonds.



Known binder: methylergonovine

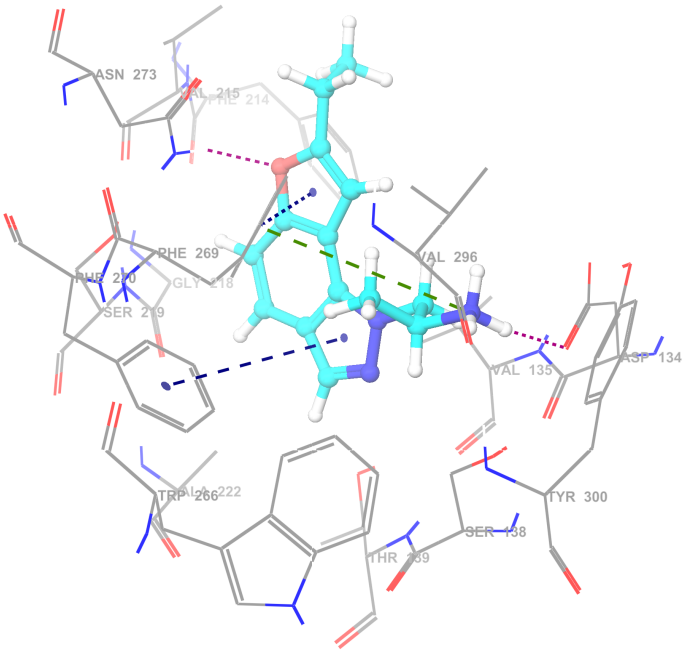
5-HT_{2A} agonist bound receptor: Model 2 from IFD



Known binder: asenapine (R)

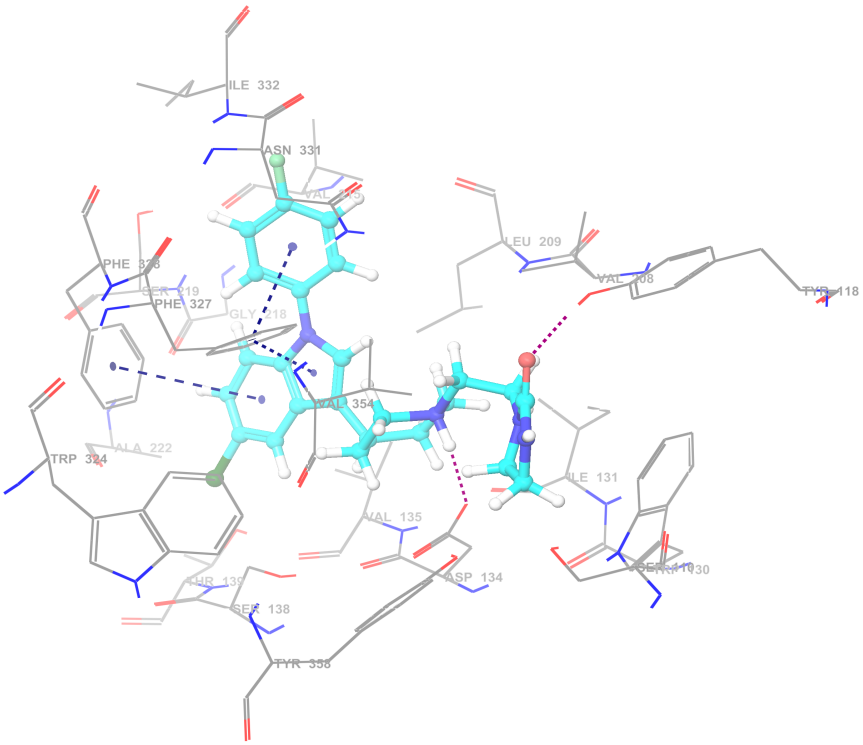
5-HT_{2A} antagonist bound receptor: Model 27 from IFD

Fig. 11. 3D representation of ligand interactions of the best homology models obtained by IFD with known binders. Residues within 3Å sphere radius of the docked ligand are shown. Interactions are marked off with dashed line: blue/turquoise = aromatic, green = π-cation, purple = H-bonds.



Known binder: YM348 (S)

5-HT$_{2C}$ agonist bound receptor: Model 8 from IFD



Known binder: sertindole (R)

5-HT$_{2C}$ antagonist bound receptor: Model 11 from IFD

Fig. 12. 3D representation of ligand interactions of the best homology models obtained by IFD with known binders. Residues within 3Å sphere radius of the docked ligand are shown. Interactions are marked off with dashed line: blue/turquoise = aromatic, green = π-cation, purple = H-bonds.

4.2.3 Docking of known binders into receptor models obtained by IFD

The known binders and decoys were docked into the new conformations (one for each receptor) by the VSW to predict their binding to the target receptors. The number of known binders docked in 5-HT_{1A} agonist bound state was 82 agonists with 19 857 decoys. 5-HT_{1A} antagonist bound state was docked with 61 known antagonists and 10 119 decoys, 5-HT_{2A} agonist bound state with 66 known agonists and 12 873 decoys, 5-HT_{2A} antagonist bound state with 55 known antagonists and 9 851 decoys, 5-HT_{2C} agonist bound state with 70 known agonists and 13 446 decoys, and 5-HT_{2C} antagonist bound state with 89 known antagonists and 19 328 decoys.

The 10 best Glide score results of the known binders were chosen to predict how the Glide scoring correlated to the experimentally determined affinity values (Tables 6, 7 and 8).

Table 6. Showing the 10 compounds of the known binders with best Glide score for the 5-HT_{1A} receptor models. Mean score of the 10 compounds was set to be a threshold value for scoring of the toxicants. Some ligands appear more than one time due to the R and S configuration states. Affinity values imported from www.guidetopharmacology.org (40).

5-HT_{1A} agonist bound state (model 7)			**5-HT_{1A} antagonist bound state (model 5)**		
Known binders	*Glide score (kcal/mol)*	*Affinity (pKi)*	*Known binders*	*Glide score (kcal/mol)*	*Affinity (pKi)*
"L-772,405"	-8.34	7.2*	"(S)-flurocarazolol"	-9.17	7.5
"L-694,247"	-8.12	9.3	"(S)-flurocarazolol"	-8.65	7.5
"lisuride"	-7.70	9.7 - 9.8	"ketanserin"	-8.51	5
"FG-5893"	-7.69	8.7	"rec 15/3079"	-8.50	9.7
"rizatriptan"	-7.67	6.4	"(R)-flurocarazolol"	-8.46	6.5
"L-772,405"	-7.66	7.2*	"WAY-100635"	-8.39	7.9 – 9.2
"5-CT"	-7.63	9.4 - 10.3	"[3H]p-MPPF"	-8.27	8.4**
"LY334370"	-7.55	7.8	"pipamperone"	-8.25	5.6
"donitriptan"	-7.48	7.6	"(-)-propranolol"	-8.14	7.5
"eletriptan"	-7.34	7.4	"pindolol"	-8.08	8.1
Mean score	**-7.72**		**Mean score**	**-8.44**	

*- these affinity values have pIC_{50} unit.
**- these affinity values have p*K*d unit.

Table 7. Showing the 10 compounds of the known binders with best Glide score for the 5-HT_{2A} receptor models. Mean score of the 10 compounds was set to be a threshold value for scoring of the toxicants. Affinity values imported from www.guidetopharmacology.org (40).

5-HT_{2A} agonist bound state (model 2)			**5-HT_{2A} antagonist bound state (model 27)**		
Known binders	*Glide score (kcal/mol)*	*Affinity (pKi)*	*Known binders*	*Glide score (kcal/mol)*	*Affinity (pKi)*
"donitriptan"	-9.79	6.7	"methiothepin"	-7.86	8.5
"5-CT"	-8.86	6.5	"cyamemazine"	-7.55	8.8
"AL-37350A"	-8.81	8.7	"fluspirilene"	-7.48	8.0
"Ro 60-0175"	-8.74	7.4	"pimozide"	-7.48	7.1 - 7.7
"tryptamine"	-8.72	7.1	"mianserin"	-7.33	7.7 - 9.6
"BW723C86"	-8.60	7.2	"metergoline"	-7.20	8.6
"lisuride"	-8.52	8.6	"risperidone"	-6.90	9.3 - 10.0
"terguride"	-8.51	8.3	"clozapine"	-6.88	7.6 - 9.0
"5-MeOT"	-8.46	8.9	"asenapine"	-6.86	10.2
"5-hydroxy-tryptamine"	-8.23	8.9	"pipamperone"	-6.77	8.3
Mean score	**-8.72**		**Mean score**	**-7.23**	

Table 8. Showing the 10 compounds of the known binders with best Glide score for the 5-HT_{2C} receptor models. Mean score of the 10 compounds was set to be a threshold value for scoring of the toxicants. Some ligands appear more than one time due to the R and S configuration states. Affinity values imported from www.guidetopharmacology.org (40).

5-HT_{2C} agonist bound state (model 8)			**5-HT_{2C} antagonist bound state (model 11)**		
Known binders	*Glide score (kcal/mol)*	*Affinity (pKi)*	*Known binders*	*Glide score (kcal/mol)*	*Affinity (pKi)*
"5-CT"	-7.87	5.2 - 6.7	"trifluoperazine"	-9.04	6.4
"aripiprazole"	-7.83	7.6	"trifluoperazine"	-8.68	6.4
"YM348"	-7.60	9.0	"zotepine"	-8.64	8.6
"YM348"	-7.58	9.0	"chlorpromazine"	-8.38	7.6 - 8.2
"Ro 60-0175"	-7.45	7.7 - 8.2	"amitriptyline"	-8.37	8.1
"S 16924"	-7.37	7.7	"sertindole"	-8.11	9.0 - 9.2
"RU 24969"	-7.23	6.8	"mesulergine"	-8.10	8.7 - 9.3
"BRL-15572"	-7.21	6.2	"volinanserin"	-7.99	7.5 - 7.7
"VER-3323"	-7.15	8.2	"clozapine"	-7.83	7.4 - 8.7
"TFMPP"	-7.10	6.5 - 7.8	"perphenazine"	-7.83	6.9
Mean score	**-7.44**		**Mean score**	**-8.30**	

4.2.4 Docking of known binders into crystal structures of SERT

The know binders (inhibitors) of SERT were directly docked with Glide into all three protein prepared structures of SERT (PDB ID: 5I6X; 5I71; 5I73). No constraint option was added for docking, such that ligands could freely orient in the binding pocket to obtain the best free energy of binding.

The 10 best Glide score results of the known binders were chosen to predict how the Glide scoring correlated to the experimentally determined affinity values (Tables 9, 10 and 11).

Table 9. Showing the 10 compounds of the known binders with best Glide score for the orthosteric binding site in SERT (PDB ID: 5I6X). Mean score of the 10 compounds was set to be a threshold value for scoring of the toxicants. Some ligands appear more than one time due to the R and S configuration states. Affinity values imported from www.guidetopharmacology.org (40).

SERT orthosteric binding site (PDB ID: 5I6X)		
Known binders	*Glide score (kcal/mol)*	*Affinity (pKi)*
"[^{3}H]paroxetine"	-9.73	9.7**
"paroxetine"	-9.56	9.6
"fluoxetine"	-9.09	8.5
"fluoxetine"	-9.02	8.5
"[^{3}H]citalopram"	-9.00	8.3**
"citalopram"	-9.00	8.4
"[^{3}H]citalopram"	-8.87	8.3**
"escitalopram"	-8.87	9.0*
"citalopram"	-8.87	8.4
"vortioxetine"	-8.83	8.8
Mean score	**-9.08**	

*- these affinity values have pIC_{50} unit.
**- these affinity values have p*K*d unit.

Table 10. Showing the 10 compounds of the known binders with best Glide score for the orthosteric binding site in SERT (PDB ID: 5I71). Mean score of the 10 compounds was set to be a threshold value for scoring of the toxicants. Some ligands appear more than one time due to the R and S configuration states. Affinity values imported from www.guidetopharmacology.org (40).

SERT orthosteric binding site (PDB ID: 5I71)		
Known binders	*Glide score (kcal/mol)*	*Affinity (pKi)*
"vilazodone"	-9.60	8.8 - 9.3*
"fluvoxamine"	-9.23	8.7**
"fluvoxamine"	-9.23	8.7**
"lofepramine"	-9.15	7.2
"[^{3}H]citalopram"	-9.12	8.3**
"escitalopram"	-9.12	9.0*
"citalopram"	-9.12	8.4
"[^{3}H]citalopram"	-9.09	8.3**
"citalopram"	-9.09	8.4
"protriptyline"	-8.98	7.7**
Mean score	**-9.17**	

*- these affinity values have pIC_{50} unit.
**- these affinity values have p*K*d unit.

Table 11. Showing the 10 compounds of the known binders with best Glide score for the allosteric binding site in SERT (PDB ID: 5I73). Mean score of the 10 compounds was set to be a threshold value for scoring of the toxicants. Some ligands appear more than one time due to the R and S configuration states. Affinity values imported from www.guidetopharmacology.org (40).

SERT allosteric binding site (PDB ID: 5I73)		
Known binders	*Glide score (kcal/mol)*	*Affinity (pKi)*
"nefazodone"	-8.45	6.7**
"vilazodone"	-8.19	8.8 - 9.3*
"lofepramine"	-8.14	7.2
"desipramine"	-7.74	7.7
"desipramine"	-7.74	7.7
"lofepramine"	-7.73	7.2
"amoxapine"	-7.50	7.7
"ziprasidone"	-7.48	7.3
"nortriptyline"	-7.46	8.2
"dosulepin"	-7.38	8.1
Mean score	**-7.78**	

*- these affinity values have pIC_{50} unit.
**- these affinity values have p*K*d unit.

4.2.5 Screening scores for exogenous toxicants

The total of 9 757 ligand prepared toxicants were docked into constructed serotonergic receptor models and protein prepared structures of SERT. Out of 9 757 toxicants, 6 803 had a CNS MPO score ≥ 4.

For serotonergic receptor models the mean Glide scores of the 10 best Glide score results (Tables 6, 7 and 8) of the known binders were used as a threshold. Toxicants with a score better than the threshold value for known binders were considered as putative binders for the target. The number of toxicants that exceeded threshold mean score for each receptor model is shown in Table 12.

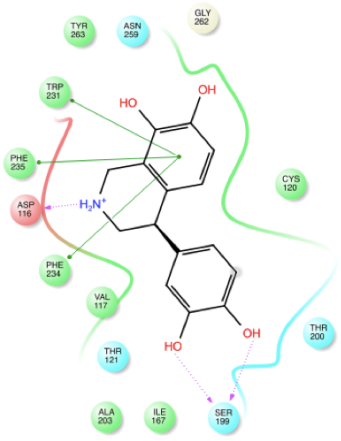
Table 12. Showing the number of toxicants that exceeded the threshold mean score.

Receptor model	Mean Glide score	Number of toxicants over threshold
5-HT_{1A} agonist bound	-7.72	33
5-HT_{1A} antagonist bound	-8.44	57
5-HT_{2A} agonist bound	-8.72	4
5-HT_{2A} antagonist bound	-7.23	186
5-HT_{2C} agonist bound	-7.44	74
5-HT_{2C} antagonist bound	-8.30	144

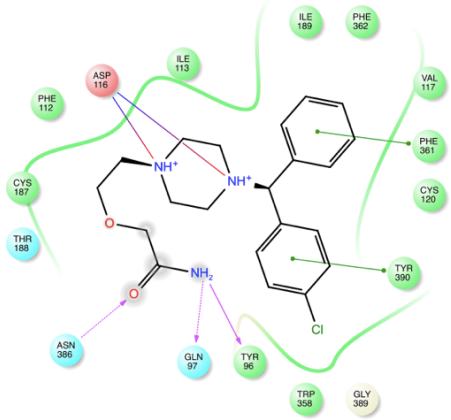
The 10 best Glide score results of the toxicants for each receptor model were also presented (Appendix, A1). From these tables, the toxicants with the highest CNS MPO score were selected for detailed inspection (Table 13 and Fig. 13).

Table 13. Showing toxicants that gained the best CNS MPO and Glide score for the receptor models.

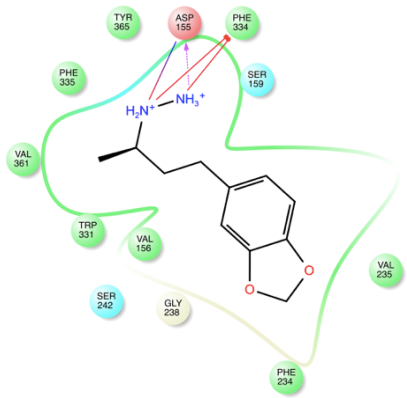
Toxicant	CNS MPO	Glide score	Target
zelandopam	4.16	-8.18	5-HT_{1A}
cetrizine amide	4.86	-9.15	5-HT_{1A}
safrazine	5.50	-8.98	5-HT_{2A}
eletriptan	4.13	-8.60	5-HT_{2A}
rac nebivolol	4.17	-8.05	5-HT_{2C}
trifluperidol	3.86	-9.74	5-HT_{2C}



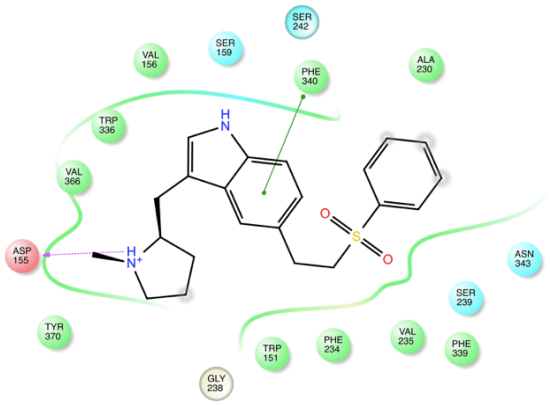
Toxicant: zelandopam
5-HT_{1A} agonist bound receptor: Model 7 from IFD



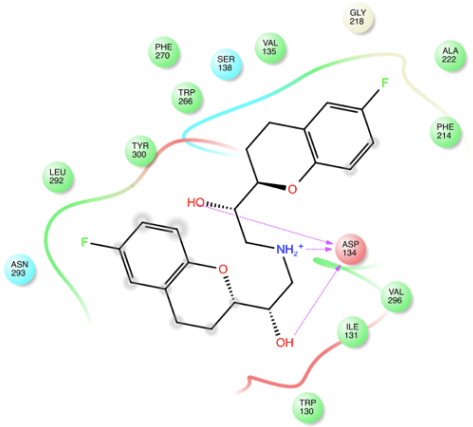
Toxicant: cetrizine amide
5-HT_{1A} antagonist bound receptor: Model 5 from IFD



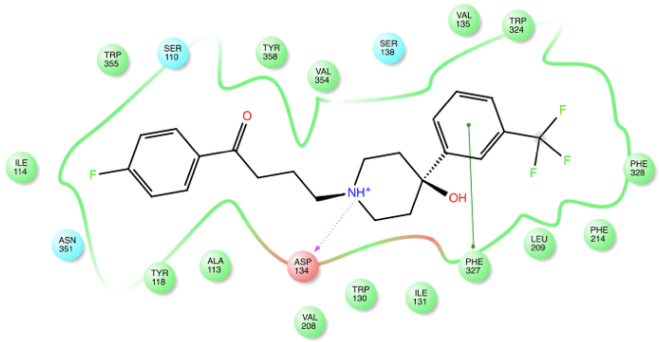
Toxicant: safrazine
5-HT_{2A} agonist bound receptor: Model 2 from IFD



Toxicant: eletriptan
5-HT_{2A} antagonist bound receptor: Model 27 from IFD



Toxicant: rac nebivolol
5-HT_{2C} agonist bound receptor: Model 8 from IFD



Toxicant: trifluperidol
5-HT_{2C} antagonist bound receptor: Model 11 from IFD

Fig. 13. The toxicant-receptor complexes of the models with the toxicants that gained the best CNS MPO and Glide score. The amino acid residues are in a 3Å sphere radius around the toxicants.

The protein prepared structures of SERT were also docked with known binders (inhibitors) and toxicants without use of the constraint option in Maestro.

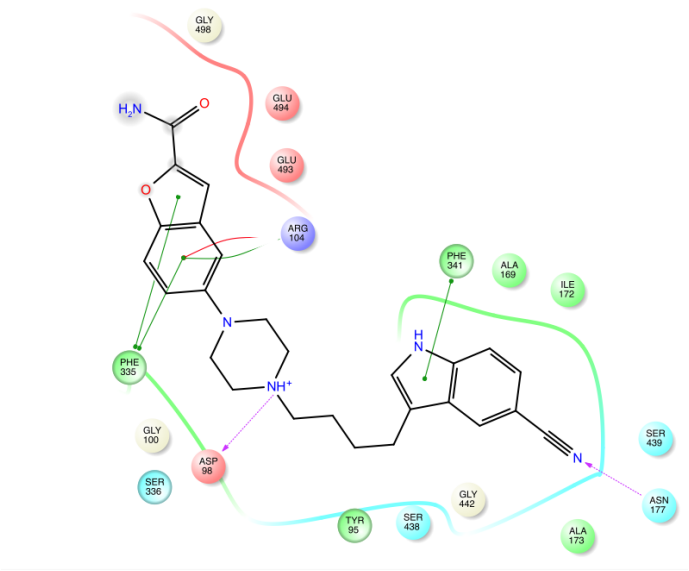
For SERT structures the mean Glide scores of the 10 best Glide score results (calculated in Tables 9, 10 and 11) of the known binders were used as a threshold. The number of toxicants that exceeded the threshold mean score for each SERT structure is shown in Table 14.

Table 14. Showing the number of toxicants that exceeded the threshold mean score.

X-ray crystal structure	Mean Glide score	Number of toxicants over threshold
SERT (PDB ID: 5I6X)	-9.08	114
SERT (PDB ID: 5I71)	-9.17	207
SERT (PDB ID: 5I73)	-7.78	370

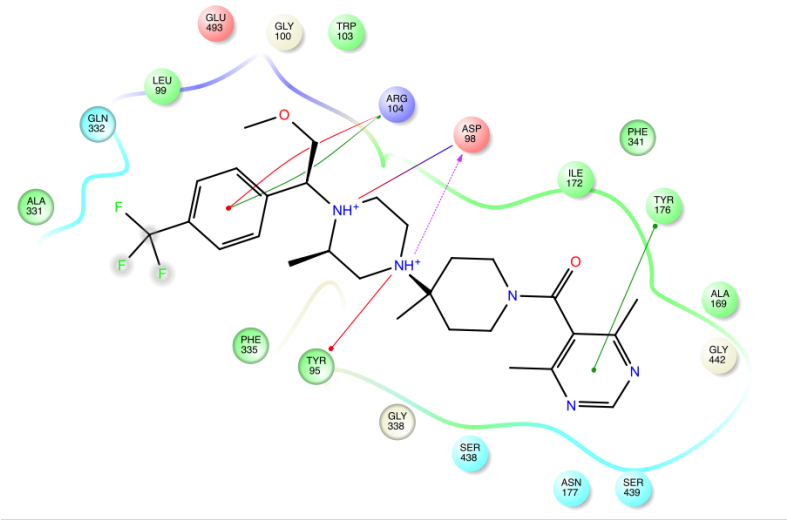
The 10 best Glide score results of the toxicants for each SERT structure were also selected (Appendix, A2) for further analysis. However, none of the 10 selected toxicants, except for bamifylline which was 9th on the Glide score list (Appendix, A2, Table b), had a CNS MPO score ≥ 4. The selected vicriviroc was 29th on the Glide score list and trelanserin was 54th on the Glide score list.

The SERT complexes with the toxicants that gained the best CNS MPO and Glide scores are presented in Figs. 14, 15 and 16.



Known binder: vilazodone

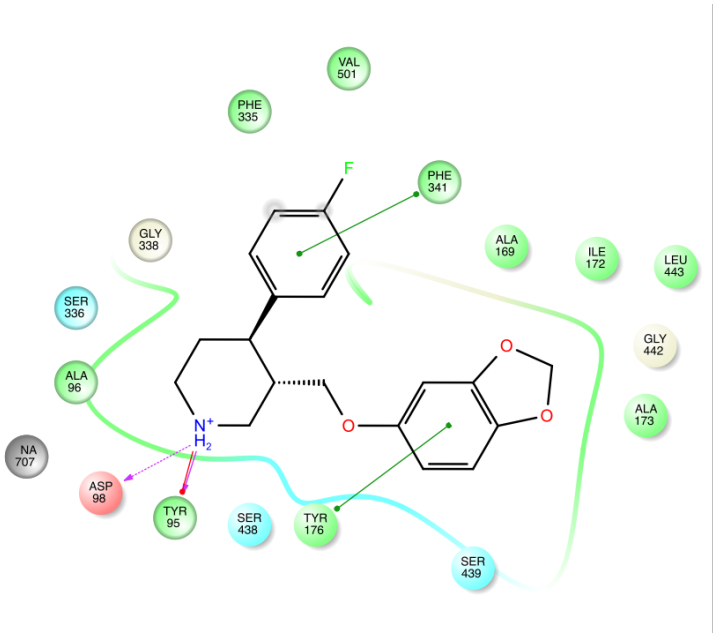
Glide score: -9.60 pIC_{50}: 8.8 – 9.3



Toxicant: vicriviroc

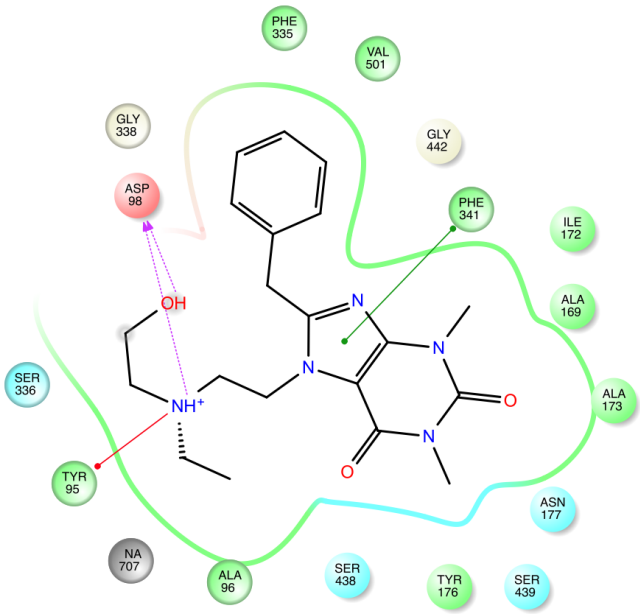
Glide score: -10.29 CNS MPO: 4.32

Fig. 14. SERT (S)-citalopram bound (central site) in complex with the known binder and the toxicant that gained the best CNS MPO and Glide score. The amino acid residues are in a 3Å sphere radius around the known binder and the toxicant. Affinity values imported from www.guidetopharmacology.org (40).



Known binder: [^{3}H] paroxetine

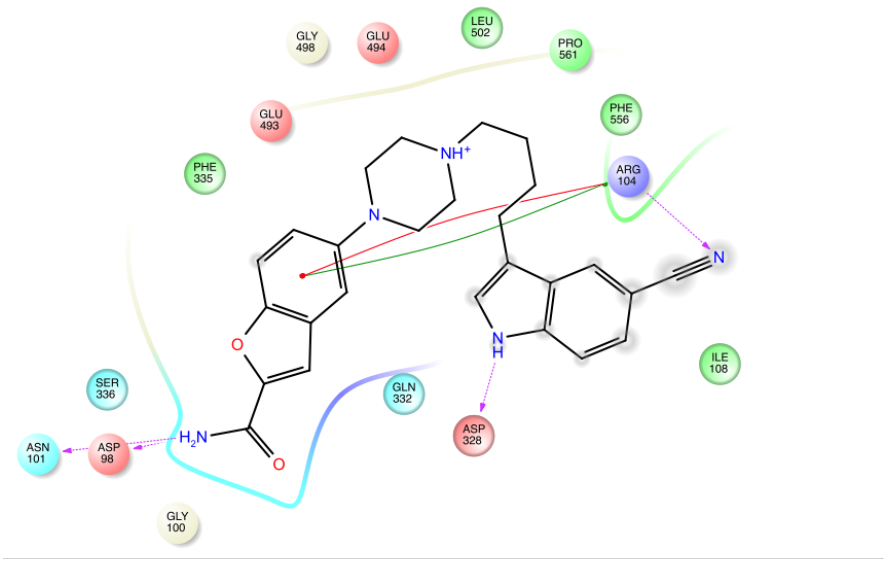
Glide score: -9.73 pK_d: 9.7



Toxicant: bamifylline

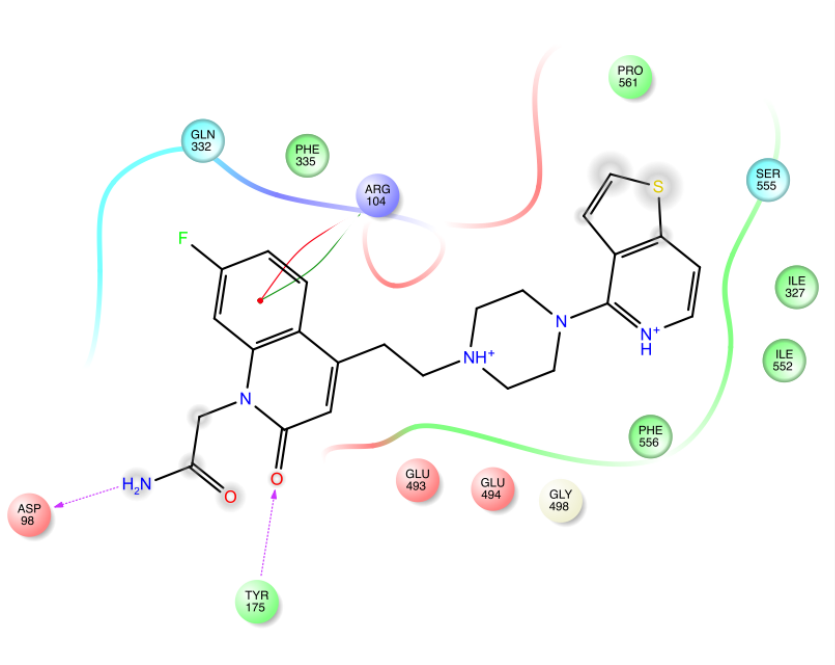
Glide score: -10.34 CNS MPO: 4.70

Fig. 15. SERT paroxetine bound (central site) in complex with known binder and toxicant that gained the best CNS MPO and Glide score. The amino acid residues are in a 3Å sphere radius around the known binder and the toxicant. Affinity values imported from www.guidetopharmacology.org (40).



Known binder: vilazodone

Glide score: -8.19 pIC_{50}: 8.8 – 9.3



Toxicant: trenlanserin

Glide score: -8.85 CNS MPO: 4.41

Fig. 16. SERT (S)-citalopram bound (allosteric site) in complex with known binder and toxicant that gained the best CNS MPO and Glide score. The amino acid residues are in a 3Å sphere radius around the known binder and the toxicant. Affinity values imported from www.guidetopharmacology.org (40).

5. Discussion

In the present project we have constructed theoretical homology models of the 5-HT_{1A}, 5-HT_{2A} and 5-HT_{2C} receptors and optimized the X-ray crystal structures of SERT for docking by molecular modelling techniques. The optimized models were used for docking of known binders for these important drug targets, and for predicting the ligand-receptor interactions of 8 164 exogenous toxicants from the Tox21 database.

As of today, there is no available crystal structures for 5-HT_{1A}, 5-HT_{2A} and 5-HT_{2C} receptors. The homology modelling approach was the method chosen for predicting the 3D structure of these receptors. The models were constructed by using known receptor structures as templates with similar biological and chemical properties imported from the PDB. High sequential similarity between the model targets and the available X-ray crystal structures of 5-HT_{1B}, 5-HT_{2B}, D_3 and β_2 (Table 3) has enabled a reliable homology modelling. However, a certain degree of template bias will always be present, as the templates are not completely identical to structure of the receptors constructed.
Sequence alignments of serotonergic and melatonin receptors were performed in order to show the conserved regions. High amino acid conservation between melatonin and serotonin receptors was observed in all TMHs (Fig. 6). Their structural similarity could indicate their similar interactions to environmental toxicants, which might lead to potential physiological consequences.

The sequences of target receptors were aligned with different crystal structures to accordingly yield an agonist and antagonist bound state for each receptor (Fig. 8). This was done due to the differences in the size of the binding pockets, where we assumed that an antagonist bound state have somewhat bigger binding site relative to the agonist bound state. This ensured that some molecules of bigger size could be docked into the receptor models.
The protein health score of serotonergic receptors (in the binding pockets) was < 6 after ICM refinements. The protein health value < 6 means that the strain energy of structures is acceptable. The models were further optimized by IFD, while the X-ray crystal structures of SERT were directly prepared for Glide docking, and IFD was not performed for the SERT structures. As the focus was on the binding pockets, the loop regions were neglected in the homology modelling process. However, loops are highly flexible parts of the receptors, and sometimes they can be important in the ligand-receptor interactions.

The docking and scoring calculations performed were based on Prime and Glide modules integrated in Maestro program. In the IFD protocol, a particular ligand (high affinity binder) was docked with Glide to the protein prepared receptor structure, with many poses generated. Prime was used to optimize the receptor structure with each particular ligand pose (the induced-fit part). The ligand was then redocked into the new receptor conformations, and the complexes were scored based both on the redocking Glide score and the Prime energy from the optimization.

The scoring functions Glide and Prime are considered to have very good performance at predicting if the ligand can bind or not, but they are not so reliable for prediction of the ligands free energy (ΔG).

Different IFD and Glide scores were calculated for serotonergic receptor models docked with high affinity binders. As the more negative scores indicate stronger interactions with the receptor, the best IFD score was -541.02 for the 5-HT_{2C} receptor docked with YM348 and the best Glide score was -9.83 for 5-HT_{2C} receptor docked with serindole (Table 5). IFD and Glide scores for docking of the other models also performed well. Seemingly, Glide scores for 5-HT_{1A} receptor docked with LY293284 (-6.35) and for 5-HT_{2A} receptor docked with asenapine (-7.09) were somewhat lower than expected, relative to their experimentally determined affinity values (10.1 and 10.2 p*K*i).

One of the approaches utilized for docking to serotonergic receptors, was to include the crucial constrained interaction between the carboxylate oxygen atom(s) in Asp3.32 in TM helix III and the protonated amine of ligands (Figs. 10, 11 and 12). It was important to include this constraint to obtain realistic poses in the docking calculations. Ligand-receptor complexes presented in these figures indicate that besides aspartic acid, there is a frequent aromatic interaction between the benzene ring in ligands and phenylalanine (Phe) from receptors side. Serine (Ser242 in TMH V) from 5-HT_{2A} receptor makes hydrogen bond interaction to an amine in the ligand methylergonovine, while tyrosine (Tyr118 in TMH III) in 5-HT_{2C} receptor makes hydrogen bond interaction to a carbonyl group in the ligand sertindole, and asparagine (Asn273 in TMH VI) from 5-HT_{2C} receptor makes hydrogen bond interaction to a cyclic ether in the ligand YM348 (Fig. 12).

To evaluate the constructed models of serotonergic receptors, the known binders and decoys were docked by the VSW. All the serotonergic receptor models performed good at differentiating between actives and decoys, with BEDROC values ranging from 0.498 (for antagonist bound state of 5-HT_{2A}) to 0.828 (for agonist bound state of 5-HT_{2A}). However, not all the actives could be docked into receptor models, and that could be seen in the case of

antagonist bound state of 5-HT_{2C} receptor (Fig. 9, down-left corner). This is probably a result of a too small binding pocket, since antagonist actives generally are molecules of bigger size than agonists. It is important to mention that BEDROC metric can not recognize if the size of the binding pockets modelled is real. If the binding pockets modelled are too big, the compounds which naturally wouldn't bind will be docked.

The BEDROC values, however, suggest that all the constructed models are predictive, and thus can be used for docking studies to examine the ligand-receptor interactions of the environmental toxicants.

The total of 9 757 ligand prepared toxicants (originally 8 164 before LigPrep that also generates enantiomers) were docked into serotonergic receptor models and the protein prepared structures of SERT. Out of 9 757 toxicants, 6 803 had a CNS MPO score ≥ 4. A CNS MPO score ≥ 4, indicates that a ligand has CNS drug-like properties, and might be able to reach the CNS. Accordingly, around 70 % of all the toxicants in the Tox21 database have such physicochemical properties and could possibly interfere with neurotransmission in the brain.

To be able to predict the interactions of environmental toxicants with the constructed serotonergic receptors and the SERT models, we had to firstly inspect their interactions with the known binders. It was observed that most of the 10 compounds with most favourable scoring also had strong experimentally determined affinity values (Tables 6, 7 and 8). The mean Glide score of known binders was calculated to establish a threshold value for the docking of the toxicants. The threshold value was simply utilized to get an insight on how many toxicants have the potential of making interactions with serotonergic receptors and the SERT. Toxicants with a score better than the threshold value for known binders were considered as putative binders for the target. They ranged from 4 toxicants over threshold Glide score -8.72 for 5-HT_{2A} agonist bound receptor (Table 12), to 370 toxicants over threshold Glide score -7.78 for SERT (PDB ID: 5I73) (Table 14).

The CNS MPO score was used in combination with the Glide score to select out toxicants with higher risk to reach and affect serotonergic receptors and the SERT in the CNS.

It could be observed that important interactions between toxicants and serotonergic receptors, similarly to the known binders, are in many cases aromatic interactions with the amino acid residue phenylalanine (Fig. 13). In model 7 from IFD of the 5-HT_{1A} agonist bound receptor, it was observed that both the known binder (LY293284) and the toxicant (zelandopam) make

interactions to Asp116 (TMH III) and Phe235 (TMH VI) (Figs. 10 and 13). In model 5 of 5-HT_{1A} antagonist bound receptor, both the known binder (rec 15/3079) and the toxicant (cetrizine amide) make similar interactions to Asp116 (TMH III) and Phe361 (TMH VI). In model 2 of 5-HT_{2A} agonist bound receptor, the known binder (methylergonovine) makes interactions to Asp155 (TMH III), Ser242 (TMH V) and Phe335 (TMH VI), while the only common interaction of the toxicant (rec nebivolol) is to Asp155, and the other dissimilar interaction (ionic) is to Phe334 (TMH VI) (Figs. 11 and 13). In model 27 of 5-HT_{2A} antagonist bound receptor, both the known binder (asenapine) and the toxicant (trifluperidol) bind to Asp155 (TMH III) and Phe340 (TMH VI). Asenapine binds in addition to Phe339 (TMH VI). In model 8 of 5-HT_{2C} agonist bound receptor, it was observed that the known binder (YM348) binds to Asp134 (TMH III), Phe270 (TMH VI) and Asn273 (TMH VI). However, only the constrained Asp134 interaction is present between the toxicant (safrazine) and the receptor (Figs. 12 and 13). In model 11 of 5-HT_{2C} antagonist bound receptor, the known binder (sertindole) makes interactions to Tyr118 (TMH II), Asp134 (TMH III), Phe327 (TMH VI) and Phe328 (TMH VI). Of these, the toxicant (eletriptan) interacts only with Asp134 and Phe327.

For the protein prepared structures of SERT, Glide scores of known binders were generally better than those of serotonergic models (Tables 9 and 10), except for the allosteric structure of SERT (Table 11). The correlation between the Glide scores and the experimental affinity values was also very good.

The Glide scores of toxicants for SERT structures (Appendix, A2) were better relative to the Glide scores of toxicants for serotonin receptor models (Appendix, A1).

The SERT receptor structures were docked without the constraint option, such that ligands could freely orient in the binding pocket to obtain the best free energy of binding. The interactions between known binders and SERT and toxicants and SERT are presented in Figs. 14, 15 and 16. The important hydrogen bond interaction between the protonated aliphatic amine in ligands and the carboxylate oxygen atom of Asp98 (TMH I) in transporter is present in all the complexes. Similarly, as for the serotonergic receptors, aromatic interactions with phenylalanine is also present in some ligands. Ligand interactions with tyrosine (Tyr) and arginine (Arg) were also observed to be present in some of the complexes.

In the SERT structure co-crystallized with (S)-citalopram at the central site, the known binder (vilazodone) makes interactions with Asp98 (TMH I), Phe335 (TMH VI), Phe341 (TMH VI), Arg104 (TMH I) and Asn177 (TMH III). Of these, the toxicant (vicriviroc) interacts with

Asp98 (TMH I) and Arg104 (TMH I) (Fig. 14). In the SERT structure co-crystallized with paroxetine at the central site, the known binder ([^{3}H] paroxetine) interacts with Asp98 (TMH I), Phe341 (TMH VI) and Tyr176 (TMH III), while the toxicant (bamifylline) makes similar interactions to Asp98 (TMH I) and Phe341 (TMH VI). In the SERT structure co-crystallized with (S)-citalopram at the allosteric site, the known binder (vilazodone) interacts with both Asp98 (TMH I) (important for central site) and Asp328 (TMH VI) (important for allosteric site), with this stretching out to both the orthosteric and allosteric sites, and Arg104 (TMH I). Of these, the toxicant (trenlanserin) was able to interact with Asp98 (TMH I) and Arg104 (TMH I). However, the toxicant failed to get docked at the allosteric site and make the crucial interaction to Asp328 (TMH VI).

In the work presented, many toxic compounds from the Tox21 database were predicted to interact with serotonergic receptors and the SERT. Potential interaction of environmental toxicants could affect the actions of neurotransmitters, drugs, hormones and inflammatorial mediators. Many environmental toxicants had CNS MPO ≥ 4 and are likely to cross the BBB and reach the CNS.

The serotonergic system is a very important drug targeting field, and it is of great interest to understand the structural and functional properties of its receptors and transporters.

Cumulative evidences suggest that 5-HT_{1A}, 5-HT_{2A} and 5-HT_{2C} receptors and SERT have a role in pathology of depression and they might be a key for future development of more efficacious and faster acting drugs. Many of the known binders tested in this study are approved antidepressants and their interactions and physicochemical properties were important for analysis of the environmental toxicant's ability to interfere with the CNS neurotransmission of serotonin.

6. Conclusion

The homology modelling approach has its weaknesses and errors, and the generated models may have uncertainties influenced by the profound impacts of the utilized templates. The constructed models of serotonergic receptors were, however, able to differentiate between actives and decoys and the BEDROC scores proved the models to be predictive.
Detailed interaction analysis of the selected compounds of serotonergic receptors and the SERT indicate that besides the crucial interaction with the conserved aspartic acid, aromatic interactions with phenylalanine are also very important. The obtained high CNS MPO scores and similar Glide scores between the known high affinity binders and toxicants could suggest harmful effects and drug interactions in serotonergic system of the CNS.
Future studies should include *in vitro* tests of the high ranking environmental toxicants for these receptors and transporter and the work presented may serve as basis for that.

References

1. Harrison EP, Grason H. Environmental Toxicants and Maternal and Child Health: An Emerging Public Health Challenge. Johns Hopkins Bloom Sch Public Health. 2009;

2. Environmental Neurotoxicology [Internet]. Washington, D.C.: National Academies Press; 1992 [cited 2016 Apr 29]. Available from: http://www.nap.edu/catalog/1801

3. Harris JB, Blain PG. Neurotoxicology: what the neurologist needs to know. J Neurol Neurosurg Psychiatry. 2004 Sep 1;75(suppl 3):iii29–34.

4. E. K. Silbergeld, "Toxicology," in ILO Encyclopedia of Occupational Health & Safety, chapter 33, part 4, International Labour Office, Washington, DC, USA, 4th edition, 1998.

5. About Tox21 [Internet]. National Center for Advancing Translational Sciences. 2015 [cited 2016 Apr 30]. Available from: https://ncats.nih.gov/tox21/about

6. Grandjean P, Landrigan PJ. Developmental neurotoxicity of industrial chemicals. Lancet Lond Engl. 2006 Dec 16;368(9553):2167–78.

7. Pamphlett R. Uptake of environmental toxicants by the locus ceruleus: A potential trigger for neurodegenerative, demyelinating and psychiatric disorders. Med Hypotheses. 2014 Jan;82(1):97–104.

8. Lee S-M, Booe JM, Pioszak AA. Structural insights into ligand recognition and selectivity for classes A, B, and C GPCRs. Eur J Pharmacol. 2015 Sep 15;763, Part B:196–205.

9. Katritch V, Cherezov V, Stevens RC. Structure-function of the G protein-coupled receptor superfamily. Annu Rev Pharmacol Toxicol. 2013;53:531–56.

10. Kobilka BK, Deupi X. Conformational complexity of G-protein-coupled receptors. Trends Pharmacol Sci. 2007 Aug;28(8):397–406.

11. Strasser A, Wittmann H-J. Modelling of GPCRs [Internet]. Dordrecht: Springer Netherlands; 2013 [cited 2016 Apr 30]. Available from: http://link.springer.com/10.1007/978-94-007-4596-4

12. Rang HP, Dale MM. Rang and Dale's pharmacology. Edinburgh; New York: Elsevier/Churchill Livingstone; 2012.

13. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, et al. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science. 2007 Nov 23;318(5854):1258–65.

14. 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. 2016 May;34(5):952–70.

15. Baou M, Boumba VA, Petrikis P, Rallis G, Vougiouklakis T, Mavreas V. A review of genetic alterations in the serotonin pathway and their correlation with psychotic diseases and response to atypical antipsychotics. Schizophr Res. 2016 Jan;170(1):18–29.

16. Gordon N, Goelman G. Understanding alterations in serotonin connectivity in a rat model of depression within the monoamine-deficiency and the hippocampal-neurogenesis frameworks. Behav Brain Res. 2016 Jan 1;296:141–8.

17. Gabrielsen M. Structure, function and the inhibition of the serotonin transporter studied by molecular docking, -dynamics and virtual screening. [Internet]. 2011. Available from: http://www.ub.uit.no/munin/bitstream/10037/6403/7/paper_1.pdf.

18. Wang C, Jiang Y, Ma J, Wu H, Wacker D, Katritch V, et al. Structural basis for molecular recognition at serotonin receptors. Science. 2013 May 3;340(6132):610–4.

19. Artigas F. Serotonin receptors involved in antidepressant effects. Pharmacol Ther. 2013 Jan;137(1):119–31.

20. Gabrielsen M, Kurczab R, Siwek A, Wolak M, Ravna AW, Kristiansen K, et al. Identification of novel serotonin transporter compounds by virtual screening. J Chem Inf Model. 2014 Mar 24;54(3):933–43.

21. Chou Y-H, Hsieh W-C, Chen L-C, Lirng J-F, Wang S-J. Association between the serotonin transporter and cytokines: Implications for the pathophysiology of bipolar disorder. J Affect Disord. 2016 Feb;191:29–35.

22. Cannon DM, Ichise M, Fromm SJ, Nugent AC, Rollis D, Gandhi SK, et al. Serotonin transporter binding in bipolar disorder assessed using [11C]DASB and positron emission tomography. Biol Psychiatry. 2006 Aug 1;60(3):207–17.

23. Coleman JA, Green EM, Gouaux E. X-ray structures and mechanism of the human serotonin transporter. Nature. 2016 Apr 21;532(7599):334–9.

24. Simon, G. Ciechanowski, P. (2015, Dec 10). Depression treatment options for adults [Internet]. Available from http://www.uptodate.com.

25. Røsjø B. (2014, May 3). Anxiety and depression are the most expansive (Angst og depresjon koster mest) [Internet]. Retrieved from http://forskning.no.

26. Schildkraut JJ. The catecholamine hypothesis of affective disorders: a review of supporting evidence. Am J Psychiatry. 1965 Nov;122(5):509–22.

27. Voksne med depresjon (Depression in adults) [Internet]. Nasjonal retningslinje for diagnostisering og behandling av voksne med depresjon i primær- og spesialisthelsetjenesten. 2015 [cited 2016 Apr 30]. Available from: https://helsedirektoratet.no.

28. Sliwoski G, Kothiwale S, Meiler J, Lowe EW. Computational Methods in Drug Discovery. Pharmacol Rev. 2014 Jan 1;66(1):334–95.

29. PhD TLL, PhD DAW, editors. Foye's Principles of Medicinal Chemistry. Sixth edition. Philadelphia: LWW; 2007. 1377 p.

30. Höltje H-D, Sippl W, Rognan D, Folkers G. Molecular Modeling. 3 edition. Weinheim: Wiley-VCH; 2008. 320 p.

31. Orry AJW, Abagyan R, editors. Homology Modeling: Methods and Protocols. 2012 edition. New York: Humana Press; 2012. 419 p.

32. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov. 2004 Nov;3(11):935–49.

33. Leach AR, Shoichet BK, Peishoff CE. Prediction of protein-ligand interactions. Docking and scoring: successes and gaps. J Med Chem. 2006 Oct 5;49(20):5851–5.

34. Huang S-Y, Grinter SZ, Zou X. Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions. Phys Chem Chem Phys PCCP. 2010 Oct 28;12(40):12899–908.

35. 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. 1994 May 1;15(5):488–506.

36. Schrödinger Release 2015-3: Maestro, version 10.3, Schrödinger, LLC, New York, NY, 2015.

37. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res. 2000 Jan 1;28(1):235–42.

38. Consortium TU. UniProt: a hub for protein information. Nucleic Acids Res. 2015 Jan 28;43(D1):D204–12.

39. Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, et al. ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res. 2011 Sep 23;gkr777.

40. Southan C, Sharman JL, Benson HE, Faccenda E, Pawson AJ, Alexander SPH, et al. The IUPHAR/BPS Guide to PHARMACOLOGY in 2016: towards curated quantitative interactions between 1300 protein targets and 6000 ligands. Nucleic Acids Res. 2016 Jan 4;44(D1):D1054–68.

41. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. J Med Chem. 2012 Jul 26;55(14):6582–94.

42. Truchon J-F, Bayly CI. Evaluating virtual screening methods: good and bad metrics for the “early recognition” problem. J Chem Inf Model. 2007 Apr;47(2):488–508.

43. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem. 2004 Mar 25;47(7):1739–49.

44. Wager TT, Hou X, Verhoest PR, Villalobos A. Moving beyond rules: the development of a central nervous system multiparameter optimization (CNS MPO) approach to enable alignment of druglike properties. ACS Chem Neurosci. 2010 Jun 16;1(6):435–49.

Appendix

A1. Glide scores of environmental toxicants in serotonergic receptors: a), b), c), d), e) and f).

a)

5-HT$_{1A}$ agonist bound state (model 7)	
Toxicants	*Glide score (kcal/mol)*
"Efonidipine"	-8.94
"Lapatinib"	-8.45
"3-(1-{3-[(3S)-1-benzoyl-3-(3,4-dichlorophenyl)piperidin-3-yl]propyl}-4-phenylpiperidin-4-yl)-1,1-dimethylurea"	-8.38
"Hesperidin"	-8.37
"Oxatomide"	-8.31
"3-(1-{3-[(3S)-1-benzoyl-3-(3,4-dichlorophenyl)piperidin-3-yl]propyl}-4-phenylpiperidin-4-yl)-1,1-dimethylurea"	-8.27
"3-(1-{2-[(2R)-4-benzoyl-2-(3,4-difluorophenyl)morpholin-2-yl]ethyl}-4-phenylpiperidin-4-yl)-1,1-dimethylurea hydrochloride (1:1)"	-8.20
"Carvedilol tartrate"	-8.18
"Zelandopam"	-8.18
"cis-Flupentixol"	-8.16

b)

5-HT$_{1A}$ antagonist bound state (model 5)	
Toxicants	*Glide score (kcal/mol)*
"Indinavir sulfate"	-9.68
"N-[1-{2-[(2R)-2-(3,4-dichlorophenyl)-5-oxo-4-phenylmorpholin-2-yl]ethyl}-4-(3-fluorophenyl)piperidin-4-yl]acetamide butanedioate"	-9.64
"Nelfinavir mesylate"	-9.44
"Nelfinavir mesylate"	-9.44
"Oxatomide"	-9.41
"Manidipine dihydrochloride"	-9.27
"Flunarizine"	-9.18
"Cetirizine amide dihydrochloride"	-9.15
"Hydroxyzine hydrochloride"	-9.14
"Hydroxyzine"	-9.14

c)

5-HT$_{2A}$ agonist bound state (model 2)	
Toxicants	*Glide score (kcal/mol)*
"Spiperone"	-9.55
"Safrazine hydrochloride"	-8.98
"Carvedilol tartrate"	-8.83
"Mabuterol hydrochloride"	-8.73
"Bucindolol"	-8.71
"Methergine"	-8.63
"Methergine"	-8.63
"Amoxicillin"	-8.62
"Cefadroxil"	-8.62
"3-({(3R,4R)-6-[(5-fluoro-1,3-benzothiazol-2-yl)methoxy]-4-hydroxy-3,4-dihydro-2H-chromen-3-yl}methyl)benzoic acid"	-8.56

d)

5-HT_{2A} antagonist bound state (model 27)	
Toxicants	*Glide score (kcal/mol)*
"Formoterol hemifumarate"	-8.81
"2-[3,5-Bis(trifluoromethyl)phenyl]-N-{4-(4-fluoro-2-methylphenyl)-6-[(2R,3R)-3-hydroxy-2-(hydroxymethyl)pyrrolidin-1-yl]pyridin-3-yl}-N,2-dimethylpropanamide"	-8.64
"Eletriptan"	-8.60
"Eletriptan"	-8.60
"Oxatomide"	-8.44
"Bisphenol AF"	-8.41
"Ractopamine hydrochloride"	-8.41
"Dibekacin"	-8.38
"Ractopamine hydrochloride"	-8.34
"Bisphenol Z"	-8.23

e)

5-HT_{2C} agonist bound state (model 8)	
Toxicants	*Glide score (kcal/mol)*
"Xaliproden hydrochloride"	-8.40
"Xaliproden hydrochloride"	-8.40
"Naftopidil"	-8.36
"Cinacalcet hydrochloride"	-8.16
"Trifluperidol hydrochloride"	-8.13
"Fluprostenol"	-8.11
"Naftopidil"	-8.11
"6-Hydroxy-2-naphthyl disulfide"	-8.10
"Oxatomide"	-8.07
"rac Nebivolol hydrochloride"	-8.05

f)

5-HT_{2C} antagonist bound state (model 11)	
Toxicants	*Glide score (kcal/mol)*
"GBR 12909 dihydrochloride"	-10.38
"GBR 12909 dihydrochloride"	-10.38
"GBR 12909 dihydrochloride"	-10.32
"GBR 12909 dihydrochloride"	-10.32
"Sertindole"	-10.15
"Sertindole"	-10.15
"Salmeterol xinafoate"	-9.82
"Salmeterol"	-9.82
"Zuclopenthixol dihydrochloride"	-9.75
"Trifluperidol hydrochloride"	-9.74

A2. Glide scores of environmental toxicants in SERT structures: a), b) and c)

a)

SERT orthosteric binding site (PDB ID: 5I6X)	
Toxicants	*Glide score (kcal/mol)*
"Fluspirilene"	-11.07
"Lymecycline"	-11.03
"3-chloro-2-[(3R)-5-chloro-1-(2,4-dimethoxybenzyl)-3-methyl-2-oxo-2,3-dihydro-1H-indol-3-yl]-N-ethyl-N-(pyridin-3-ylmethyl)benzamide hydrochloride"	-10.87
"3-(1-{3-[(3S)-1-benzoyl-3-(3,4-dichlorophenyl)piperidin-3-yl]propyl}-4-phenylpiperidin-4-yl)-1,1-dimethylurea"	-10.62
"Indinavir sulfate"	-10.54
"Tipranavir"	-10.38
"Indinavir sulfate"	-10.37
"3-(1-{2-[(2R)-4-benzoyl-2-(3,4-difluorophenyl)morpholin-2-yl]ethyl}-4-phenylpiperidin-4-yl)-1,1-dimethylurea hydrochloride (1:1)"	-10.37
"Bamifylline Hydrochloride"	-10.34
"Bamifylline Hydrochloride"	-10.34

b)

SERT orthosteric binding site (PDB ID: 5I71)	
Toxicants	*Glide score (kcal/mol)*
"3-(1-{3-[(3S)-1-benzoyl-3-(3,4-dichlorophenyl)piperidin-3-yl]propyl}-4-phenylpiperidin-4-yl)-1,1-dimethylurea"	-12.56
"3-(1-{3-[(3S)-1-benzoyl-3-(3,4-dichlorophenyl)piperidin-3-yl]propyl}-4-phenylpiperidin-4-yl)-1,1-dimethylurea"	-11.92
"Indinavir sulfate"	-11.37
"3-(1-{3-[(3S)-1-benzoyl-3-(3,4-dichlorophenyl)piperidin-3-yl]propyl}-4-phenylpiperidin-4-yl)-1,1-dimethylurea"	-11.22
"Indinavir sulfate"	-11.07
"Talampicillin hydrochloride"	-11.04
"N-[1-{2-[(2R)-2-(3,4-dichlorophenyl)-5-oxo-4-phenylmorpholin-2-yl]ethyl}-4-(3-fluorophenyl)piperidin-4-yl]acetamide butanedioate"	-11.01
"3-(1-{2-[(2R)-4-benzoyl-2-(3,4-difluorophenyl)morpholin-2-yl]ethyl}-4-phenylpiperidin-4-yl)-1,1-dimethylurea hydrochloride (1:1)"	-10.99
"Manidipine dihydrochloride"	-10.91
"Indinavir sulfate"	-10.91

c)

SERT allosteric binding site (PDB ID: 5I73)	
Toxicants	*Glide score (kcal/mol)*
"4-chloro-2-fluoro-5-{[4-(3-fluorophenyl)-4-{2-[3-(2-methyl-1H-benzimidazol-1-yl)-8-azabicyclo[3.2.1]oct-8-yl]ethyl}piperidin-1-yl]carbonyl}-N-methylbenzenesulfonamide"	-10.70
"Carminomycin"	-10.40
"Idarubicin hydrochloride"	-10.04
"Daunorubicin"	-9.90
"4-chloro-2-fluoro-5-{[4-(3-fluorophenyl)-4-{2-[3-(2-methyl-1H-benzimidazol-1-yl)-8-azabicyclo[3.2.1]oct-8-yl]ethyl}piperidin-1-yl]carbonyl}-N-methylbenzenesulfonamide"	-9.81
"Daunomycin hydrochloride"	-9.72
"Bimosiamose"	-9.69
"Ketoconazole"	-9.55
"Nelfinavir mesylate"	-9.51
"Nelfinavir mesylate"	-9.51