



universität
wien

DIPLOMARBEIT

Titel der Diplomarbeit

Flexible Docking Studies of a Series of Ligands in a Homology Model of the
Human Serotonin Transporter

angestrebter akademischer Grad

Magister der Pharmazie (Mag.pharm.)

Verfasserin / Verfasser: Andreas Jurik
Matrikel-Nummer: 0308019
Studienrichtung /Studienzweig Pharmazie
(lt. Studienblatt):
Betreuerin / Betreuer: Univ.-Prof. Mag. pharm. Dr. Gerhard Ecker

Wien, im Mai 2010

Gewidmet meinem Großvater, MR Dr. Hans Cihak

Ich möchte an dieser Stelle dem Betreuer der Diplomarbeit, Herrn Univ. Prof. Dr. Gerhard Ecker, meinen besonderen Dank aussprechen. Ohne seinem trotz vielfältigster Verpflichtungen stets nach Möglichkeit offenes Ohr, dem wertvollen und konstruktiv kritischen Feedback und nicht zuletzt der Möglichkeit, in seiner Gruppe tätig zu sein, wäre das Verfassen der vorliegenden Arbeit nicht möglich gewesen.

Ausdrücklicher Dank gilt auch meinem geschätzten Kollegen, Mag. pharm. René Weissensteiner, der mir vom ersten Tag an mit Rat und Tat zur Seite gestanden ist, und dessen Meinung in fachlichen Dingen für mich stets von Relevanz ist.

Auch der ganzen Arbeitsgruppe will ich für das gute und motivierende Klima danken, sowie für die Selbstverständlichkeit, mit der Probleme anderer Gruppenmitglieder immer Gehör finden.

Ich danke meinen Geschwistern Florian, Peter und Margaretha für ihre beständige Unterstützung in allen Lebenslagen.

Mein spezieller Dank gilt meinen Eltern Dr. Brigitte Jurik-Cihak und Dr. Karl Jurik, die mir nicht nur die Ausbildung ermöglichten, sondern auch die Freude daran vermittelt haben.

Abschließend will ich meiner ganzen Familie und allen Freunden sowie Kollegen danken, die mich während meines Studiums begleitet und unterstützt haben.

Danke

Table of Contents

1. Introduction and Aim of the Study	- 1 -
2. Biology	- 2 -
2.1 Nerve Cells.....	- 2 -
2.2 Neurotransmission	- 2 -
2.2.1 Monoamine Transporters	- 4 -
2.2.2 SERT Structure and Function	- 6 -
2.2.3 Neurotransmitters	- 8 -
3. Pharmacology	- 9 -
3.1 Serotonin	- 10 -
3.2 MDMA	- 11 -
3.3 Methylphenidate.....	- 13 -
3.4 Cocaine	- 14 -
3.5 Benztropine	- 14 -
3.6 CFT	- 15 -
3.7 CIT.....	- 15 -
4. Computational Background.....	- 17 -
4.1 Molecular Modeling	- 17 -
4.2 Homology Modeling.....	- 19 -
4.3 Docking.....	- 21 -
4.4 Scoring.....	- 22 -
5. Materials and Methods	- 25 -
5.1 The Model	- 25 -
5.1.1 The Template	- 25 -
5.1.2 Sequence Alignment	- 26 -
5.2 Schrödinger Suite	- 27 -
5.2.1 Ligand Preparation.....	- 28 -
5.2.2 Prime Structure Prediction	- 28 -

5.2.3	PrimeGlide Induced Fit Docking.....	- 31 -
5.3	Data Processing	- 33 -
5.3.1	IFDScore	- 33 -
5.3.2	Clustering	- 33 -
5.3.3	MM-GBSA Rescoring.....	- 35 -
5.4	Validation of the Method.....	- 35 -
6.	Results.....	- 37 -
6.1	Serotonin.....	- 39 -
6.2	MDMA	- 42 -
6.3	Methylphenidate.....	- 46 -
6.4	Benzatropine	- 48 -
6.5	Cocaine.....	- 51 -
6.6	β -CFT.....	- 53 -
6.7	β -CIT	- 55 -
7.	Discussion and Outlook.....	- 58 -
	References	- 60 -
	List of Abbreviations	- 66 -
	Table of Figures.....	- 68 -
	Abstract.....	- 70 -
	Zusammenfassung	- 71 -
	Curriculum Vitae	- 72 -

1. Introduction and Aim of the Study

One of the most complex mechanisms in the human body is the regulation of signal transduction between nerve cells. As a workhorse in the termination of neurotransmission, the serotonin transporter (SERT), a member of the neurotransmitter sodium symporter (NSS) or solute carrier 6 family (SLC 6), responsible for the reuptake of serotonin out of the synaptic cleft back in the nerve cell. Though being quite selective for its natural ligand, the SERT is targeted by many antidepressants and illicit drugs in order to modify the signal and therefore generally achieve a mood lifting effect, but little is known about the molecular basics of the interaction both of the natural ligand and other substrates with the receptor protein. Especially challenging in studies of membrane-bound proteins is the fact that due to technical difficulties in the crystallizing process, X-ray structures and other high-resolution structural data are mostly unavailable, so potent computational methods such as homology modeling and docking are needed to determine molecular binding modes.

In case of the serotonin transporter, the closest relative protein offering a crystal structure is a bacterial leucine transporter found in *Aquifex aeolicus*, very similar in structure and function. Taking it as a template, a homology model of its eukaryotic counterpart can be created. (Beuming et al. 2006)

The aims of the study include the creation of a hSERT homology model on the basis of the LeuT, then to use it for Induced Fit docking with several ligands and the further evaluation of binding modes, if available, in comparison with already published data, to determine the predictive potency of the method. The results yielded by the experiments were expected to improve the knowledge about already established ligand binding modes, but also they should allow sufficiently reliable suggestions for yet unknown protein-ligand interaction patterns, being able to explore poses that were up to now inapproachable due to not considering the flexibility of the protein.

2. Biology

To understand the function of a protein, it is necessary to know the physiological system it belongs to. The following section should provide an insight to the biological background of neurotransmission in the human central nervous system, emphasizing the molecular level.

2.1 *Nerve Cells*

A typical mammalian nerve cell, independent of its further structural classification, consists of the soma, which is the cell body including the nucleus, the dendrites, responsible for the majority of information input, and the axon, transporting the signal to target neurons. This division in an incoming and an out-bound part explains the unidirectionality of nerve impulses, what is ensured by the fact that on the one hand dendrites have no possibility to secrete neurotransmitters and on the other hand no chemoreceptors are available on the terminal axon.

2.2 *Neurotransmission*

In most cases, the communication between nerve cells requires that neurons transform an electrical into a chemical signal in order to pass on incoming impulses. The arriving action potential leads to opening of potential-dependent Ca^{++} -channels, causing exocytosis of vesicle-stored neurotransmitters. This stimulated neurotransmitter release on the one hand and specific detection of the neurotransmitter by one or more contiguous – up to 15.000 - cells on the other hand is located at specialized junctions called synapses between the terminal axon and the acceptor dendrites. Therefore, orchestrated release and uptake of signal molecules are crucial to achieve orderly communication between neurons. (Gouaux 2009) An incoming

impulse can immediately raise extracellular neurotransmitter concentrations by 10^5 fold. That leaves behind a challenging task for the adjacent cells to quickly re-establish pre-signal conditions which is necessary to ensure reasonable impulse intervals. Therefore, the removal of transmitters from the adjacencies of their receptors can be performed by three different mechanisms. At first, the transmitter diffuses into the interstitial fluid, being diluted to inoperative concentrations. Second, some substances as ATP, acetylcholine or neuropeptides are directly enzymatically degraded in the synaptic cleft. The third mechanism is the specific uptake of the substances by membrane-bound transporter proteins to either the emitting presynaptic cell (reuptake) with the possibility of recycling them by anew vesicular storage, or into other surrounding cells, that lack of specific carriers and therefore just have the possibility of degradation. There are rather large differences between the vesicular, axolemma-situated and adjacent cell membrane-bound neurotransmitter transporters in driving force as well as in structure and pharmacology. For instance, vesicular uptake is granted by an in-place generated electrochemical proton gradient, whereas the reuptake through the axolemma is ensured by a sodium gradient between the extra- and intracellular side of the neuron. Acting as Na^+ -neurotransmitter co-transporters, there are highly specific carriers for serotonin, dopamine, norepinephrine, glutamate, GABA and glycine. (Aktories et al. 2005)

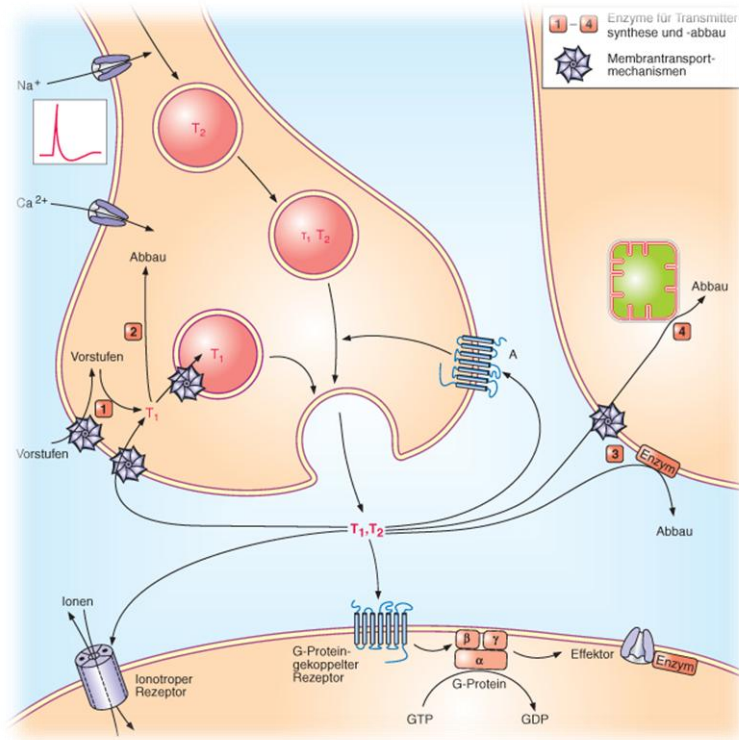


Figure 1: Synaptic signal transduction, taken from Aktories et al., 2005

2.2.1 Monoamine Transporters

As mentioned before, monoamine transporters are the main instruments to clear the synaptic cleft of certain small-molecule neurotransmitters with a positively charged amino group. Below there is a scheme describing the main monoamine cycles, also showing some prominent drugs interacting with the respective transporter.

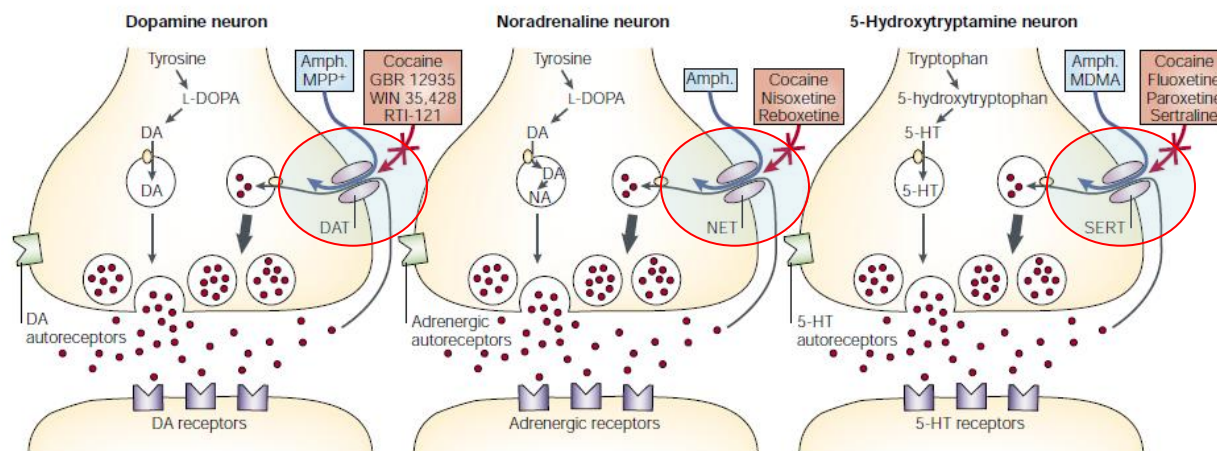


Figure 2: Scheme of monoamine-related synaptic terminals, modified from Torres, 2003

All human monoamine transporters are members of the solute carrier 6 (SLC 6) family, but each one is localized on a different gene. NET (SLC6A2) spans 45 kb respectively 14 exons on chromosome 16q12.2, the DAT (SLC6A3) gene has a size of 65 kb spread over 15 exons on the 5p15.3 chromosome, whereas the human SERT (SLC6A4) is divided into 13 exons on the 17q11 chromosome, covering 24 kb. (Torres et al. 2003)

The expression of monoamine transporters in the central nervous system is very tightly bound to neurons containing their substrate transmitter. Dopamine transporter (DAT) expression is prominent in the cell bodies of the substantia nigra and the ventral tegmentum, both located in the mesencephalon floor. In case of the noradrenaline transporter (NET), the area of interest is the locus coeruleus and several other nuclei in the brain stem, as serotonin transporters (SERT) can be found in the dorsal and median raphe nuclei.

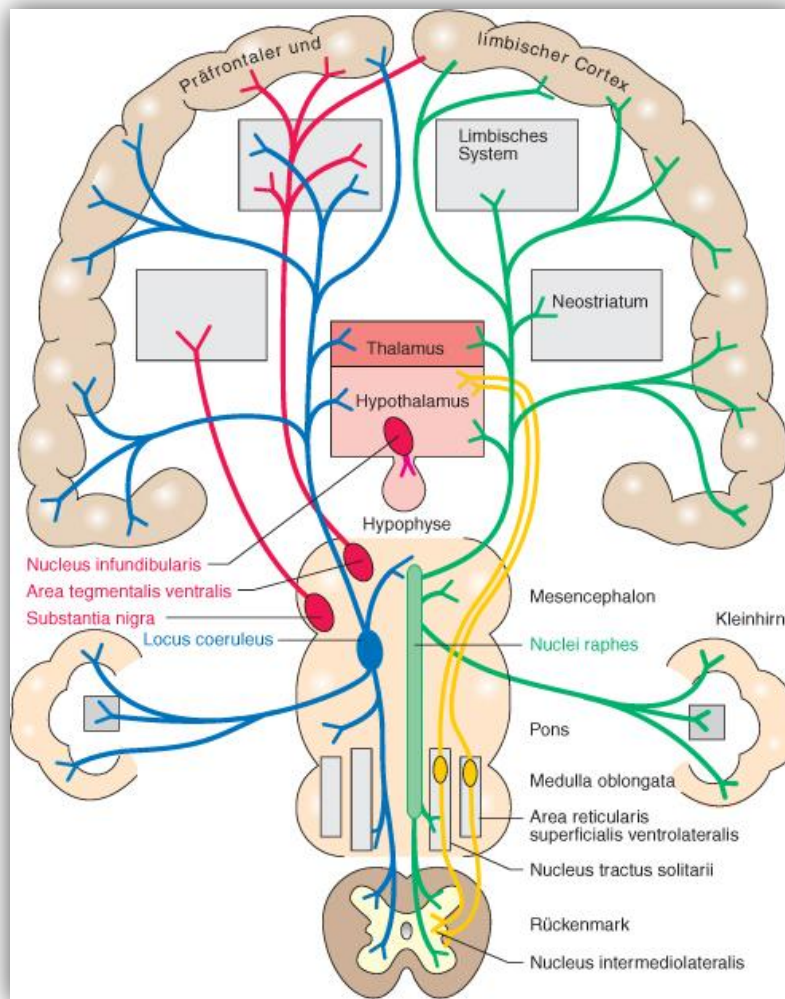


Figure 3: Serotonergic (green), dopaminergic (blue) and noradrenergic (red) nerve tracts, taken from Aktories et al., 2005

2.2.2 SERT Structure and Function

The monoamine transporters are essentially needed during the first step of transmitter recycling, so the transport itself has to be tightly regulated to avoid uncontrolled flux between nerve cell and the surrounding synaptic cleft. The unique structure of the SLC 6 protein family assures regulated transport by different key properties. First, the binding site is responsible for highly selective substrate binding. As the SERT is a co-transporter, also ion binding sites in the

correct relative position to the substrate binding site are critical for the transport. Second, accessing the binding site must not be possible from both sides of the membrane simultaneously to avoid uncoupled flux. Just the coupled binding of serotonin and Na^+/Cl^- allows the required conformational changes of the protein from the open-to-out state to the occluded state and finally the open-to-in configuration, where substrate and ions can be released. Subsequently, binding of a potassium ion initiates the return to the first state. (Gary Rudnick 2006)

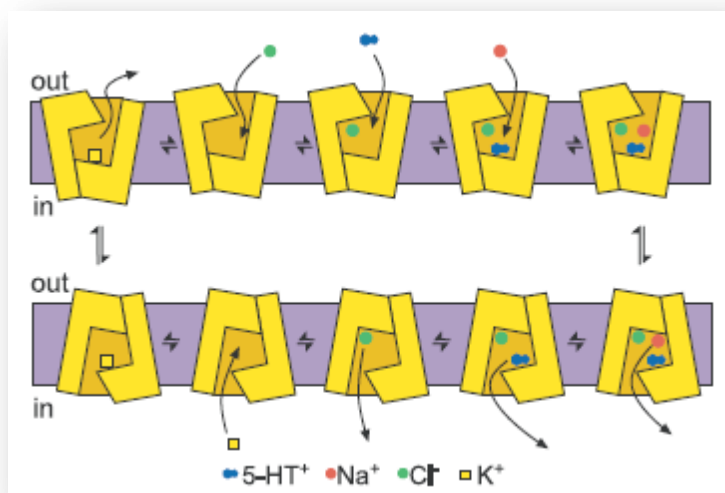


Figure 4: Possible mechanism of serotonin transport, taken from Rudnick, 2006

After the structure determination of the LeuT by the group of Eric Gouaux in 2005 the highly resolved data granted insights to the topology of the substrate binding site of monoamine transporters. Two sodium ions bound together with leucine in the active site were identified. Comparing the residues involved in ion binding with the corresponding residues in SERT showed that most were identical or conservatively substituted, as for example two serines were replaced by threonines. The effect of presence or absence of the sodium ions has a different impact on the binding quality in SERT depending on the particular ligand. Serotonin, for instance, is not affected by missing sodium, but β -CIT seems to need at least one sodium ion

coordinated by residues corresponding to those in LeuT. (Humphreys et al. 1994) To be on the safe side, two sodium ions were included during the modeling.

2.2.3 Neurotransmitters

Neurotransmitters are presynaptic released substances with influence on the following cell. They are members of four major chemical groups: Amines, amino acids, neuropeptides and nucleotides. Very prominent examples of these neurotransmitters are biogenic amines such as 5-hydroxytryptamine (serotonin), norepinephrine (noradrenaline) or dopamine. (Torres et al. 2003) The importance of those small molecules is underlined by the fact that they are all closely related to proteinogenic amino acids, as they can be derived from them by decarboxylation and eventually one or more hydroxylation steps. (Fig. 1) They control various functions of the human body ranging from locomotion to hormone secretion, which requires precisely regulated signal duration and intensity.

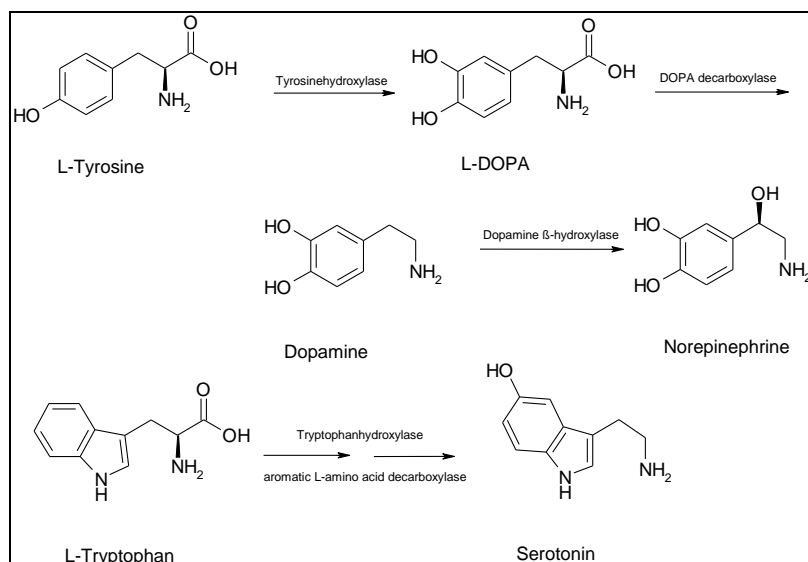


Figure 5: Amino acids as neurotransmitter precursors

All neurotransmitters except neuropeptides, that are built in ribosomes and delivered to the nerve ends by axonal transport, are locally synthesized out of precursor molecules and packed in the nerve ends. Vesicular storage not only holds ready well-defined doses of signal

molecules, but also prevents the freshly produced neurotransmitters from enzymatic degradation.

3. Pharmacology

Many agents affecting the brain function including antidepressants, neurotoxins and psychostimulants target monoamine transporters. The mechanism of action depends on the drug, as cocaine acts as a competitive, non-selective inhibitor, and amphetamine-like substances achieve reverse transport of neurotransmitters, thus pumping them into the extracellular space instead of taking them up. Both result in high extracellular monoamine levels. For the studies carried out on the hSERT a set of seven compounds was used, as for some the interaction pattern with the target is quite well-defined and for some it is rather unknown.

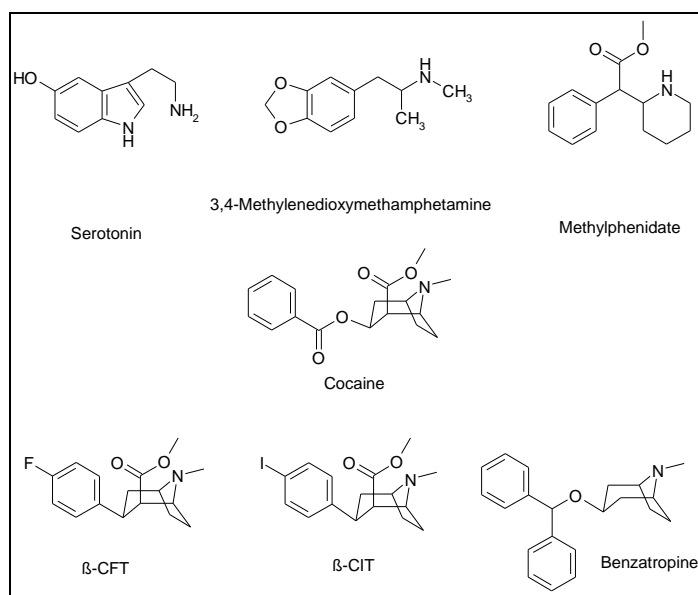


Figure 6: Ligands used for the study

3.1 Serotonin

The first used ligand was the natural substrate of the SERT, serotonin. The highest serotonin quantities in the human body can be found in the enterochromaffin gut cells of the gastrointestinal tract and in thrombocytes. Compared to them, neuronal serotonin levels are rather low; nevertheless serotonergic neurons are assumed to be responsible for the regulation of sleep patterns, mood, sensation of pain, ingestion and body temperature. As mentioned earlier, the amino acid tryptophan is the precursor of serotonin. As it is an essential amino acid, the speed of the synthesis is limited by the supply of tryptophan. In the first step of biosynthesis, it is hydroxylated to 5-Hydroxytryptophan by the cytoplasmatic tryptophanhydroxylase. During the second step 5-Hydroxytryptophan is decarboxylated to serotonin by the aromatic L-amino acid decarboxylase, the same enzyme that is also involved in catecholamine synthesis. This also explains the usual abbreviation '5-HT' for serotonin, as the resulting decarboxylation product is 5-HydroxyTryptamin. After the final step serotonin is transported to storage vesicles, waiting for its use. With the exception of 5-HT₃ as the only ionotropic receptor, serotonin targets a variety of G protein coupled receptors activating intracellular second messenger cascades. After reuptake, serotonin is either stored in the also partially recycled vesicles or degraded to 5-Hydroxyindolacetaldehyd by monoamine oxidases with the subtypes MAO-A and MAO-B, being further metabolized to 5-Hydroxyindoleacetic acid (5-HIAA) as the main metabolite in the urine. Serotonin and its metabolites can also be conjugated with sulphuric or glucuronic acid, partially already in the central nervous system.

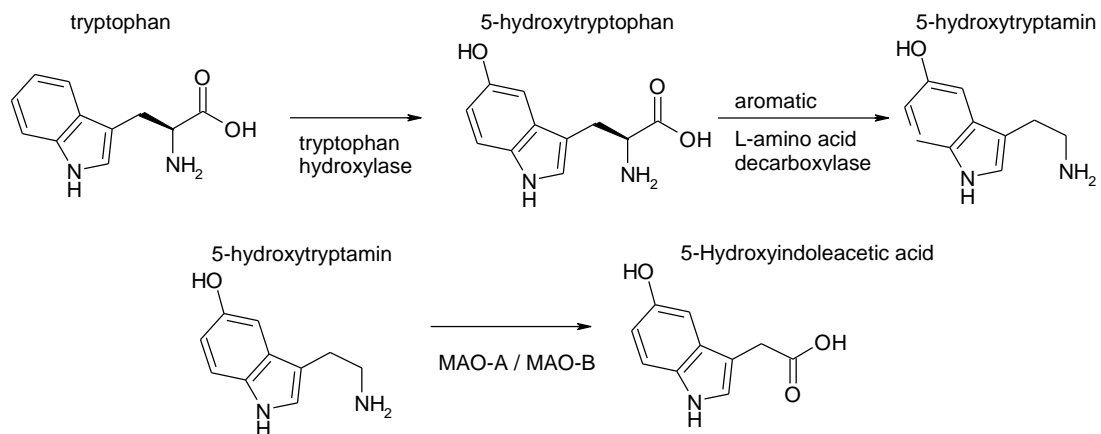
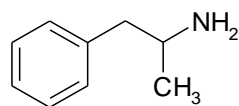


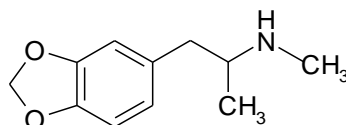
Figure 7: Serotonin pathway

3.2 MDMA

3,4-Methylenedioxyamphetamine ('Ecstasy') is a semi-synthetic hallucinogen with high structural similarity to amphetamine. It was first synthesized and patented by Merck Pharmaceuticals in 1914. Precursor compounds can be found in plants like nutmeg, saffron and vanilla beans. (Lyles & Cadet 2003) Its action on the serotonergic system depends on the block of 5-HT reuptake and induction of 5-HT release, whereas a variety of other targets is also affected. Producing euphoria and mild stimulation, its use as a recreational drug in Europe and North America is increasingly popular among young people. Though often described as a relatively harmless drug, acute toxic side effects like malignant hyperthermia, hepatitis and convulsions can occur. In addition, long-term use of MDMA is strongly associated with neurotoxic effects observed at frequent consumers, including selective degeneration of serotonergic axons (White et al. 1996)



amphetamine



3,4-methylenedioxyamphetamine

Figure 8: Comparison of Amphetamine and MDMA

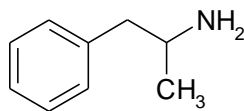
The precise role of MDMA in the process of destroying brain serotonin terminals is still unclear, but recent studies also suggest the involvement of metabolites like 3,4-methylenedioxyamphetamine (MDA). There are two main pathways in MDMA metabolism. The first pathway leads via O-demethylenation to 3,4-dihydroxymethamphetamine (HHMA) and further O-methylation to 4-hydroxy-3-methoxymethamphetamine (HMMA), and is followed by subsequent conjugation with glucuronic acid or sulfate. The second pathway starts with N-demethylation to MDA and ends up in corresponding glycine-conjugated benzoic acid derivatives formed by several desamination and oxidation steps. (Mueller et al. 2009)

	MDA ¹	6-OHMDMA ^{2,4}	THMA ^{2,3,4}	HHMA ^{3,5,6}	HMMA	HHA ⁷	HMA ⁷
R ₁	H	CH ₃	CH ₃	CH ₃	CH ₃	H	H
R ₂			OH	OH	CH ₃	OH	CH ₃
R ₃			OH	OH	OH	OH	OH
R ₄	H	OH	OH	H	H	H	H

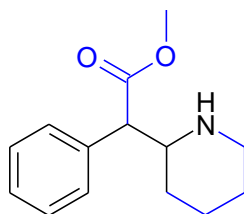
Figure 9: MDMA metabolites, taken from Mueller et al., 2009

3.3 Methylphenidate

This prominent ring-substituted amphetamine-like psychostimulant currently mainly used in attention-deficit/hyperactivity disorder (ADHD) and narcolepsy has a high inhibition potency for DAT and NET ($K_i = 0.1 \mu\text{M}$). Also a certain potency, even if much weaker, for SERT ($K_i = 100 \mu\text{M}$) can be observed. (Han & Gu 2006) The potency in blocking DAT is as high as the one of cocaine, with less side effects. (N D Volkow et al. 1999) Despite the low SERT affinity it was chosen for interaction studies in that protein, both to see if a noteworthy difference to high-affinity ligands is being reflected in the number of docking poses retrieved and calculated binding energy. Also, at least in knock-out mice a mentionable effect of methylphenidate on the SERT has strongly been suggested, hence being subject of discussion. (R R Gainetdinov & M G Caron 2001)



amphetamine



methylphenidate

Figure 10: Amphetamine - methylphenidate relationship

The cyclic structure of methylphenidate causes the formation of two chiral centers in 2 and 2' position, but experiments have shown that the R,R-(+)-threo configuration of the molecule is far more potent than the S,S-(−)-threo enantiomer. (Aoyama et al. 1996) Degradation is 90% done by hepatic cytochromes, leading to oxidized metabolites like ritalinic acid, hydroxy- and

oxymethylphenidate, less than 1% are excreted unmodified with the urine. The main metabolite, ritalinic acid, was also supposed to possess pharmacological activity, but later studies led to the assumption that it does not seem to have important activity contributions. (Yu-Shin Ding et al. 2004)

3.4 Cocaine

The illegal status of the tropane alkaloid originates from its psychotropic and highly addictive potential. Despite this, the drug itself is a potent inhibitor of all monoamine reuptake transporters, making it highly interesting for this study. The substance extracted from leaves of *Erythroxylum coca* was primarily used as stimulant among South American natives. The original use of cocaine in the western world as a local anesthetic today is very limited due to potential cardiotoxicity and its high vasoconstrictor activity, which yet can be desirable during surgery to avoid bleeding, but derivatives with a more favorable side effect profile outstripped it almost entirely. (Kovacic 2005) Apart from that, the importance of cocaine as a drug of abuse is enormous. Nearly any route of administration can be considered for cocaine abuse, whereas the major ones are intranasal, intravenous and through smoking. This is followed by a rapid metabolism by various liver enzymes, whereas the main metabolite benzoylecgonine is formed by a nonspecific liver carboxylesterase, also responsible for the formation of coca ethylene in the presence of ethanol, both metabolites being important for toxicological screening and quantification procedures. (Brzezinski et al. 1994) Also, cytochrome P-450 enzymes are involved in cocaine metabolism, as the 3A4 subtype acts as cocaine N-demethylase resulting in norcocaine, followed by several other CYP subfamily related metabolism steps. (Pellinen et al. 2000)

3.5 Benztropine

Being a cocaine-like tropane alkaloid, Benztropine is also an inhibitor of monoamine transporters, yet showing higher DAT specificity, which was the reason to take it as a lead compound for even more DAT-selective drugs. (Rothman et al. 2008) Due to its anticholinergic action it is sold under the name Cogentin® for the palliative treatment of Parkinson disease, as the imbalance between dopamine and acetylcholine can be slightly reduced. Though SERT affinity is low, its structural relationship to cocaine made it a candidate for further investigation.

3.6 CFT

(-)-2 β -Carboxymethoxy-3 β -(4-fluorophenyl)tropane (WIN 35,428, β -CFT) is a cocaine analogue with stimulating properties. Although the higher affinity and extended half-life it is not an explicitly restricted substance, as the synthesis is comparably difficult and involves the restricted intermediate ecgonine, making it highly unlikely to be used as a recreational drug. The main use of β -CFT is in radiolabeled imaging studies, when cocaine itself is not convenient because of its inferior pharmacokinetic profile. Metabolism can be considered similar to cocaine, with respect to its longer half-life. Compared to cocaine, β -CFT has higher affinity for both DAT and SERT, but decreased potency at the NE transporter. (Boja et al. 1992)

3.7 CIT

(-)-2 β -Carbomethoxy-3 β -(4-iodophenyl)tropane (RTI-55, β -CIT) is another analogue of cocaine mainly for the use in scientific research showing highly increased potency. By increasing the halogen substitution on the 4-position of the phenyl ring from fluorine to iodine SERT affinity was meant to be increased without affecting DAT inhibiting potency, what was successful to some degree. (Weed et al. 1995) Possible therapeutic use is very limited due to its high addictive potential.

4. Computational Background

In addition to the initially mentioned fact that an X-ray structure of the SERT remains missing so far, several other challenges already encountered in previous studies had to be faced and managed, as in reality both receptor and ligand are not rigid but flexible, implicating increased complexity of the calculations. Also, the use of different software packages often caused compatibility problems.

Generally, the use of information technology in medicinal chemistry has increased nearly exponentially within the last decades. As the development of new drugs has become immensely expensive no pharmaceutical company can afford many failures, so the call for cheap methods to help predicting potent compounds became loud. Today the basis of investigations on protein-ligand interactions is prevalently a combination of software packages relying on a precise set of formulae describing different forces that can occur between atoms and molecules and, on the other hand, high-resolution structural data of the target protein, usually yielded from X-ray or NMR experiments. If this structural data is not available, one has to fall back on models. Suggestions arising out such in silico experiments can further be confirmed or vitiated by – mostly more expensive - in vitro or in vivo experiments.

4.1 *Molecular Modeling*

'A Model must be wrong, in some respects; else it would be the thing itself. The trick is to see where it is right' – Henry A. Bent, American chemist

Going down to the scale of molecules or even atoms, the representation of the objects of interest in a conventional way like taking pictures is by far more difficult, as the wavelength of visible light is not quite well suited to investigate structures of – even roughly - similar scale. So

even if other, better applicable techniques like NMR or X-ray crystallography are used, the human eye needs the information to be converted to the visible optical range between 380 and 780 nanometers, thus any effort of visualizing molecular structures must lead to one or another sort of a simplified model. In the history of structure determination, the huge expense of arithmetical calculations stood in a glaring contrast to its output of two-dimensional representations.

The only true 3D alternative was the use of molecular kits like the illustrious Dreiding Models. To demonstrate the fussiness but nonetheless efficiency of those models one has to keep in mind that even Watson and Crick had to use such kits to model base pairing. During the 1970s, computational power became efficient enough for the first pseudo-3D visualizations of small molecules, further improving through the fast evolution of processors. Then, as now, to grant the generated structures the highest possible plausibility, the geometry always should be optimized to the energy minimum state, which can be done by so-called force fields. They consider atoms as a kind of rubber balls of different size, according to the atom type, and bonds as a spring-like connection of different length and angles, tending to certain standard values. Based on Hooke's law the overall total energy of a molecule is minimized according to the formula:

$$E_{\text{tot}} = E_{\text{str}} + E_{\text{bend}} + E_{\text{tors}} + E_{\text{vdw}} + E_{\text{elec}} + \dots$$

This describes the total energy E_{tot} of a molecule by summing up bond-stretching energies with bending, torsional, van der Waals and electrostatic energy terms. (Höltje & Folkers 1996)

Depending on the algorithm, such equations perform well on certain types of molecules, but may fail totally on others, as atoms have slightly different properties according to their molecular arrangement.

Force Field	Small Molecules	Organic	Proteins	Nucleic Acids	Carbohydrates
MMFF94x	X				X
AMBER99			X	X	
CHARMM27			X	X	
CHARMM22			X		
OPLS-AA	(X)		X		
EngH-Huber			X		
PEF95SAC					X
TAFF	X				
MM3*	X				

Table 1: Common force fields used in commercial software packages

4.2 Homology Modeling

As brought up in the introduction, the structures of membrane-bound proteins are often not available due to crystallization difficulties, so it is quite common to try constructing 3D-models based on homologous proteins offering structures. In doing so, the primary sequence of the target protein has to be aligned with the sequence of the template protein. This can be done manually or by the software itself. If there are some unconserved regions where template and target vary a lot, the loops have to be sampled, i.e. predicted, by the molecular modeling software. The procedure of these approaches is quite well documented, but the accuracy always depends on the sequence homology. Typically, if the sequence identity is greater than 50%, the model performs approximately as well as the crystal structure. Generally, more than

30% are said to be needed to construct reliable homology models. (Oshiro et al. 2004) Of course, a model built has to be carefully refined and validated to ensure the highest possible reliability. Concerning the geometry several parameters have to be checked.

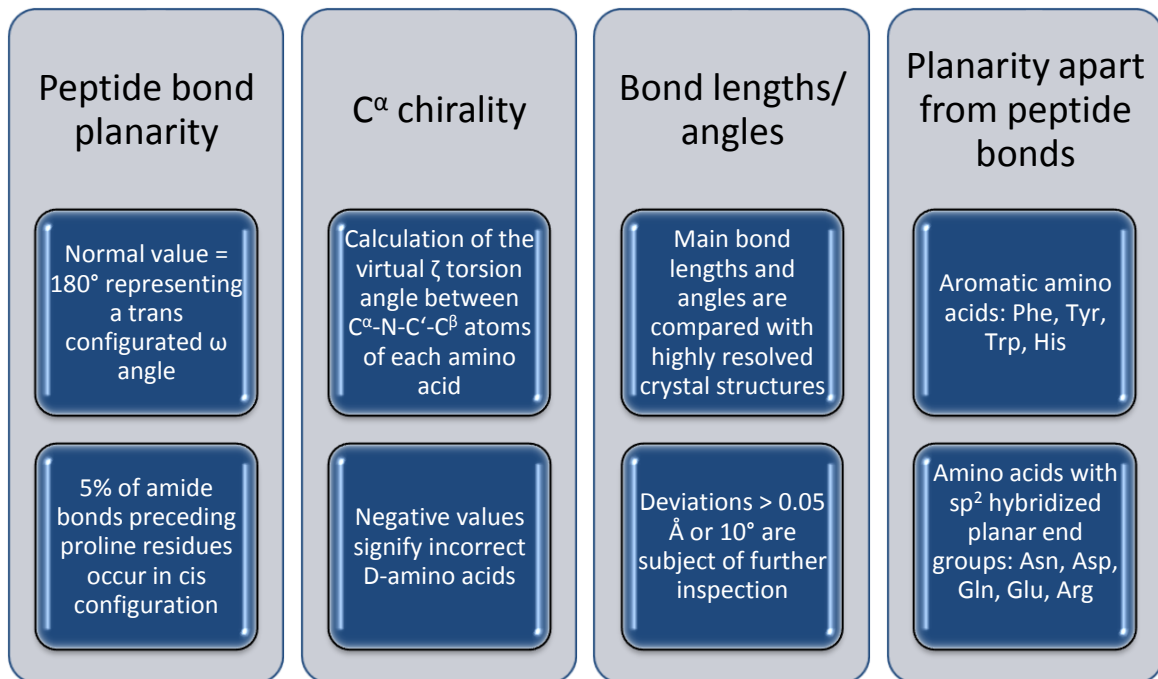


Figure 11: Checklist for the quality of protein models; taken from Höltje and Folkers, 1996

Also the packing quality plays an important role. The tight packing of secondary structures can stabilize the overall stability of a protein structure, so it can be a valuable measure for the reliability of a model.

4.3 Docking

The crucial part of calculating interactions between a protein and its possible ligands is called the docking procedure. There are several docking tools using different algorithms, but basically there are three different methods:

- a) Rigid body docking: protein and ligand both are kept inflexible
- b) Semi flexible docking: rigid protein and flexible ligand
- c) Fully flexible docking: both protein and ligand are kept flexible

There were several molecular modeling software packages available in our research group, so the choice had to follow the needs of the project. Here an overview over the most important programs available:

Name	Molecular Environment	Operating Environment	Schrödinger Suite	GOLD	SYBYL
Company	Chemical Group	Computing	Schrödinger, LLC	CCDC	Tripos, L.P.
Graphical User Interface	MOE		Maestro	HERMES	SYBYL
Operating System(s)	Windows, Linux		(Windows), Linux	Windows, Linux	Linux, SGI
Latest Released Version	MOE 2008.10		Schrödinger Suite 2009	GOLD 4.1	SYBYL 8.1
Molecule Building Module	Yes		Yes	No	Yes
Ligand Flexibility	Yes		Yes	Yes	Yes
Protein Flexibility	No, Subsequent Minimization	Energy	Yes	Limited to 10 Side chains	No (Developing)

Table 2: Comparison of Main Software Features

The more flexibility is granted to the system, the more accurate results can be achieved, at the cost of large computational expense. A high grade of flexibility is especially useful while a homology model based on a low-resolution template is being used, as it is the case in this study, so it was advisable to follow the path of fully flexible docking. According to the needs of the experiments the Schrödinger software package was chosen for fitting best to the requirements and combined several modules that would have been needed to be taken from different providers, exposing the data to the omnipresent threat of compatibility problems.

4.4 Scoring

The results of a docking run need to be ranked in regard to their quality. Also, during a docking process it is useful to estimate the fitness of a ligand in the receptor during the placement optimization procedure.

Usually one will try to predict the free energy of binding according to the Gibbs-Helmholtz equation, which can be done by applying one or more of various scoring functions, differing in speed and accuracy.

$$\Delta G = \Delta H - T\Delta S$$

ΔG is the free energy of binding, ΔH describes the enthalpy, T stands for the temperature and ΔS is the entropy term.

To transform ΔG into the binding constant K_i the following equation can be used, involving the gas constant R :

$$\Delta G = - RT \ln K_i$$

There are three main approaches for scoring functions. At first the so-called Empirical Scoring Functions can be mentioned. They aim to calculate binding affinities by adding property terms known to contribute to drug binding by using multilinear regression to optimize the particular coefficients of the computed terms with the aid of a training set of appropriate protein-ligand complexes, limiting the significance of the function to structures similar to the training set.

The members of the second group are called Force-field-based Scoring Functions. As the name implies, they are based on force fields that use molecular mechanics. Van der Waals interactions are expressed by a Lennard-Jones potential, electrostatic interactions are described by the Coulomb energy. The energy of those nonbonded interactions can be described by the following term:

$$E = \sum_{i=1}^{lig} \sum_{j=1}^{rec} \left[\frac{A_{ij}}{r^{12}} - \frac{B_{ij}}{r^6} + 332 \frac{q_i q_j}{D r_{ij}} \right]$$

A_{ij} and B_{ij} describe the van der Waals repulsion and attraction of two atoms i and j at a distance of r_{ij} , the point charges of the atoms i and j are given by q_i and q_j , and the dielectric function is D . The factor of 332 allows the conversion of electrostatic energy to kilocalories per mole.

As entropy contributions are not considered in force-field calculations, large and polar molecules tend to be scored best, which should be kept in mind when looking at the results.

Knowledge-based Scoring Functions deny breaking down protein-ligand binding energy into all mentionable contributions for further detailed analysis. Also, they try to avoid the threat posed by the limited chemical space provided by the training set in empirical scoring functions by using potentials of mean force. Those define the occurrence of protein-ligand atom pairs of

certain atom types and distances as a measure for the quality of the fit. The sum of all interatomic protein-ligand complex interactions results in the score.

Still, accurate predictions of binding free energies are impossible, as some parameters as the role of water or exact contributions of protein flexibility remain hardly computationally accessible. So the focus lies on post-processing docking outputs to facilitate the job of the scoring functions. (Höltje & Folkers 1996)

As an effort to improve the performance of scoring-based ligand ranking, several different scoring functions can be combined. There are two hypotheses why the results should improve by rescoring strategies. The one called 'consensus scoring' draws on the assumption that active compounds should be top-ranked by several scoring functions, whereas false positives should be reduced by fusing the data. (Charifson et al. 1999) The other strategy is to utilize the different strengths of scoring functions, called the complementary hypothesis, implying that the combined power leads to improved results. However, it can happen that one of the greatest benefits of rescoring, which is the reduction of computational costs for improving the results by several scoring algorithms that are generally faster than multiple docking procedures, gets lost if highly accurate but computationally expensive calculations as MMGBSA are involved. (O'Boyle et al. 2009)

5. Materials and Methods

5.1 *The Model*

The first requirement to allow docking studies on the hSERT was the presence of a 3D structure. As mentioned before, an X-ray structure is not yet available, so the task was to create a homology model processible by well-established molecular modeling software. To be able to meet the demands for creating a powerful model there was an urgent need for a suitable template, a quite difficult task for a membrane-bound protein.

5.1.1 *The Template*

At a first glance it seems quite unlikely that the human serotonin transporter and a bacterial amino acid transporter have very much in common, but their structural similarity turned out to be a piece of luck for the modeling procedure. The availability of a crystal structure and the high structural similarity made the leucine transporter of the underwater bacterium *Aquifex aeolicus* (LeuT_{aa}) the best template on-hand. Both belong to the Na⁺/Cl⁻ dependent transporter family catalyzing a thermodynamically uphill carriage of their substrate. The advantage of the LeuT_{aa} was its ability to form crystals, a crucial step in X-ray structure determination. The yielded 3D information of the protein 70Å in height and about 48Å in diameter revealed the presence of twelve helical transmembrane segments (TM 1-12), with both amino and carboxy termini looming into the cytoplasmic side of the membrane. The protein core consists of two pseudo-symmetric counterparts formed by TM 1-5 and TM 6-10 and is responsible for substrate and ion binding. This similarity cannot be seen in the primary sequence, but it is affirmed by the fact that the 130 α-carbons of the residues belonging to TM1-TM5 and the corresponding atoms of helices TM6-TM10 show a root mean standard deviation (RMSD) of just 5.3Å by a rotation of 176.5°. (Yamashita et al. 2005)

5.1.2 Sequence Alignment

Comparing the primary sequence of LeuT with the one of hSERT, one will just find 21% identical residues. The homology although can be numbered 40 to 45%. (Zhou et al. 2007) That means, amino acids with common properties, e.g. aliphatic residues like leucine and isoleucine or acidic amino acids like glutamic acid and aspartic acid, are not identical but homologue in function, so even with a relatively poor sequence identity of two protein sequences, their homology can be significantly higher. In addition, despite the problematically low overall similarity of LeuT and hSERT the main binding site is highly conserved, strongly suggesting a common mechanism of transport. In that crucial region, 19 out of 25 residues in a range of 5 Å around the centroid of the binding site are identical. Nonetheless, many regions remote from the direct substrate interaction site but important for the transporter function are quite poorly conserved and therefore are difficult to align. The group of Beuming et al. compared 177 eucaryotic and 167 procaryotic NSS protein sequences to optimize those problematic regions, resulting in a comprehensive alignment of the LeuT with important eucaryotic members of the NSS family. (Beuming et al. 2006)

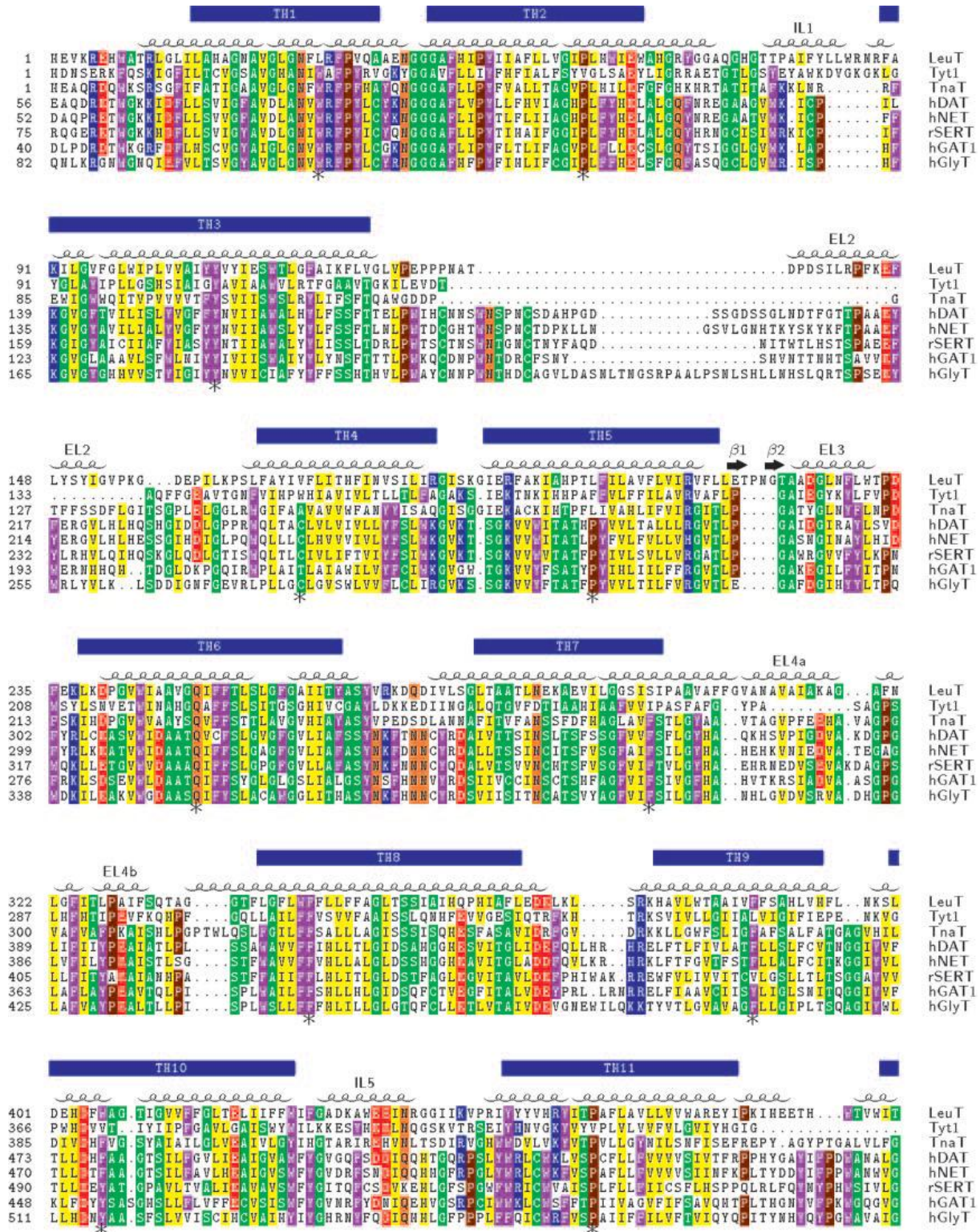


Figure 12: Sequence alignment of LeuT with eucaryotic transporters, Beuming et al., 2006

5.2 Schrödinger Suite

Having the main information needed for building the homology model at hand, still there was the need for a tool suitable for the actual performing of the task. The Schrödinger Suite 2008 of Schrödinger, LLC was chosen for several reasons. First, the embedded docking application allows both protein and ligand flexibility, which was mentioned before to be important for homology models with uncertain exact structure, second, the MAESTRO graphical user interface (GUI) combines most of the applications needed from the primary alignment to the final evaluation, which is beneficial to avoid compatibility problems that can easily occur when using several different software packages.

5.2.1 Ligand Preparation

The ligands introduced in section 3 were drawn in Maestro (Version 8.5, Schrödinger, LLC, New York, NY, 2008) according to the structural information yielded from Chemical Abstracts Service (CAS) via SciFinder (Chemical Abstracts Service, Brandywine, MD, 2007). If due to chiral centers more stereoisomers were available, the most active described in literature was chosen. The generated ligand library was adapted to physiological state, e.g. the primary amine of serotonin was protonated as it would be under binding conditions. The further energy minimization was performed using the OPLS force field.

5.2.2 Prime Structure Prediction

For the actual building of the homology model the Prime structure prediction application needed the input of the primary sequence of hSERT. For this purpose the SWISSPROT/Expasy site (www.expasy.org/sprot, now www.uniprot.org) delivered the essential information under the access number P31645 (SLC6A4_HUMAN). The 630 amino acid long sequence was saved in

the FASTA file format containing also information about the species and the protein family, and used as a query input for Prime. Out of several LeuT template structures with similar resolution that were found the crystal structure with the PDB code 2QB4 was chosen, also offering the co-bound substrate leucine, the antidepressant desipramine and two sodium ions in the inner binding site. For the proper use as a template, the protein was refined using the Protein Preparation Wizard that deleted needless water and solvent molecules. Furthermore, the residues Asn133 and Ala134 missing in all available LeuT X-ray structures had to be added in order to close the gap between Pro132 and Thr135. Although the loop region containing the gap is not supposed to be important for ligand binding the missing residues had to be replaced to facilitate an energetically stable template structure. To do so, the sequence of the pdb file was aligned with its own but complete primary sequence again obtained from the SWISSPROT site using the Prime structure prediction module.

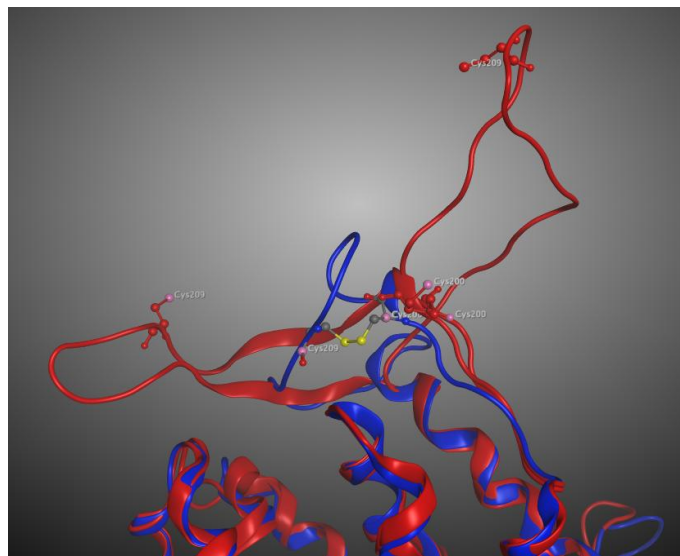


Figure 13: Comparison of EL 1 in 2A65 (red) and 2QB4 (blue) based models

The processed, so to say repaired LeuT structure now was able to serve as a proper template for creating the SERT model. According to the prior alignment scheme of Beuming et al. the two

sequences were manually aligned, allowing the sequence of the hSERT to adapt to the 3D structure of the 519 residue-long prokaryotic relative. Some loops in unconserved regions of the protein needed to be built and/or fixed, like a disulfide bond between Cys200 and Cys209 formed using a detection tool downloaded in the Schrödinger Script Center.

After the compulsory energy minimizing step the model was checked against SERT models previously built in our group and by the group of Celik et al., based on LeuT structure entry 2A65. The overall RMSD value of 4.17 Å between the different models was comparably high resulting from differences mainly in the larger extracellular loop regions. For example, the 60 residues between TM3 and TM4 are extremely flexible, though the cystine moiety suggested in the topology diagram of Rudnick (Gary Rudnick 2006) also detected in the newly built model seemed to slightly stabilize the internal structure of the loop.

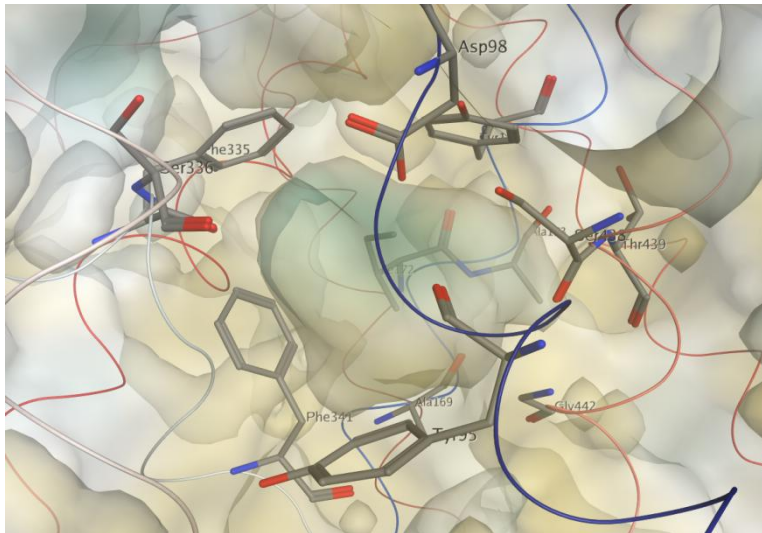


Figure 14: Inner binding site

Nevertheless the region is not of direct interest for substrate binding, so even major deviations do not pose a problem for the equivalence of different models. The considerably low RMSD

value of 0.43 Å concerning the residues forming the binding site stands in contrast to the higher values in the peripheral regions, thus granting comparable results.

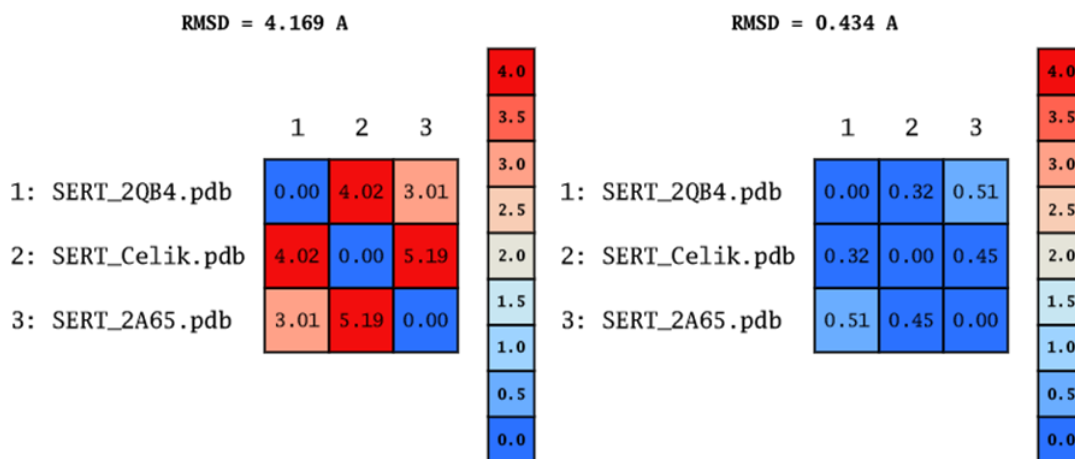


Figure 15: RMSD values of different SERT models

5.2.3 PrimeGlide Induced Fit Docking

The main difference between standard GLIDE docking and the Induced Fit Docking workflow is, as mentioned before, the protein flexibility apart from normal energy minimizing steps used by nearly any docking software. That means, in principle every ligand pose has its corresponding receptor with a unique structure, making the data processing especially challenging. For the generation of the receptor grid either the centroid of an already placed ligand or the centroid of defined residues were selectable. It turned out that the best choice were the residues forming the inner binding site.

Localization	Residues
TM 1	Tyr95, Asp98
TM 3	Ala169, Ile172, Ala173, Tyr176
TM 6	Phe335, Ser336, Phe341
TM 8	Ser438, Thr439, Gly442

Table 3: Localization of the residues forming the inner binding site

As the template structure represented the occluded state of the transporter, the space for the ligand placement inside the binding site was rather limited, so the side chains of the aromatic residues were initially removed to be rebuilt after the primary placement. As major influence of forces outside the protein could be excluded due to the occluded nature of the model the use of an implicit membrane meaning huge computational expenses could be omitted. With the selection of the desired precision mode responsible for the calculation time (Standard Precision/SP or Extra Precision/XP mode) and the definition of the prior generated ligand library the job was ready to be started.

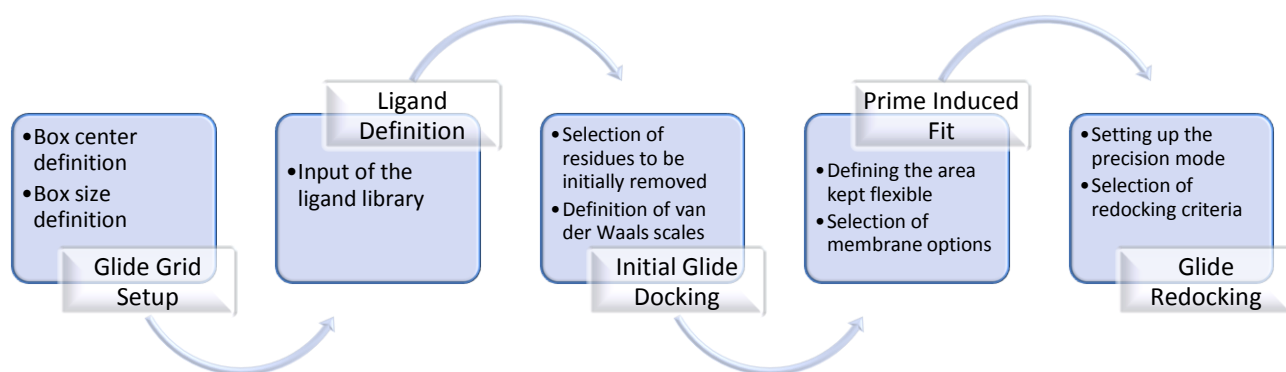


Figure 16: General workflow

5.3 *Data Processing*

The large amounts of data generated during a docking experiment have to be screened for their essential information prior to being able to speak of results. The data analysis was supported by several computational tools available for sorting and assessing the output files.

5.3.1 IFDScore

Looking at the output files created during the docking process, 49 different weighting and energy contribution values for each pose were offered. After all, the poses were by default ranked according to the IFDScore value, a virtual binding free energy term. The composite score consists of GlideScore, standing for protein-ligand interaction energy, and a 5% admixture of PrimeEnergy, the protein molecular mechanics energy term. It was developed to eliminate predicted structures with an energy gap large enough compared to the lowest PrimeEnergy value to be considered noise.

5.3.2 Clustering

To fetch also sterical information the poses had to undergo a hierarchical clustering procedure by operating the embedded XCluster application. The module embedded in the Schrodinger software package allows several modes of hierarchical clustering.

The sterical pose features are compared either superposed or in place, according to RMSD values or clustering levels corresponding to the number of different docking poses.

The application also offers a graphic for a quick overview over the clustering behavior of the members of a docking run, so called clustering statistics. Depending on the chosen RMSD or clustering level value the thresholds for a cluster membership differ, modifying the cluster sizes. In most cases it is useful to visualize the calculated clusters for final judgment of the determination quality of the chosen clustering mode.

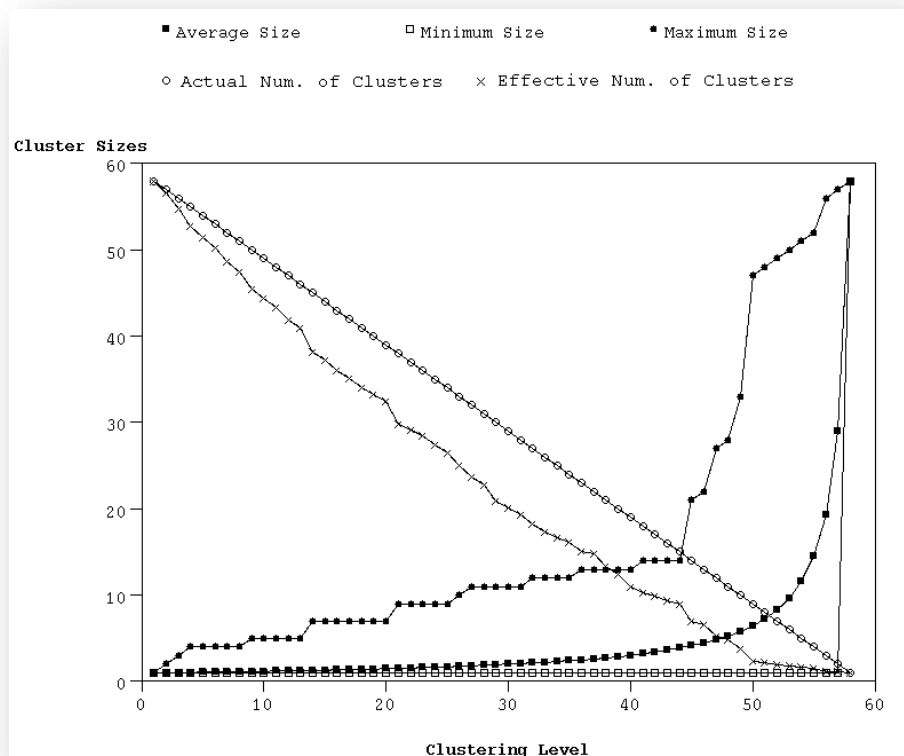


Figure 17: Clustering statistics for a docking run with serotonin

The figure above plots the clustering level against the cluster sizes, each level standing for a pose assigned to a cluster aggregation step. The faster the cluster size increases, the more similar are its members. Several other visualization modes exist that can be applied to judge the clustering behavior depending on the respective requirements. In the results section the clustering value is plotted against the cluster sizes, clearly showing the threshold value for the first cluster to form.

5.3.3 MM-GBSA Rescoring

Although the internal IFDScore seemed to be quite reliable for reasonable ligand ranking, it was desirable to introduce an external ranking tool, which was found in the MM-GBSA rescoring application. This post docking scoring protocol had proved its ability to correctly rank a series of kinase inhibitors, so it was chosen for the experiment. (Lyne et al. 2006) It turned out that the program was designed for rescoring docking results of Schrödinger's Glide application with a single protein input file and a ligand pose library, but it was not able to handle target proteins with slightly changing binding pockets as produced by induced fit docking runs. As the main task of the rescoring procedure was considered the confirmation of the overall IFDScore ligand ranking and not weighting the different ligand poses, the most reasonable way to handle the challenge was to take the best scored pose complex of each ligand and re-divide it into protein and ligand. Stored as separate files the data could be taken as input for the rescoring tool.

The X-Score program obtained from Dr. Renxiao Wang from the University of Michigan was planned to be used as a third, also external, rescoring tool, but failed due to compatibility problems.

5.4 *Validation of the Method*

The fact that the template structure was published with a co-crystallized substrate bound to the binding site offered a good possibility to prove the ability of the method to retrieve valid binding modes, while it has to be kept in mind that due to the nature of crystallizing procedures the states showed by X-ray structures do not necessarily need to represent the actual in-vivo binding mode.

During the first step, the raw pdb file downloaded from the Protein Data Base (PDB) had to be processed for the further use as a docking target structure. This was done by the Protein

Preparation Wizard, capable of adding missing hydrogen atoms, removing co-crystallized solvent or simple water molecules and categorizing missing atom types. The latter is necessary because sometimes some of the residues are not fully recognized due to the fact that during the X-ray structure determination nitrogen and oxygen atoms are just categorized by their position in the particular amino acid, as the electron densities of the two atom types are similar.

After that procedure, the protein complex consisted just of the 516 residue-long chain, two sodium ions and still the bound substrate. This leucine was subject to further deletion, to offer a free binding site for its re-docking. Applying similar settings also used for the actual experiments the program was offered the isolated leucine structure as the ligand file to be docked into the empty LeuT binding site consisting of the residues corresponding to those in the hSERT. Out of initially sixteen valid poses from the primary docking stage, eight successfully passed the sorting and filtering steps to the scoring stage, taking an overall of 6774 seconds equal to two hours calculation time in standard precision mode. Compared with the crystal structure, the best-ranked docking pose showed an overall RMSD of 0.47 Å respectively 0.24 Å in crucial binding areas. (Fig.13) Also the standard clustering procedure included the best-ranked pose in the top cluster, showing that both methods have a reasonable informative value.

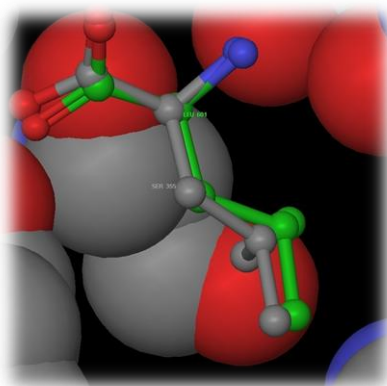


Figure 18: Original and best-ranked docking pose of leucine in LeuT

6. Results

After a series of test runs to determine the optimal calculation settings the main docking runs in standard and extra precision mode were set up. Compared to the redocking procedure for the validation experiment, the time consumption of the main calculations was by far higher, taking 138 hours for the SP and 136 hours in XP mode. Also, on first sight it might seem strange that the more precise mode took less time, but taking into account that the SP docking run had to deal with 493 valid initial docking poses and an overall final output of 126 unique structure files, the 107 final pose files of the XP run resulted out of initially 478 valid poses. That means an overall of 66 minutes per final pose in SP mode compared to 76 minutes in XP mode or an increase of 15% of average calculation time for the more precise experiments. In general, results yielded in XP mode were considered to be more precise, so the results section refers to SP results just in case that further information gain could be achieved by their detailed examination.

	SP	XP
Initial Poses	493	478
Final Poses	126	107
Total Time (min.)	8307,5	8181,1
Time per initial pose	16,9	17,1
Time per final pose	65,9	76,5
Time per overall pose	13,4	14,0

Table 4: Docking time validation

The distribution of poses per ligand was similar for both precision modes, as it can be seen in the following table.

Standard Precision (SP) Mode:		Extra Precision (XP) Mode:	
•Serotonin	$\Sigma = 58$ poses	•Serotonin	$\Sigma = 37$ poses
•Cocain	$\Sigma = 16$ poses	•Cocain	$\Sigma = 12$ poses
•Methylphenidate	$\Sigma = 19$ poses	•Methylphenidate	$\Sigma = 23$ poses
•MDMA	$\Sigma = 26$ poses	•MDMA	$\Sigma = 28$ poses
• β -CFT	$\Sigma = 1$ pose	• β -CFT	$\Sigma = 2$ poses
• β -CIT	$\Sigma = 1$ pose	• β -CIT	$\Sigma = 2$ poses
•Benzatropine	$\Sigma = 5$ poses	•Benzatropine	$\Sigma = 3$ poses

Table 5: Pose numbers

As the score of each pose was broken down into several energy contribution values, during the further steps a large amount of data had to be processed. By default, the primary ranking was not according to the ligand type but to the IFD Score value. Henceforward the ligands had to be moved into separate groups to be made comparable for the clustering tools. For the selection of the favored poses we decided to combine different criteria to be met. The best-ranked complex for each ligand was automatically taken as a result due to the significance of the parameter already proven during the redocking process. In addition, the in-place clustering behavior at dedicated RMSD thresholds of 1.5, 2.0 and 2.5 Å was examined in case that the primary population was large enough to suggest possible information gain.

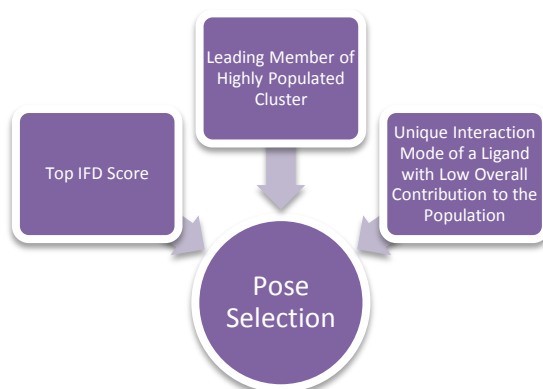


Figure 19: Pose Selection Parameters

The resulting files created out of the best-ranked pose for each ligand were also re-divided into separate protein and ligand files for the MM-GBSA rescoring step, leading to the following primary ligand ranking.

Compound	Acronym	IFD Score	MMGBSA L/R ligand energy	MMGBSA Complex Energy
Serotonin	5-HT	-855.06	-65.74	-16.91
Methylenedioxymethamphetamine	MDMA	-854.40	-11.54	-16.86
Methylphenidate	MPD	-850.22	-9.31	-16.83
Benzatropine	BTP	-845.38	37.95	-16.76
Cocaine	COC	-844.84	27.83	-16.81
2 β -Carbomethoxy-3 β -(4-fluorophenyl)tropane	CFT	-841.62	19.84	-16.79
2 β -Carbomethoxy-3 β -(4-iodophenyl)tropane	CIT	-836.80	19.02	-16.78

Table 6: IFDScores correlated to MM/GBSA Energies, $r^2=0.71$, without BTP: $r^2= 0.86$

6.1 Serotonin

The natural ligand did not just produce the most final poses, it also achieved the best overall scores. The 37 XP poses seemed perfect for applying the different pose selection tools, as they were structurally diverse but yet able to form clusters at the chosen RMSD thresholds.

For the clustering level corresponding to the 1.5 Å thresholds still six clusters with more than one member were calculated, but going up to 2.0 and 2.5 Å the poses merged to four and finally two clusters, each with the best-ranked pose as the leading pose in the highest populated cluster.

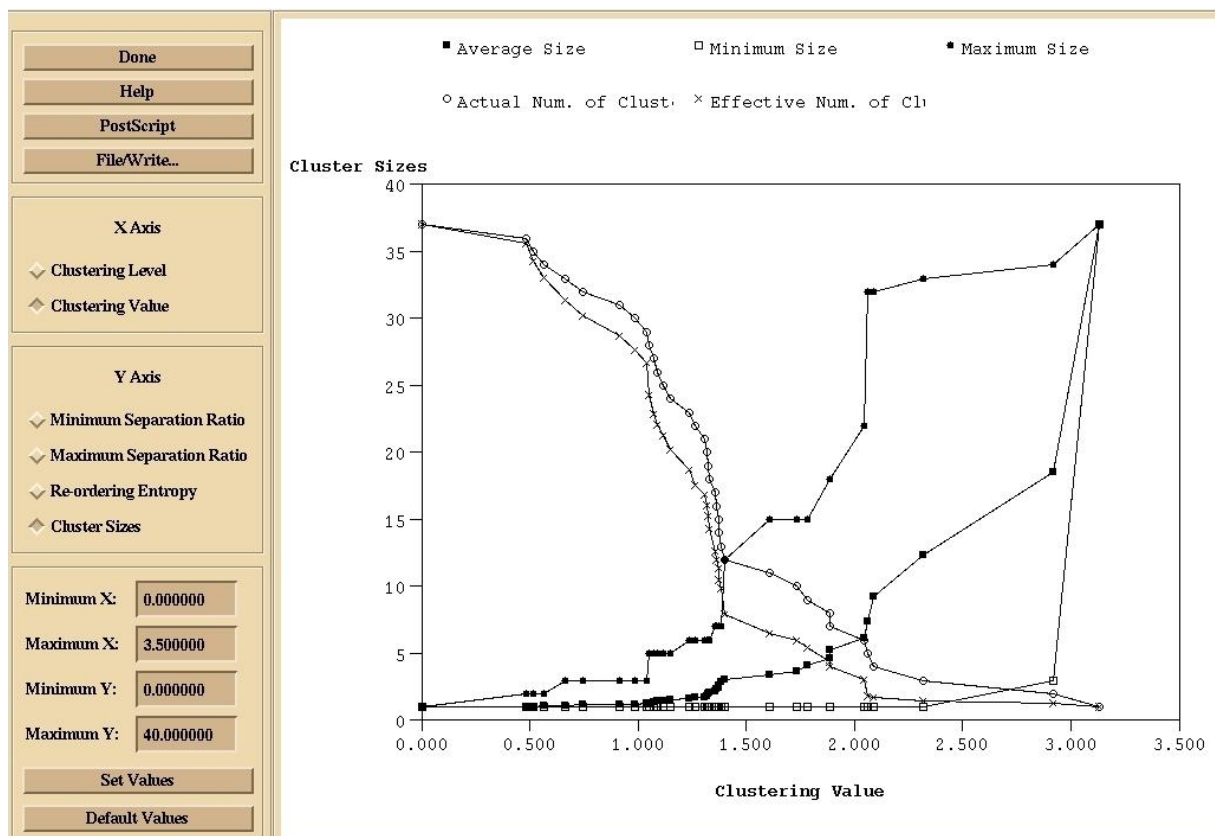


Figure 20: Clustering Statistics for Serotonin

Exhaustive investigation on the protein-ligand interactions of the best-ranked pose showed an interaction pattern similar already published data. (Celik et al. 2008)

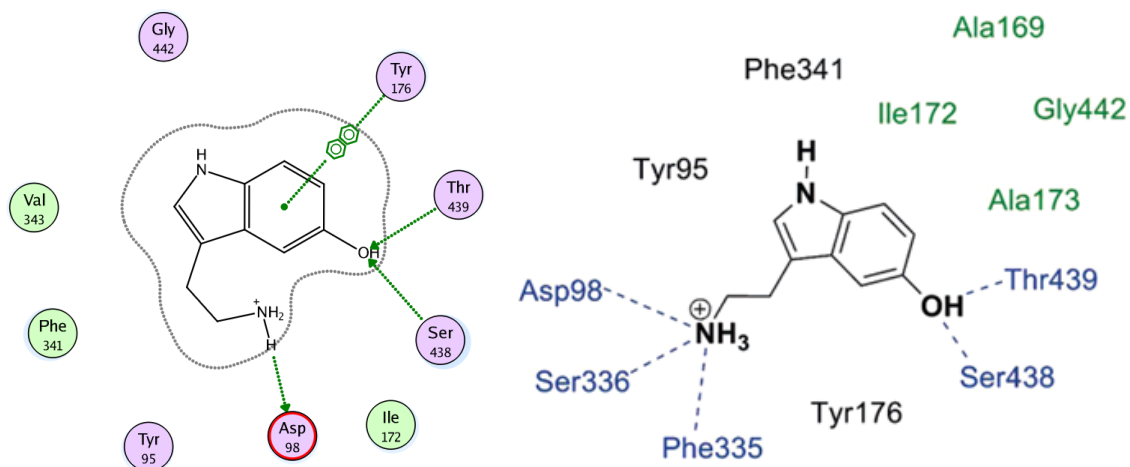


Figure 21: 5-HT interactions of the best-ranked pose and pose taken from Celik et al., 2007

A common feature of most ligands binding to the SERT is the salt bridge between the positively charged nitrogen and the aspartic acid Asp98. As a second anchor point serves the hydroxy group in position 5 of the ligand, respectively the side chains of Ser438 and Thr439, joined by hydrogen bonding. The pi-pi interaction of Tyr 176 with the phenyl ring of serotonin questionable, as only molecular dynamics simulations would be able to show the stability of this interaction

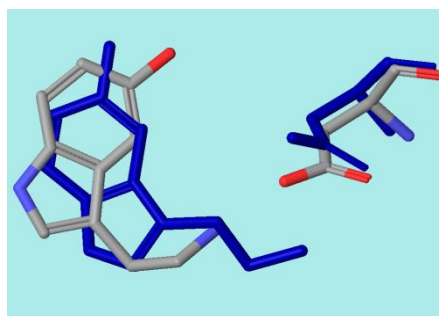


Figure 22: Comparison of MOE and Schrödinger output leading poses

Also the comparison with the leading pose yielded in earlier rigid docking experiments with the MOE program package (dark blue) showed a similar orientation.

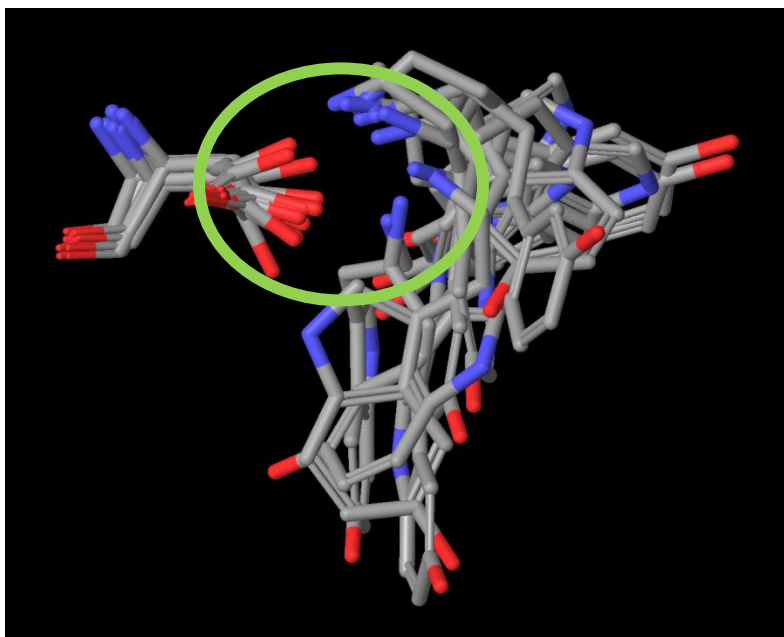


Figure 23: Orientation of Serotonin Poses around Asp98

6.2 MDMA

The internal scoring application ranked several MDMA poses higher than less favorable serotonin results, consistent to published affinity data. (Rothman & Baumann 2003) Taking a look on the clustering behavior of the MDMA poses, two main clusters became apparent with the two best-scored poses as the respective leading members. The figure on the right shows the output data for the clustering levels corresponding to the 2.0 and 2.5 Å thresholds. Looking at an overlay of the Cluster 1 members it was remarkable that they arranged themselves like snapshots of a movement through the final part of the ligand channel down into the final position inside the binding site, naturally within the comparably narrow limits of the defined receptor grid.

```
Starting command thresh 26
Clustering 26; threshold distance 2.4130697; 3 cluster(s):
Cluster 1; Leading member= 1; 8 members, sep_rat 1.0725:
 1 25 22 8 17 26 24 27
Cluster 2; Leading member= 2; 19 members, sep_rat 1.0725:
```

Figure 24: Leading Position of Poses 1 and 2

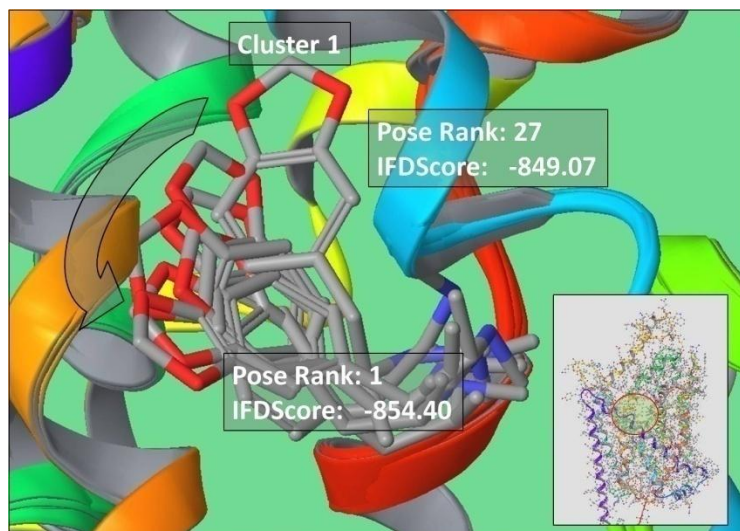


Figure 25: Superposition of MDMA Cluster 1

Taking the best ranked pose, at the same time the leading pose of cluster 1, the main binding interactions with the protein are the previously mentioned salt bridge with the aspartic acid 98 and a backbone interaction with Phe 335, supplemented by a good steric fit.

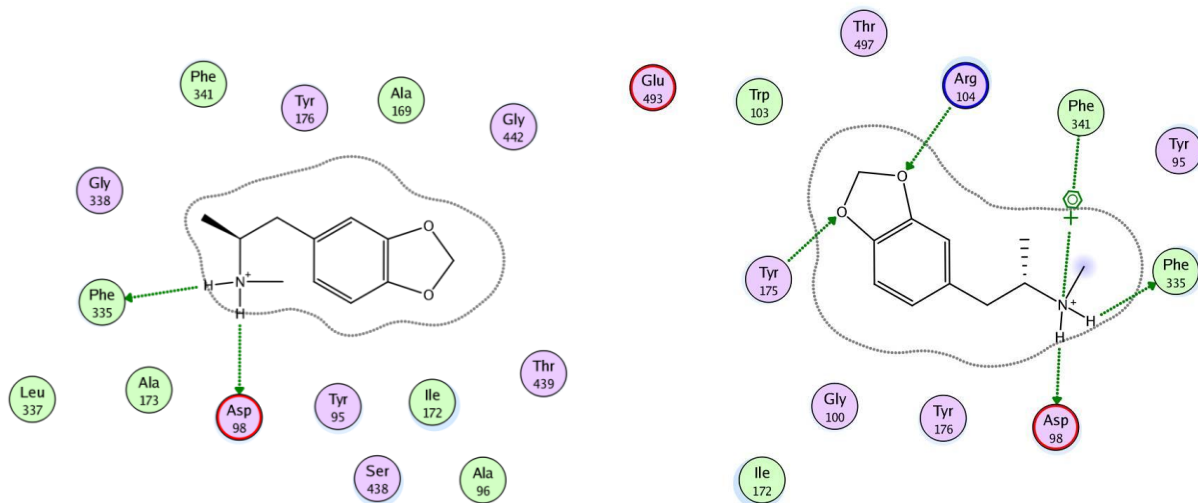


Figure 26: Interactions of the two best-ranked MDMA poses

The second pose cluster is turned by an angle of about 90° compared to the first one, sharing the positively charged nitrogen as the center of rotation. This causes the common side chain interactions with D98 and F335. Additional features can be seen regarding the two ether oxygen atoms with R104 respectively Y175 and an arene-cation interaction of F341 with once again the positive charge of the ligand nitrogen.

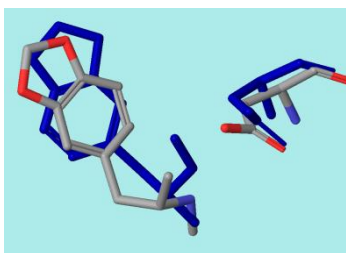


Figure 27: Superposition of pose 1 with rigid docking leading pose

Comparing the top pose of cluster 1 to the leading pose of the rigid docking experiments, again a similar binding mode could be observed. For better orientation, Asp98 is also being displayed in the figure above.

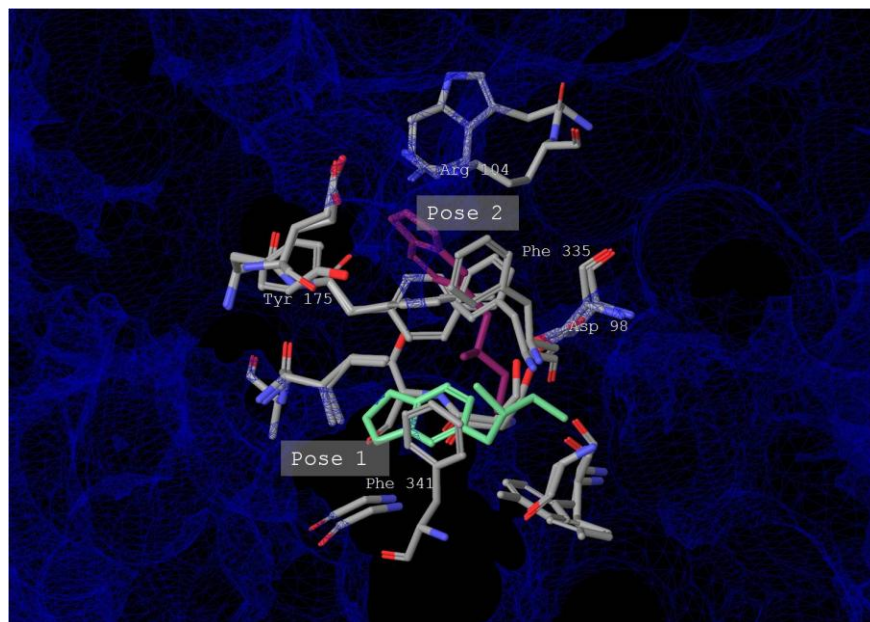


Figure 28: Position of the Leading Poses in the Binding Site

As it can be seen in the distance map of the XCluster module, the first cluster was formed at a 0.46 Å threshold, similar to serotonin. Nevertheless, the clear differentiation of the MDMA poses into two large clusters within a narrow distance range stands in contrast to the comparably late main cluster formation of serotonin, as the high structural diversity of serotonin poses is reflected in the comparison of the two distance maps.

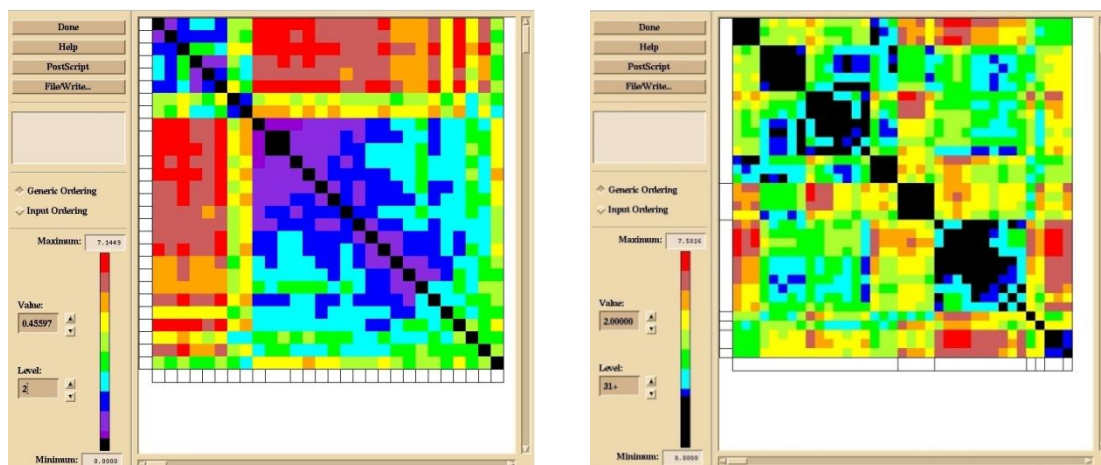


Figure 29: MDMA (left) and Serotonin (right) pose distance maps

6.3 Methylphenidate

This ligand takes an exceptional position within the tested substances. As in some studies no in-vitro binding to SERT could be detected for this drug (Markowitz et al. 2006), others show at least about 2000-fold reduced inhibition potency of methylphenidate compared to DAT or NET. (Han & Gu 2006) However well-ranked docking poses were generated. Therefore, these structures were considered artifacts and might support the theory that ligand specificity is at least partially accomplished by regions outside the binding site as for instance the substrate channel. Further considerations about the kinetics of binding were suspended to later on studies, carrying forward the previous workflow. Once again, the clustering visualization tools showed the clearly favored orientation in the binding site dictated by the best-ranked pose as leading member of cluster 1, at which 13 of 23 poses belong to the highest populated cluster at the 2.5 Å threshold. The second largest cluster with four members led by pose 9 formed at the 2.0 Å threshold. As it can be seen in the figure below, the large cluster formed quite below the first critical value of 1.5 Å.

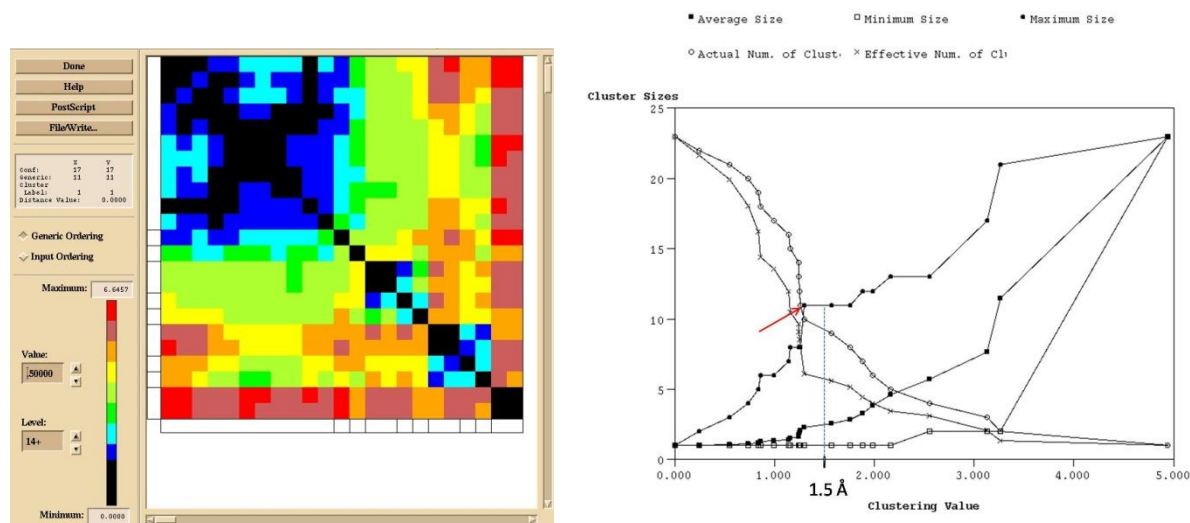


Figure 30: Formation of Methylphenidate Cluster 1

Comparing the two leading members with those of the top MDMA clusters, the similar orientation of the ligands was quite remarkable.

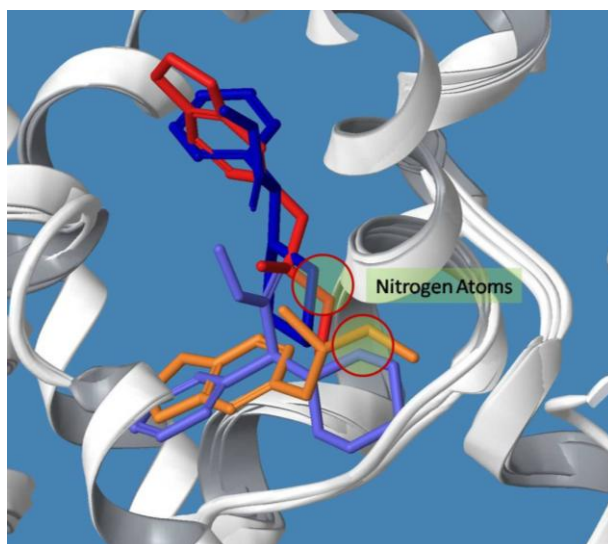


Figure 31: Overlay of leading MDMA and MPD poses

Looking at the best-ranked methylphenidate pose, an interaction pattern similar to the previous compounds arises. Along with the ionic bond affecting D98 and its nitrogen counterpart, again a backbone interaction with F335 and an additional π - π interaction with Y176 could be found.

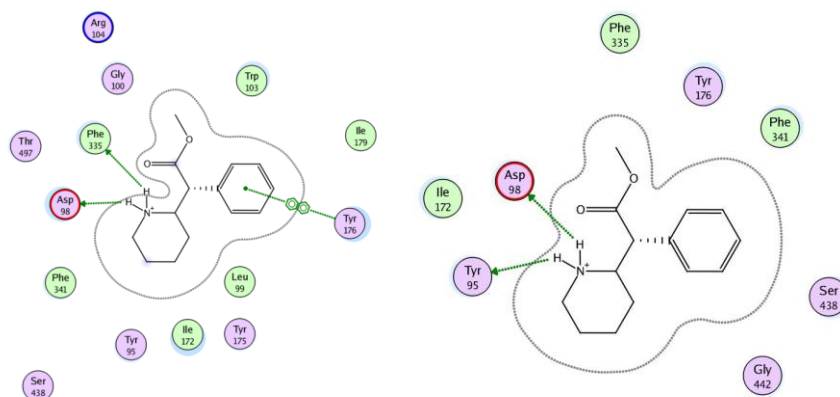


Figure 32: Best calculated methylphenidate interactions

Leading member of cluster 2, the second chosen methylphenidate structure, showed the well-known interaction with D98 and another side chain interaction with Y95.

6.4 *Benzatropine*

The first member of the docked tropane derivatives yielded just three valid poses, two of them sharing the same orientation with an RMSD value of just 1.25 Å.

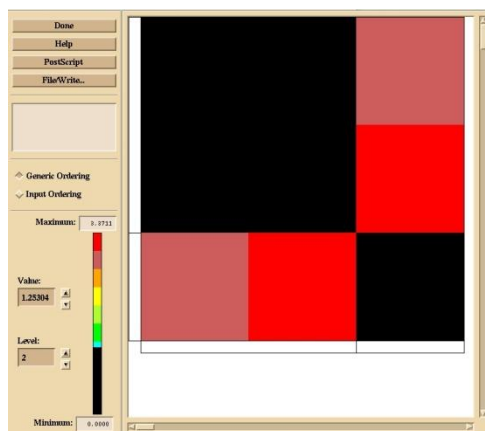


Figure 33: Benzatropine distance map

The IFD-Scores were relatively poor, taking the ranks 90 to 92 out of 107 poses. As it is mentioned later in the cocaine section, this can be subject to the relatively narrow borders of the binding site that were defined according to the corresponding substrate binding site in the LeuT. A recently published paper, though being concerned with the DAT, strongly suggests a tropane binding pocket slightly shifted to the extracellular part of the protein. (Huang et al. 2009)

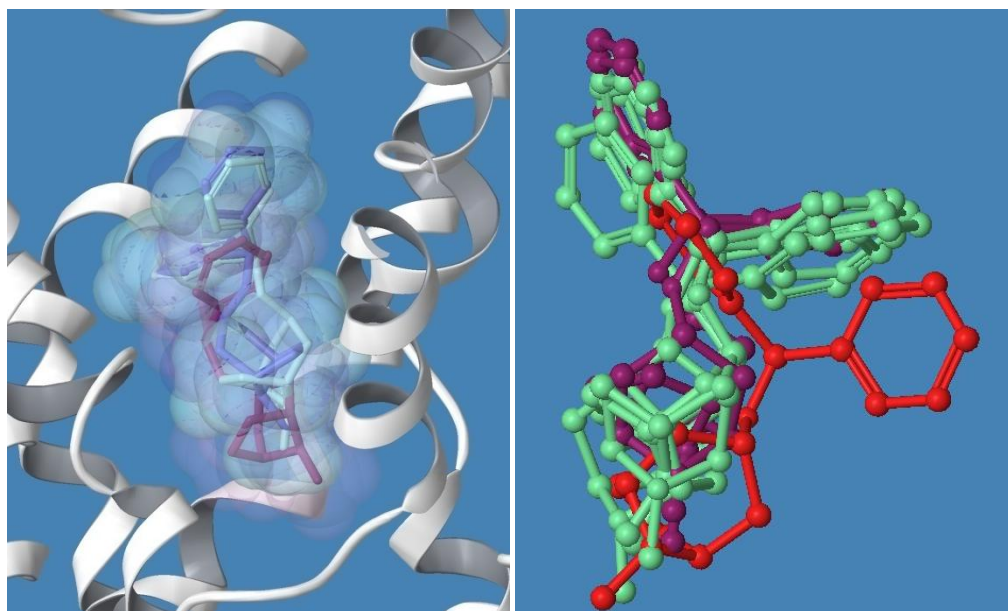


Figure 34: Superposed Benzatropine Poses (Discarded Pose in Red)

As the medium-ranked separate pose did not share its orientation with any of the other main tropane derivatives the population of Benzatropine poses was increased by the SP output data, adding five Benzatropine structures for further investigation. Within the SP population, the BTP poses were even worse ranked, but all of them shared the main XP orientation, leading to the assumption that the second XP pose could be considered an outlier.

The main conformation showed the central D98-interaction and a π - π interaction with W103. Interesting about the latter is the fact that the tryptophan at this position is conserved within the main eucaryotic monoamine transporters that usually have binding affinities in the same orders of magnitudes for tropane-like transporter inhibitors.

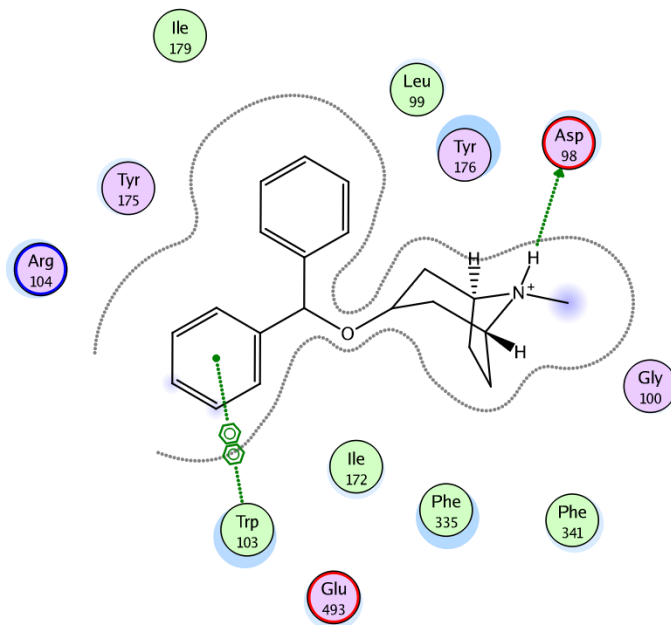


Figure 35: Best-ranked Benzatropine interactions with the hSERT

During the MM-GBSA rescoring step Benzatropine was the only ligand with an overall ranking not consistent with the IFD Score results. Though the leading pose was the best ranked tropane structure according to its IFD score value, it obtained the lowest MM-GBSA score within all tested ligands, yet another reason for further experiments with an expanded binding site.

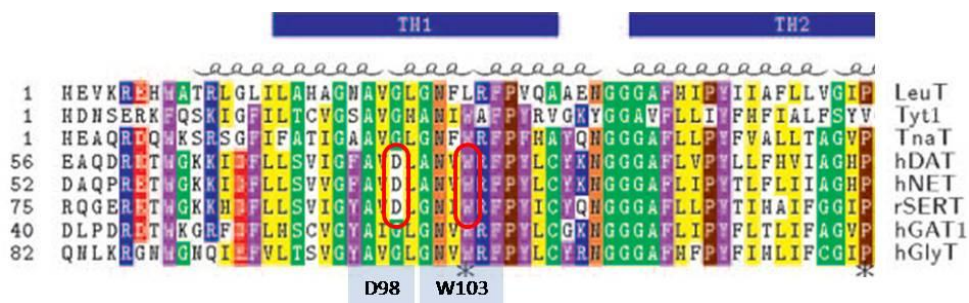


Figure 36: Residues Involved in Benzatropine Binding

6.5 Cocaine

The total amount of six poses with totally different interaction patterns in the output of this ligand made a clustering step useless. In the figure below the overlay of the poses shows the different orientations of the molecule, highlighted by the special representation of the nitrogen atoms.

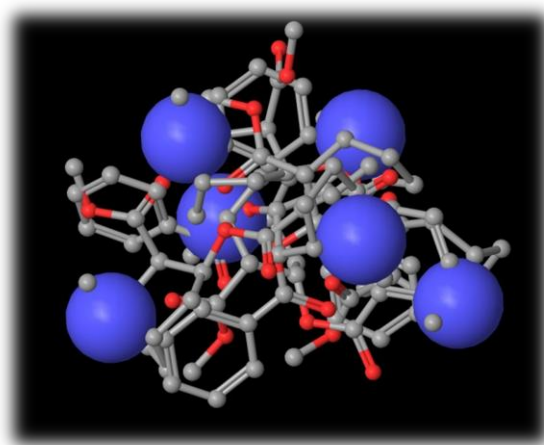


Figure 37: Superposed cocaine poses

To be consistent with the general workflow the clustering module was applied anyway, also to prove the ability to detect poor clustering behavior of a population. To do so, the clustering statistics were examined.

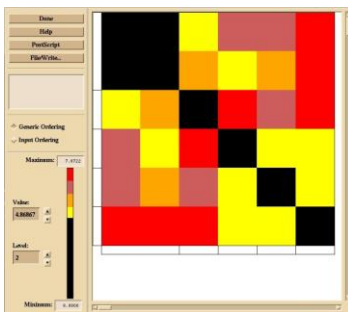


Figure 38: Distance Map of the Cocaine Poses

The figure above shows the distance map for cocaine at a clustering level of 2, revealing a threshold distance of 4.9 Å necessary to form the first cluster.

In addition to the scoring value it seemed reasonable to consider the position of the subsequently discussed cocaine analogues that yielded even less poses, apparently due to their bulky halogen moiety. An overlay of the cocaine poses with the results for β -CFT showed that the binding mode of the best-scored cocaine pose was consistent with the cocaine analogue.

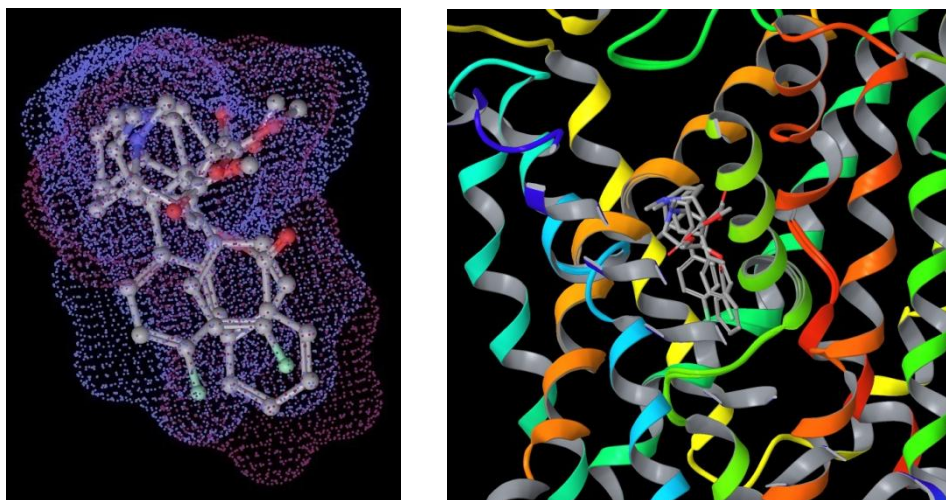


Figure 39: Overlay of cocaine and β -CFT

As the comparably few interactions of the leading cocaine pose consisting of just a π - π interaction with F341 and the steric fit inside the pocket are also not reproducible within the other poses the result is quite questionable. A paper recently published discussed the interactions of cocaine and its analogues with the dopamine transporter, where it was shown that an interaction of the positively charged nitrogen with D79 in DAT, the conserved amino acid corresponding to D98 in the SERT, is very likely to exist. F326, the residue corresponding to F341, is mentioned in the paper, but not as an essential residue. (Beuming et al. 2008) The recent work of Huang et al. even proposes that the binding site of cocaine and its analogues is about 11 Å above the current one, though the paper also just refers to the DAT.

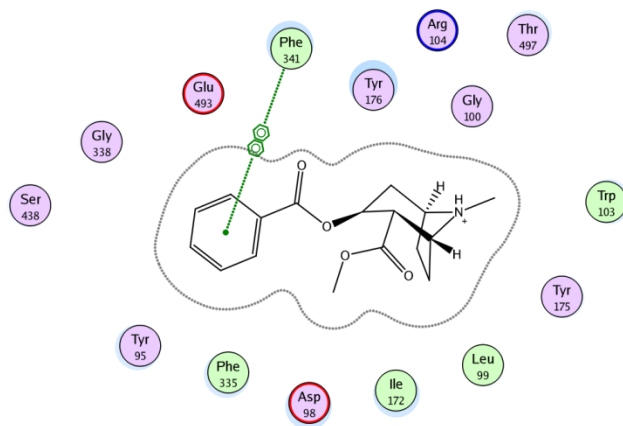


Figure 40: Proposed cocaine interactions

6.6 β -CFT

This halogen-substituted cocaine analogue with highly increased potency produced just two valid poses quite similar to each other in the extra precision mode. Additional data from the SP run was consistent with the XP poses with an RMSD of below 2 Å.

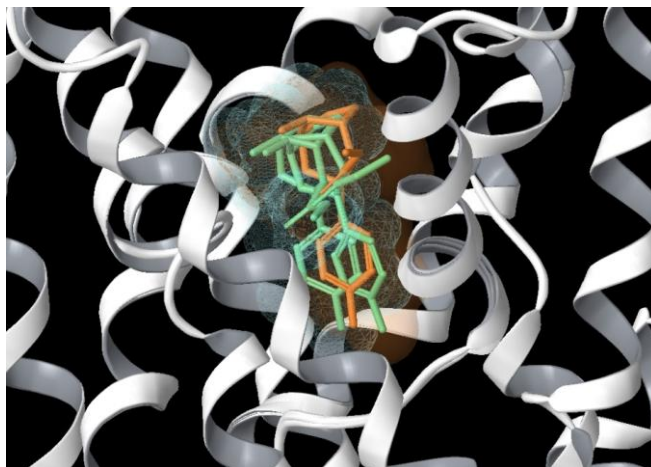


Figure 41: CFT Poses of the XP Run and Additional SP CFT Pose (Orange) in Superposition

The rank of the ligand at the lower end of the pose population both in IFD and MM-GBSA score was not very surprising at first sight, as a certain preference for DAT binding has been observed for the compound. Although an earlier in-vitro assay in rat tissue found the compound to be an even more potent SERT uptake inhibitor than cocaine with IC₅₀ values of 91.0 nM vs. 260 nM (Kuhar et al. 1999), recent studies in human embryonic kidney cells (HEK 293) showed that CFT has –apart from its unquestioned higher DAT potency- a SERT binding and transport inhibition IC₅₀ of 2300 nM (±0.3), therefore about 3-4 fold higher than cocaine with an IC₅₀ of about 580 nM (±110), justifying the CFT scores being lower than the cocaine values. (Riss et al. 2009)

Investigations on the leading pose showed just a π - π interaction with Y176. Earlier binding studies for this substance on the DAT would lead to the assumption that D98 is involved in the binding, but the 6.3 Å distance between the negatively charged residue and the quaternary nitrogen atom makes this interaction highly unlikely.

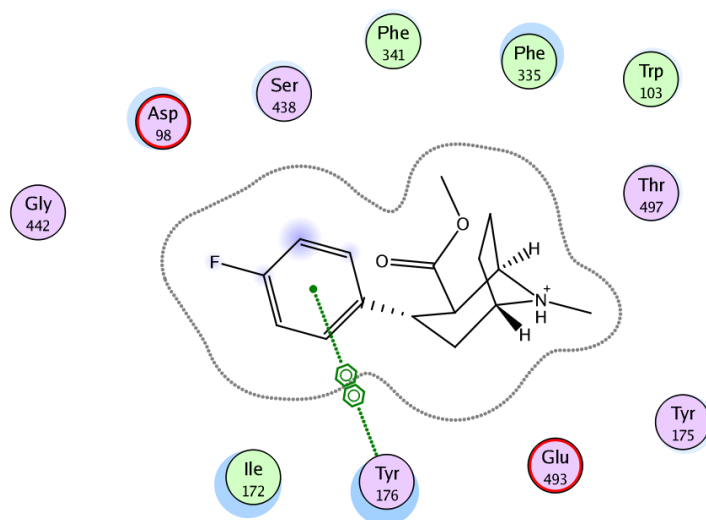


Figure 42: Computationally preferred β -CFT binding pattern

6.7 β -CIT

Again, the yield of poses with a total of two output structures was quite poor. Even looking at the results from the SP mode just one additional structure was available, not only the worst ranked of all poses but also similar to one pose already available in the XP data, so its consideration was needless. The overall orientation of the primary pose was similar to the one of β -CFT, but it did no longer show the π stacking with Y176 known from the previous ligand.

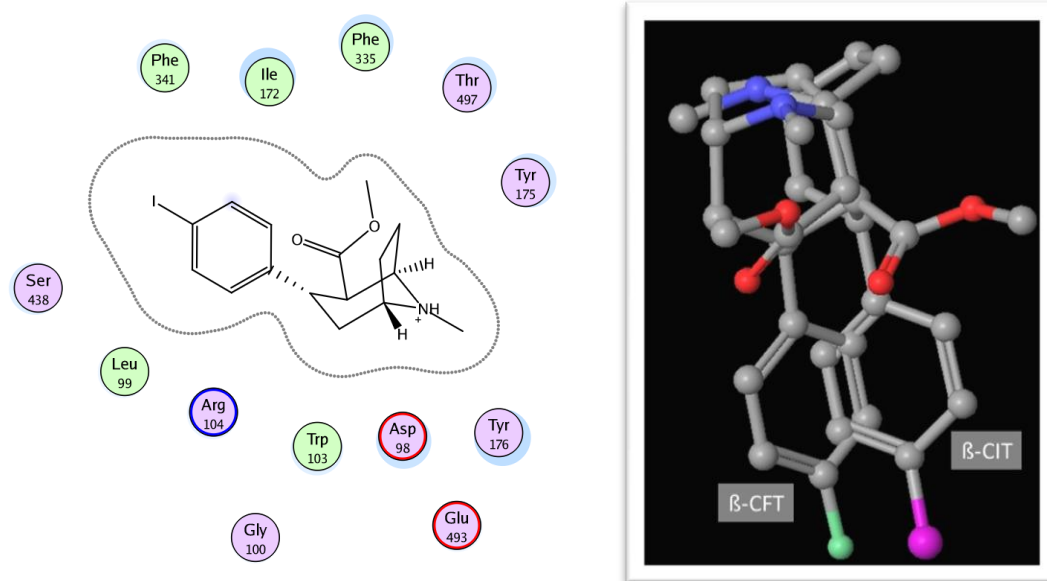


Figure 43: β -CIT in the binding pocket and in superposition with β -CFT

Taking a look at the second calculated binding mode, the interaction pattern looks different, as the ligand is nearly 180° turned in the pocket, and now shows not only the ionic interaction with D98, but also a backbone interaction with F335 and an arene-cation bond between F341 and the positively charged nitrogen, a pattern quite familiar from non-tropane ligands.

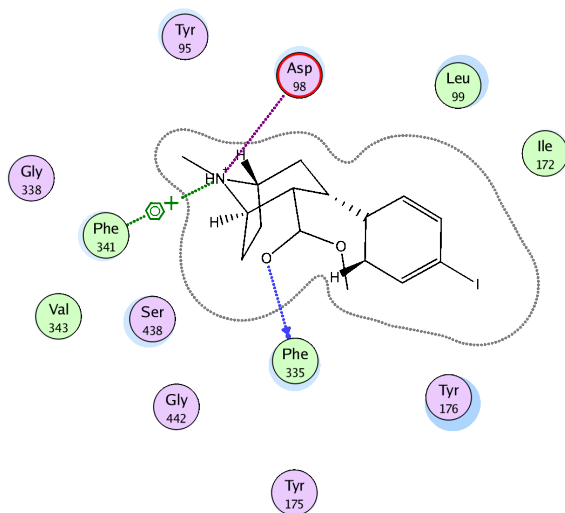


Figure 44: Secondary binding mode of β -CIT

The reason for that prima facie better pose to be less favorable for the scoring function lies in several energy contributions, as it can be seen in a variety of scores below.

Title	glide ligand efficiency	glide hbond	glide ligand efficiency	Prime vdW	Prime Solv GB	Prime Energy	IFDScore
CIT pose 1	-1.152.545	-0.095891	-0.424601	-1.255.585.456	-1.887.210.392	-16.566.098.229	-836.796.934
CIT pose 2	-1.112.714	-0.276251	-0.409927	-1.251.090.835	-1.853.241.613	-16.561.239.603	-836.260.528

Figure 45: Different energy contributions for CIT poses

Anyway, the differences in the IFD score were minimal with -836.8 vs. -836.3.

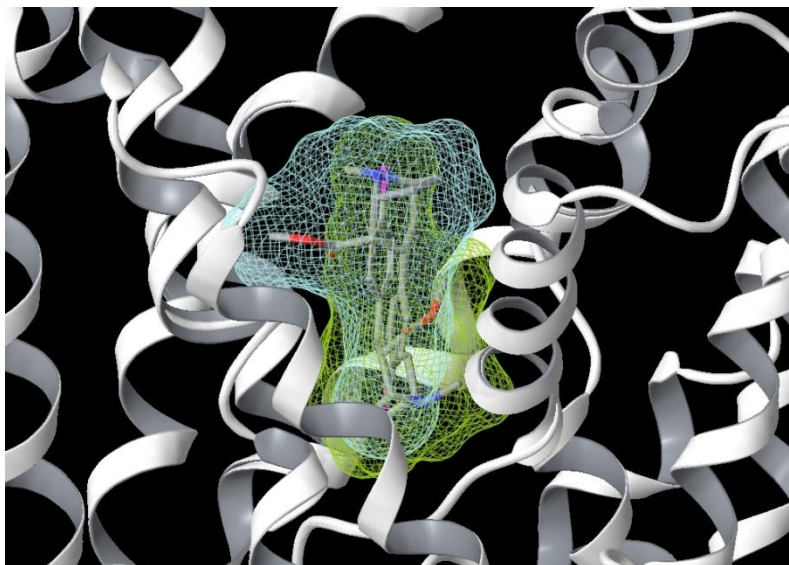


Figure 46: Antipodal Orientation of β -CFT Poses

Although the only difference to the previous compound is the fluorine substituted by iodine, the change in the experimental values is huge. Still being predominantly DAT selective ($IC_{50} = 6.34 \pm 1.68$ nM), the SERT potency ($IC_{50} = 29.17 \pm 6.40$ nM) is a lot higher than the one of cocaine ($IC_{50} = 580.73 \pm 113.16$ nM). (Okada et al. 1998) Once again it is tempting to doubt either the method or the positioning of the binding site, as the calculated data is not quite consistent with the experimental data. In any case further experiments with an expanded binding site seem inevitable.

7. Discussion and Outlook

Although we are still bound to the limitations of a model, the method showed to be quite accurate and fast applicable. Compared to earlier experiments keeping the protein backbone rigid the low number of output poses was surprising. An overlay of the poses yielded showed that the theoretical framework of induced fit backbone movements was not exploited; anyway, subsequent global energy minimization steps could be omitted due to the program architecture performing in-process minimization. Structures resulting from rigid-docking experiments in our group were available for five of the seven ligands. Comparing the final poses with those yielded in the previous work, the central interactions could be observed for most ligands except the tropane derivatives. As mentioned in the respective results section, this fact can be due to a possible shifted binding site for this group of ligands or actually a unique binding mode, calling for further investigation. The correlation of scoring values to in-vitro affinity data was especially challenging due to inhomogeneous cell systems. It is not surprising that rat tissue yields different data compared to, for instance, HEK cells, but several ligands simply could not be measured in identical test systems. As affinity fluctuations added up to several orders of magnitude for a single ligand, depending on the cell system, correlating scoring data to IC_{50} values was abandoned. Questions emerging out of these results will need further investigation both in careful watch of newly released literature and furthermore continuative experiments. Shifted or expanded binding sites will play an important role as well as the use of additional software modules and further model refinement. Maybe the largest information gain will result out of experiments using molecular dynamics (MD) simulations, promising to identify stable ligand orientations, but yet limited due to the immense use of computational power needed for MD calculations. Definitely, docking algorithms keeping the binding site flexible stand for an important improvement in the method per se, yielding structures that can be used as good binding mode suggestions for new ligands and as starting structures for MD simulations. The Induced Fit protocol led to a significantly lower number of output poses, yet keeping behind the expectations in actual side chain and backbone movements. As far as it can be judged by hitherto yielded data, the results seem reliable. Still the data processing is difficult due to the

fact that many applications can not deal with a series of ligand-protein complexes with changing protein backbone. However, as it becomes more and more common to keep both ligand and protein flexible it will be just a matter of time until most software modules can handle Induced Fit output data.

References

Aktories, K. et al., 2005. *Allgemeine und spezielle Pharmakologie und Toxikologie*, Elsevier GmbH.

Aoyama, T. et al., 1996. Pharmacokinetics and pharmacodynamics of methylphenidate enantiomers in rats. *Psychopharmacology*, 127(2), 117-122.

Beuming, T. et al., 2008. The binding sites for cocaine and dopamine in the dopamine transporter overlap. *Nature Neuroscience*, 11(7), 780-789.

Beuming, T. et al., 2006. A comprehensive structure-based alignment of prokaryotic and eukaryotic neurotransmitter/Na⁺ symporters (NSS) aids in the use of the LeuT structure to probe NSS structure and function. *Molecular Pharmacology*, 70(5), 1630-1642.

Boja, J.W. et al., 1992. High potency cocaine analogs: neurochemical, imaging, and behavioral studies. *Annals of the New York Academy of Sciences*, 654, 282-291.

Brzezinski, M.R. et al., 1994. Purification and characterization of a human liver cocaine carboxylesterase that catalyzes the production of benzoylecgonine and the formation of cocaethylene from alcohol and cocaine. *Biochemical Pharmacology*, 48(9), 1747-1755.

Celik, L. et al., 2008. Binding of serotonin to the human serotonin transporter. Molecular modeling and experimental validation. *Journal of the American Chemical Society*, 130(12), 3853-3865.

Charifson, P.S. et al., 1999. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *Journal of Medicinal Chemistry*, 42(25), 5100-5109.

Ding, Y. et al., 2004. Brain kinetics of methylphenidate (Ritalin) enantiomers after oral administration. *Synapse (New York, N.Y.)*, 53(3), 168-175.

Gainetdinov, R.R. & Caron, M.G., 2001. Genetics of childhood disorders: XXIV. ADHD, part 8: hyperdopaminergic mice as an animal model of ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40(3), 380-382.

Gouaux, E., 2009. Review. The molecular logic of sodium-coupled neurotransmitter transporters. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1514), 149-154.

Han, D.D. & Gu, H.H., 2006. Comparison of the monoamine transporters from human and mouse in their sensitivities to psychostimulant drugs. *BMC Pharmacology*, 6, 6.

Höltje, H. & Folkers, G., 1996. *Molecular Modeling - Basic Principles and Applications*, VCH Publishers Inc., New York NY (USA).

Huang, X., Gu, H.H. & Zhan, C., 2009. Mechanism for Cocaine Blocking the Transport of Dopamine: Insights from Molecular Modeling and Dynamics Simulations. *The Journal of Physical Chemistry. B*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19831380> [Accessed October 22, 2009].

Humphreys, C.J., Wall, S.C. & Rudnick, G., 1994. Ligand binding to the serotonin transporter: equilibria, kinetics, and ion dependence. *Biochemistry*, 33(31), 9118-9125.

Kovacic, P., 2005. Role of oxidative metabolites of cocaine in toxicity and addiction: oxidative stress and electron transfer. *Medical Hypotheses*, 64(2), 350-356.

Kuhar, M.J. et al., 1999. Studies of selected phenyltropanes at monoamine transporters. *Drug and Alcohol Dependence*, 56(1), 9-15.

Lyles, J. & Cadet, J.L., 2003. Methylenedioxymethamphetamine (MDMA, Ecstasy) neurotoxicity: cellular and molecular mechanisms. *Brain Research Reviews*, 42(2), 155-168.

Lyne, P.D., Lamb, M.L. & Saeh, J.C., 2006. Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *Journal of Medicinal Chemistry*, 49(16), 4805-4808.

- Markowitz, J.S. et al., 2006. A comprehensive in vitro screening of d-, l-, and dl-threo-methylphenidate: an exploratory study. *Journal of Child and Adolescent Psychopharmacology*, 16(6), 687-698.
- Mueller, M.A. et al., 2009. Further Studies on the Role of Metabolites in MDMA-induced Serotonergic Neurotoxicity. *Drug Metabolism and Disposition: The Biological Fate of Chemicals*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19628751> [Accessed August 13, 2009].
- O'Boyle, N.M., Liebeschuetz, J.W. & Cole, J.C., 2009. Testing Assumptions and Hypotheses for Rescoring Success in Protein-Ligand Docking. *Journal of Chemical Information and Modeling*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19645429> [Accessed September 28, 2009].
- Okada, T. et al., 1998. Assessment of affinities of beta-CIT, beta-CIT-FE, and beta-CIT-FP for monoamine transporters permanently expressed in cell lines. *Nuclear Medicine and Biology*, 25(1), 53-58.
- Oshiro, C. et al., 2004. Performance of 3D-database molecular docking studies into homology models. *Journal of Medicinal Chemistry*, 47(3), 764-767.
- Pellinen, P. et al., 2000. Kinetic characteristics of norcocaine N-hydroxylation in mouse and human liver microsomes: involvement of CYP enzymes. *Archives of Toxicology*, 74(9), 511-520.

Riss, P.J., Hummerich, R. & Schloss, P., 2009. Synthesis and monoamine uptake inhibition of conformationally constrained 2beta-carbomethoxy-3beta-phenyl tropanes. *Organic & Biomolecular Chemistry*, 7(13), 2688-2698.

Rothman, R.B. & Baumann, M.H., 2003. Monoamine transporters and psychostimulant drugs. *European Journal of Pharmacology*, 479(1-3), 23-40.

Rothman, R.B. et al., 2008. Dopamine transport inhibitors based on GBR12909 and benztropine as potential medications to treat cocaine addiction. *Biochemical Pharmacology*, 75(1), 2-16.

Rudnick, G., 2006. Serotonin transporters--structure and function. *The Journal of Membrane Biology*, 213(2), 101-110.

Torres, G.E., Gainetdinov, R.R. & Caron, M.G., 2003. Plasma membrane monoamine transporters: structure, regulation and function. *Nature Reviews. Neuroscience*, 4(1), 13-25.

Volkow, N.D. et al., 1999. Methylphenidate and cocaine have a similar in vivo potency to block dopamine transporters in the human brain. *Life Sciences*, 65(1), PL7-12.

Weed, M.R. et al., 1995. Reinforcing and discriminative stimulus effects of [beta]-CIT in rhesus monkeys. *Pharmacology Biochemistry and Behavior*, 51(4), 953-956.

White, S.R. et al., 1996. The effects of methylenedioxymethamphetamine (MDMA, "Ecstasy") on monoaminergic neurotransmission in the central nervous system. *Progress in Neurobiology*, 49(5), 455-479.

Yamashita, A. et al., 2005. Crystal structure of a bacterial homologue of Na⁺/Cl⁻-dependent neurotransmitter transporters. *Nature*, 437(7056), 215-223.

Zhou, Z. et al., 2007. LeuT-desipramine structure reveals how antidepressants block neurotransmitter reuptake. *Science (New York, N.Y.)*, 317(5843), 1390-1393.

List of Abbreviations

ADHD	Attention-deficit hyperactivity disorder
ATP	Adenosine triphosphate
BTP	Benzatropine
CAS	Chemical Abstracts Service
β -CFT	(-)-2 β -Carbomethoxy-3 β -(4-fluorophenyl)tropane
β -CIT	(-)-2 β -Carbomethoxy-3 β -(4-iodophenyl)tropane
COC	Cocaine
CYP	Cytochrome P450
DAT	Dopamine active transporter
L-DOPA	L-3,4-dihydroxyphenylalanine
GABA	γ -AminoButyric Acid
GUI	Graphical user interface
HEK	Human embryonic kidney
HHMA	3,4-Dihydroxymethamphetamine
5-HIAA	5-Hydroxyindoleacetic
HMMA	4-hydroxy-3-methoxymethamphetamine
5-HT	5-Hydroxytryptamine
IFD	Induced fit docking
LeuT	Leucine transporter

MAO	Monoamine oxidase
MD	Molecular dynamics
MDA	3,4-Methylenedioxyamphetamine
MDMA	3,4-Methylenedioxymethamphetamine
MM/GBSA	Molecular mechanics/generalized Born model, solvent accessibility
MPD	Methylphenidate
NET	Norepinephrine Transporter
NMR	Nuclear magnetic resonance
NSS	Neurotransmitter: sodium symporters
OPLS	Optimized potentials for liquid simulations
PDB	Protein Data Base
RMSD	Root mean square deviation
SERT	Serotonin transporter
SLC	Solute Carrier
SP	Standard precision
XP	Extra precision

Table of Figures

Figure 1: Synaptic signal transduction, taken from Aktories et al., 2005.....	- 4 -
Figure 2: Scheme of monoamine-related synaptic terminals, modified from Torres, 2003.....	- 5 -
Figure 3: Serotonergic (green), dopaminergic (blue) and noradrenalinergic (red) nerve tracts, taken from Aktories et al., 2005	- 6 -
Figure 4: Possible mechanism of serotonin transport, taken from Rudnick, 2006	- 7 -
Figure 5: Amino acids as neurotransmitter precursors	- 8 -
Figure 6: Ligands used for the study.....	- 9 -
Figure 7: Serotonin pathway.....	- 11 -
Figure 8: Comparison of Amphetamine and MDMA	- 12 -
Figure 9: MDMA metabolites, taken from Mueller et al., 2009	- 12 -
Figure 10: Amphetamine - methylphenidate relationship.....	- 13 -
Figure 11: Checklist for the quality of protein models; taken from Höltje and Folkers, 1996 .	- 20 -
Figure 12: Sequence alignment of LeuT with eucaryotic transporters, Beuming et al., 2006 .	- 27 -
Figure 13: Comparison of EL 1 in 2A65 (red) and 2QB4 (blue) based models.....	- 29 -
Figure 14: Inner binding site	- 30 -
Figure 15: RMSD values of different SERT models	- 31 -
Figure 16: General workflow	- 32 -
Figure 17: Clustering statistics for a docking run with serotonin.....	- 34 -
Figure 18: Original and best-ranked docking pose of leucine in LeuT.....	- 36 -
Figure 19: Pose Selection Parameters	- 38 -
Figure 20: Clustering Statistics for Serotonin	- 40 -
Figure 21: 5-HT interactions of the best-ranked pose and pose taken from Celik et al., 2007	- 41 -
Figure 22: Comparison of MOE and Schrödinger output leading poses	- 41 -
Figure 23: Orientation of Serotonin Poses around Asp98.....	- 42 -
Figure 24: Leading Position of Poses 1 and 2.....	- 42 -
Figure 25: Superposition of MDMA Cluster 1.....	- 43 -
Figure 26: Interactions of the two best-ranked MDMA poses.....	- 44 -

Figure 27: Superposition of pose 1 with rigid docking leading pose	- 44 -
Figure 28: Position of the Leading Poses in the Binding Site.....	- 45 -
Figure 29: MDMA (left) and Serotonin (right) pose distance maps	- 45 -
Figure 30: Formation of Methylphenidate Cluster 1	- 46 -
Figure 31: Overlay of leading MDMA and MPD poses	- 47 -
Figure 32: Best calculated methylphenidate interactions.....	- 47 -
Figure 33: Benzatropine distance map	- 48 -
Figure 34: Superposed Benzatropine Poses (Discarded Pose in Red)	- 49 -
Figure 35: Best-ranked Benzatropine interactions with the hSERT.....	- 50 -
Figure 36: Residues Involved in Benzatropine Binding.....	- 50 -
Figure 37: Superposed cocaine poses.....	- 51 -
Figure 38: Distance Map of the Cocaine Poses.....	- 51 -
Figure 39: Overlay of cocaine and β -CFT	- 52 -
Figure 40: Proposed cocaine interactions	- 53 -
Figure 41: CFT Poses of the XP Run and Additional SP CFT Pose (Orange) in Superposition ...	- 53 -
Figure 42: Computationally preferred β -CFT binding pattern.....	- 54 -
Figure 43: β -CIT in the binding pocket and in superposition with β -CFT	- 55 -
Figure 44: Secondary binding mode of β -CIT.....	- 56 -
Figure 45: Different energy contributions for CIT poses	- 56 -
Figure 46: Antipodal Orientation of β -CFT Poses	- 57 -

I strived to locate all owners of picture copyrights and to get their acceptance for the use of their pictures in this thesis. In case that anyway a copyright violation is claimed, I request being informed.

Abstract

The selective reuptake of serotonin (5-hydroxytryptamine, 5-HT) from the synaptic cleft back into the nerve cell is performed by the presynaptic membrane-embedded serotonin transporter (SERT), thus terminating neurotransmission. An experimentally resolved structure of the SERT was not yet published, but the determination of binding mode hypotheses between ligands and a protein binding site with missing high-resolution structural data could be realized by creating a homology model of the protein of interest. As the SERT belongs to the neurotransmitter sodium symporter (NSS) or solute carrier 6 family (SLC6), its sequence could be aligned with a bacterial analogue, the prokaryotic leucine transporter LeuT. A detailed alignment of the SERT with the LeuT (PDB Code 2QB4) is available. Despite relatively poor sequence identity, it offers a highly conserved binding site, making homology modeling possible. For obtaining accurate information on binding interactions, flexibility both of ligand and target had to be taken into account, as it was done by a recently introduced efficient docking method, called Induced Fit Docking. The software necessary for applying the method was embedded in the Schrödinger molecular modeling software package by Schrödinger, LLC. It granted the binding site more flexibility than other docking programs and allowed to perform most of the working steps from the drawing of the ligands to the final evaluation steps within one software package. The ligand library consisting of the natural ligand serotonin and a selection of drugs that are known to bind into the SERT was docked into the inner serotonin binding site. After the final homology model had been created and the correct parameter settings were identified the docking experiments in two precision modes were performed. The output data was sorted according to calculated energy values and examined for its clustering behavior, leading to the selection of favored poses and therefore interaction pattern suggestions. The significance of the results was strived to be assured by several validation steps. The method itself proved its ability to retrieve experimentally resolved binding modes in a re-docking experiment with the template protein-ligand complex. For the docking experiments the selected binding modes were compared with published data. The results for serotonin were consistent with the data available, indicating similar quality for the ligands which yet lack information about their binding mode. The compounds based on the tropane scaffold were docked in a mode different from all other compounds, which needs further evaluation. However, as far as it can be judged by hitherto yielded data, the results seem reliable, and it will be just a matter of time until Induced Fit docking is implemented in most software modules.

Zusammenfassung

Die selektive Wiederaufnahme von Serotonin (5-Hydroxytryptamin, 5-HT) aus dem synaptischen Spalt erfolgt durch den präsynaptischen membrangebundenen Serotonintransporter (SERT), was zur Beendigung der Neurotransmission führt. Bisher wurde noch keine experimentell ermittelte 3D-Struktur des SERT publiziert, trotzdem können Studien über den Bindungsmodus zwischen Liganden und ihren korrespondierenden Bindungsstellen in diesem Protein realisiert werden, indem man ein Homologiemodell erstellt. Da der SERT zur Neurotransmitter-Natrium Symporter (NSS) oder Solute Carrier 6 (SLC 6) Familie gehört, kann seine Primärsequenz mit jener eines bakteriellen Analogons (LeuT, PDB Code 2QB4) übereinandergelegt werden, wofür bereits ein detaillierter Plan publiziert wurde. Trotz der verhältnismäßig geringen Sequenzidentität bietet das Protein eine gut konservierte Bindungstasche, was das Generieren eines Homologiemodells möglich macht. Um genaue Informationen über Bindungsmodi zu erhalten, muss die Flexibilität sowohl des Liganden als auch des Zielproteins berücksichtigt werden, so wie es bei einer kürzlich vorgestellten effizienten Docking Methode namens 'Induced Fit Docking' der Fall ist. Die dazu benötigte Software ist in das Schrödinger Suite Programmpaket von Schrödinger, LLC eingebettet. Die Ligandenbibliothek bestand aus dem natürlichen Substrat Serotonin sowie einer Auswahl von Verbindungen mit Affinität zum SERT. Nach der Fertigstellung des Homologiemodells und dem Erarbeiten der korrekten Parameter wurden die Docking Experimente in der inneren Bindungstasche des Proteins in zwei unterschiedlichen Präzisionsmodi durchgeführt. Die dabei erzeugten Daten wurden nach Bindungsenergie und ihrem Clusterverhalten sortiert, was zur Auswahl favorisierter Posen und verschiedenen Vorschlägen möglicher Interaktionsmuster führte. Dabei wurde versucht, die Signifikanz der Ergebnisse durch mehrere Validierungsschritte sicherzustellen. Die Fähigkeit der Methode zum Wiederauffinden experimentell ermittelter Bindungsdaten zeigte sich in einer Redocking Prozedur, bei den Experimenten wurden die ausgewählten Bindungsmuster mit der Literatur verglichen. Die Ergebnisse für Serotonin waren konsistent mit publizierten Daten, was eine ähnliche Qualität für Liganden mit bisher unbekanntem genauen Bindungsmodus verspricht. Die Substanzklasse der Tropanderivate erwies sich als problematisch und bedarf weiterer Untersuchungen. Dennoch scheinen die Ergebnisse, soweit man es bis jetzt beurteilen kann, verlässlich zu sein, und es dürfte nur eine Frage der Zeit sein, bis Induced Fit Docking in den meisten Softwarepaketen implementiert sein wird.

Curriculum Vitae

Persönliche Daten:

Name	Andreas Jurik
Geburtsdatum	15. Oktober 1983
Familienstand	ledig

Ausbildung:

seit Oktober 2008:	Diplomarbeit am Department f. Pharm. Chemie
September 2005:	1. Diplomprüfung
seit Oktober 2003:	Studium der Pharmazie an der Universität Wien
1993 – 2001:	Albertus Magnus Gymnasium der Marianisten, Wien XVIII Matura mit ausgezeichnetem Erfolg im Juni 2001
1989 – 1993:	Volksschule Wehlistraße, Wien II

Arbeitstätigkeit:

Oktober- Dezember 2009:	geringfügig beschäftigt, Universität Wien, Wiss. Personal
SS 09, WS 09/10, SS 10:	Tutor am Department f. Pharm. Chemie, bzw. FH Campus Wien
Februar- Mai 2008:	geringfügig beschäftigt, Jobs Personalberatung GmbH, 1090 Wien
Juli 2006:	Ferialpraxis Nordrand Apotheke, 1210 Wien
seit Juni 2002:	Teilzeitarbeit Ordination Dr. Karl Jurik, Ordinationshilfe
Juli/August 1999:	Ferialpraxis Ordination Dr. Karl Jurik, Ordinationshilfe
Juli 1998:	Ferialpraxis DonautankschiffahrtsgesmbH, 1020 Wien

Sonstiges:

2004-2009:	Doppelter Staatsmeister, Vierfacher Vizemeister in Rollstuhlrugby, Mitglied der Nationalmannschaft, Teilnahme an 2 Europameisterschaften und Sieg bei mehreren internationalen Turnieren
Juli 2002-Mai 2003:	Aufenthalt im RZ Weißer Hof der AUVA
Juli 2002:	Querschnittlähmung C VI/VII nach Badeunfall
Oktober 2001- Mai 2002:	Ableisten des Präsenzdienstes, Garderegiment